



**Proceedings of the 29<sup>th</sup>  
UK Academy for  
Information Systems  
(UKAIS)  
International  
Conference**

**University of Kent  
25<sup>th</sup>-26<sup>th</sup> April 2024**

**ISBN: 978-1-7390875-1-7**

## Preface

### **Proceedings of the 29<sup>th</sup> UK Academy for Information Systems (UKAIS) International Conference**

**ISBN: 978-1-7390875-1-7**

On behalf of the UKAIS and its committee, we welcome you to the conference proceedings for UKAIS 2024. This volume contains the papers presented at UKAIS2024: UK Academy for Information Systems, Annual International Conference held on April 25-26, 2024. The conference has 11 tracks and includes 4 Keynote presentations – Professor Kieran Conboy, Professor of Business Information Systems in the School of Business & Economics, University of Galway and co-Principal Investigator in the Lero Irish Software research centre; Dr Charles Knight, Assistant Director at Advance HE; Ozgur Savas, Vice President, Global Solutions Consulting, Zscaler; Professor Marina Jirotko, University of Oxford.

The UKAIS conference is the premier academic event in the Information Systems calendar within the UK and attracts leading scholars from the UK and overseas. It is a charity, whose aims are to enhance the recognition and knowledge of IS within the UK, and to provide a forum for discussing issues in IS teaching and research. UKAIS recognises the importance of including practitioners in its work.

The UK Academy for Information Systems was established in 1994 to foster a better understanding of the Information Systems field within the UK. We provide a forum for discussing issues in IS teaching and research and lobby professional/policy bodies on behalf of our field, such as the Office for Students, UKRI/Research England, UK business and UK Government. There is a conference every year, which is usually preceded by a PhD consortium. UKAIS Aims:

- To promote a better knowledge and understanding of information systems within the United Kingdom.
- To improve the practice of information systems teaching and research.
- To enable successful knowledge transfer of IS research into teaching and practice in order to provide a positive economic and societal impact.

Many thanks to all those that have given of their time so freely to review papers for the academy, it is much appreciated. Also a huge thanks to our conference administrators, Abi Hopkins and Emma Pearson, who really do keep all in check and on track and make the conference happen.

Thanks to everyone that has made this happen, the UKAIS Board and all the Track Chairs and Reviewers, we thank you all.

April 2024

The UKAIS 2024 conference chairs

## Organising Team

### Conference Chairs:

Laurence Brooks (University of Sheffield), Spyros Angelopoulos (Durham University), Yu-Chun Pan (Northeastern University), Efpraxia Zamani (Durham University)

### Conference Administrator:

Jan Williams, University of Kent

### Programme Committee

Nisreen	Ameen	Royal Holloway, University of London
Spyros	Angelopoulos	University of Durham
Laurence	Brooks	University of Sheffield
Ana Isabel	Canhoto	University of Sussex Business School
Dinara	Davlembayeva	Cardiff University
Guy	Fitzgerald	University of Loughborough
Emma	Forsgren	Leeds University
Stefanie	Gante	Durham Business School
Abigail	Green	Coventry University
Emma	Gritt	Leeds University
Najmeh	Hafezieh	Royal Holloway University London
Abi	Hopkins	University of Wolverhampton
Maddy	Hunt	Durham Business School
Andrea	Jimenez	University of Sheffield
Oliver	Kayas	Liverpool John Moores University
Maria	Kutar	Information Systems, Organisations and Society (ISOS) Research Centre, University of Salford
Yuanyuan	Lai	Royal Holloway University of London
Joyce	Lee	National Chengchi University
Shuyang	Li	Associate Professor in Business Analytics
Patrick	Mikalef	NTNU
Peiyu	Pai	National Taiwan University
Yu-Chun	Pan	Northeastern University London
Niki	Panteli	Lancaster University
Ilias	Pappas	University of Agder
Arisa	Shollo	Copenhagen Business School
Mylene	Struijk	The University of Sydney
Chekfoung	Tan	University College London
Eleni	Tzouramani	University of the West of Scotland
Sara	Vannini	University of Sheffield
Xenia	Vassilakopoulou	University of Agder
Will	Venters	London School of Economics and Political Science
Yichuan	Wang	University of Sheffield
Efpraxia	Zamani	University of Durham
Jun	Zhang	University of Sheffield

## Table of Contents

Number	Title	Authors
4	Developing the foundations for an inclusive IS education – a case method approach for promoting female digital leaders	Niki Panteli, Ling Xiao and Lucy Gill-Simmen
6	Emotional regulation for improved student learning: can smartwatches be utilized to navigate an increasingly digital student context?	Fay Giaver, Jostein Engesmo and Niki Panteli
7	Artificial Intelligence and Blockchain Technology	Patrick Buckley
11	Driving Student Success through a Data-Driven Approach in Higher Education	Yun Chen and Kate Han
12	Business models for digital sustainability: The case of Fintech startups	Su Hua Ou and Yuliani Suseno
15	Digital transformation in the public sector: Critical factors of Digital Justice Portal	Fotis Kitsios and Maria Kamariotou
16	On the Temporal Experiences of Remote Workers	Efpraxia Zamani
17	From Connections to Classrooms: Utilizing Social Media platforms for Enhanced Engagement: The case of LinkedIn.	Hai Nguyen, Micheal Parsons and Minh Ngoc Do
19	Impact of WhatsApp Groups on Teaching and Learning in Higher Education: A Review	Courage Matobobo and Godwin Pedzisa Dzvapatsva
21	The Role of Middle Managers in Steering Digital Transformation: A Dynamic Capabilities Perspective	Po-Wen Chen and Joyce Lee
23	Fostering inclusion in open strategy: three IT-enabled mechanisms	Liucen Pan, Mo Moeini and Joao Baptista
24	Uncovering the Landscape of Ethical Artificial Intelligence	Ito Abrham, Ruilin Zhu and Mahsa Honary
26	What are the Drivers of Non-Fungible Tokens (NFTs) Community Engagement?	Fulya Acikgoz, Abdelsalam Busalim and Ali Tarhini
27	Self-Determination Theory in Crowdfunding: Data from Vietnam	Tommi Tapanainen, Quoc Dat Trinh and Hai Nguyen
28	The Power of Streamers' Speech: A Signalling Approach to Live Streaming Commerce	Hua-Xuan Lin, Joyce Yi-Hui Lee, Chih-Hong Peng and Yi-Ling Lin
29	Responsible Digital Innovation with Micro-Businesses: A Resource Orchestration Perspective	Rayen Jui-Yen Chang and Niki Panteli
30	Mapping and Visualising the Digital Economy in The Context of Developing Countries: A Bibliometric Analysis	Ubongabasi Kingsley Omon, Gordon Fletcher and Mohammed Ali
31	The Role of Augmented Reality in Enhancing Sustainable Decision-Making in the Fashion Industry	Omon Fagbamigbe, Muhammad Awais Shakir and Zeeshan Bhatti
32	An Exploration of Contemporary Factors that Influence Adoption of Instant Messaging by Academics in Higher Education	Jonathan Jackson, Nicholas Day and Nataliya Mogles
34	The Technological Challenges of Blended Learning in Higher Education: A case study of Blackboard	Mohamed Daud, Ali Gheitasy, Parisa Saadati and Laden Husamaldin



35	Digital Ethics: Resolving “Wicked” Problems and Dilemmas	Ciara Heavin and Yvonne O'Connor
36	Online Health Communities for Parents of Children with Autism Spectrum Disorder	Pin-Ran Wang, Joyce Lee and Shu-Fen Tseng
37	Sharing of personal health information for secondary use: A scoping review from the perspective of trust	Georgios Tsirozidis, Sune Müller, Keld Pedersen and Jody Platt
38	Project Materiality and Integrated Analytics in Sustainability Augmented Portfolio Management	Javed Mohammad, Laden Husamaldin and Ikram Rehman
40	A soft ecosystems methodology of digital innovation through a case study of the insurance industry's response to connected cars	Anushri Gupta, Roser Pujadas and Will Venters
45	Internalization of Digital Technologies: Adapting to Organizational Inertia	Chuying Ding and Spyros Angelopoulos
46	The Delegation Paradox: A Study on How Uncertainty Impacts Experts' Willingness to Delegate Decisions to Algorithmic Solutions	Stefanie Gante and Spyros Angelopoulos
48	The impact of the digitalisation of care on older, unpaid carers	Anastasia Rousaki, Laura Sbaffi, Efraxia Zamani, Kate Hamblin and Rachael Black
49	Understanding Artificial Intelligence For Data With High Level Of Abstraction: Beyond Pixel Importance	Luca Heising, Timo Scheidel and Carol Ou
54	Mind the gap: using threshold concept theory to advance blockchain education	Trevor Clohessy, Graham Heaslip, George Onofrei and Lory Kehoe
58	Teachers' Perceptions of Students' Use of Generative AI in Summative Assessments at Higher Education Institutions: An Exploratory Study	Chekfoung Tan, Muna M. Alhammad and Marta Stelmaszak Rosa
59	Understanding Consumers' Reactance of Technology-Enabled Personalization: Perspectives on Stress and Fatigue	Peiyu Pai, Chiaodeng Fei and Tzuyu Chen
60	Responsible digital: co-creating safe, secure and wise digital interventions with vulnerable groups	Maria Rosa Lorini, G. Harindranath and Tim Unwin
65	Another tool for the toolbox? How strategists are adopting social media for sensemaking	Tiantian Qin and Josh Morton
68	Operationalizing Algorithmic Fairness: Empirical Study and Sociotechnical Framework Proposal	Fredrik Wang, Ilias Pappas and Polyxeni Vassilakopoulou
69	Factors influencing the career decisions of women software entrepreneurs: perspectives from India and Ireland	Aditi Singh, Brian O'Flaherty and Ciara Heavin
71	Smart maintenance at offshore wind farms: A digital System of Systems approach	David Wodak and Carol Ou
72	Data-driven Sustainable Cities in the UK: The Case of Public Space and Social Cohesion	Jie Qi, Suvodeep Mazumdar and Ana Vasconcelos
73	Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs	Maria Kutar, Marie Griffiths, Subrahmaniam Krishnan-Harihara, Tony Syme, Aadya Bahl and Toluwanimi Ojutiku
76	Apparel Matters? A New First-Impression Static Appearance Dataset	Jialou Wang, Honglei Li and Wai Lok Woo

77	A Conceptual Framework and Design Principles for Decision Support in Clinical Practice: Managing Knowledge and Uncertainty	Jishnu Das, Geir Inge Hausvik and Carl Erik Moe
78	Temporal Aspects in Cancer Clinical Decision-Making Through Machine Learning	Yiyu Wang, Umair UI Hassan, Anastasia Griva and Kieran Conboy
80	Teaching and Learning through the pandemic; the effects of remote work on women academics	Maria Vardaki
82	Responsible AI Principles: Findings from an Empirical Study on Practitioners' Perceptions	Pouria Akbarighatar, Ilias O. Pappas, Polyxeni Vassilakopoulou and Sandeep Puro
84	Design and Development of a Digital Transformation Canvas for SMEs in Developing Countries: a case study of Oman	Yazan Sammour, Yun Chen and Marie Griffiths
85	Cybersecurity Investments Metrics using FAIR-ROSI	Ying He, Tong Xin and Cunjin Luo
89	Digital inclusion network building: a network weaving analysis	Sharon Wagg, Sara Vannini, Efraxia Zamani, Maira Klyshbekova and Bethany Aylward
90	Conceptualising the Artificial Intelligence Divide: A Systematic Literature Review and Research Agenda	Aishatu Lawan Mohammed, Mohammed Ali and Maria Kutar
95	The Underlying Practices of Digital Transformation Leadership: Theorising the Practitioner Voice	Pat McCarthy, David Sammon and Ibrahim Alhassan
96	Generative AI in the Workplace: A Dive into Employees' Acceptance in China	Lei Yuan
98	Moving Action Learning Sets Online: Reflecting on Privacy, Intersectionality and Group Failure	Eleni Tzouramani
99	Leveraging Generative AI in Information Systems Development	Galal Galal-Edeen and Hanan Moussa
100	Virtual Representations as Boundary Objects: a Case of Building Information Modeling	Jing Wang and Linhao Fang
101	Becoming Analytical Champions: A Simple Hermeneutics-inspired Learning Analytics Model for Modules	David Sammon and Tadhg Nagle
107	A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach (TEAeM)	Laurence Brooks, Sara Cannizzarro, Nitika Bhalla and Kathleen Richardson
110	Generative AI as a Tool for Thematic Analysis: An Exploratory Study with ChatGPT	Sian Joel-Edgar and Yu-Chun Pan
111	Media Choice for Multi-motive Communication: Impersonal and Interpersonal Motives in Social Worker-Client Relationships	Linhao Fang, Aleksandra Irnazarow, Lynda Song, Kitty Yuen-Han Mo and Johnson Chun-Sing Cheung
112	Unpacking the 'Black Boxes' of Analytics: A Data Value Map Analysis	Kieran O'Driscoll, David Sammon and Tadhg Nagle



# DEVELOPING THE FOUNDATIONS FOR AN INCLUSIVE IS EDUCATION – A CASE METHOD APPROACH FOR PROMOTING FEMALE DIGITAL LEADERS

Niki Panteli, Lancaster University Management School, UK [n.panteli1@lancaster.ac.uk](mailto:n.panteli1@lancaster.ac.uk)

Ling Xiao, Royal Holloway University of London, UK [Ling.xiao@rhul.ac.uk](mailto:Ling.xiao@rhul.ac.uk)

Lucy Gill-Simmen, Royal Holloway University of London, UK [Lucy.gill-simmen@rhul.ac.uk](mailto:Lucy.gill-simmen@rhul.ac.uk)

## Abstract

*In this paper we draw attention to the need to develop inclusive pedagogy within Information Systems (IS) education. In particular, this paper aims to examine the effectiveness of the case method as an inclusive pedagogy method driven by an interest to increase gender representation and promote female students' enthusiasm in developing careers in the Information Technology (IT) sector and digital leadership in particular. We provide empirical evidence to support the role of case method pedagogy in achieving inclusive education. We report on the practice of case writing and teaching, and on focus group results with study participants to gain an understanding of students' experiences and views on a specific inclusive case. Our findings suggest that the case method could foster students' social and academic belonging through a collaborative, interactive and comfortable teaching environment. Accordingly, we posit that the case method contributes to developing inclusive teaching practices within the IS education and to paving the way for more females taking on digital leadership positions.*

**Key words:** *inclusive pedagogy, IS education, female IT professionals, female IT students, case method, digital leadership*

## 1.0 Introduction

Within the Higher Education (HE) sector there has been an increased emphasis on inclusive pedagogy, with inclusion and diversity being highlighted as strategic priorities across UK universities (Yeager and Walton, 2011). According to Florian and Black-Hawkins (2011), inclusive pedagogy ensures that all students, regardless of gender, race, religion and ethnicity participate equally in classroom life and experience rich learning opportunities. With such practices, students cannot only continue and successfully complete their studies, but also integrate their knowledge into the labour market improving their chances for employability and successful career development (Morina and Biagiotti, 2022). Therefore, calls have been made by government and industry bodies for universities to widen participation so that students from diverse backgrounds, gender and ethnicities are encouraged to join university programmes to improve their career and employability prospects (International Bureau of Education, 2016).

Nevertheless, despite these calls, programmes in Information Systems (IS) related subjects have lacked student diversity, and continued to be characterised as white and male-dominated (Lang, Freeman et al. 2022). In their editorial of a Special Issue on Equality, Diversity and Inclusion in IS education, Lang, Freeman et al. (2022) identify that IS education suffers from inequality and barriers due to gender, ethnicity, disability, sexuality, and socio-economic status. This paper is driven by an interest to overcome some of these barriers and to identify effective inclusive pedagogical practices in the field of IS education within HE.

The position we take is that it is not only universities and admission teams that have a responsibility for inclusive HE education, nor the tech industry and employer organisations more broadly - though there is no doubt that these play a crucial role in the recruitment and retention of diverse workforce – but it also comes down to the classroom educators. In particular, we posit that teaching material should be adapted to become more inclusive. As such, educators can play a significant role in promoting inclusivity with the purpose of enhancing the career and employability prospects of students from diverse backgrounds, including female students.

## **2.0 Tech workforce – need for diversity**

IT is a growing sector and central to economic growth, with a recent report referring to the industry contributing more than £82bn to the UK economy (McDonald, 2023). Despite its growth the same report argues that the diversity landscape of the industry has not improved. Existing literature has described this industry as ‘female-unfriendly’ due to the long working hours and masculine culture of work, causing many women to decide not to enter or to leave the profession (von Hellens et al., 2011). In addition to the persistent and ongoing under-representation of women in IT, academic research has also drawn attention to gender inequality in IT employment. For example, although women are employed throughout the different categories of jobs in computing, they are increasingly concentrated in areas of work that are low in status, power and rewards. As women move up the career hierarchy, their representation shrinks (Kirton and Roberston, 2018). According to the 2021 Harvey Nash Tech Survey, gender representation in this industry continues at a snail’s pace, with only 12% of digital leaders being women.

Despite government and industry efforts to increase participation of women in the IT sector, participation remains low. Structural and cultural factors have been identified as barriers to

women's development in this sector. Structural barriers include factors related to organizations, such as working schemes, whereas cultural barriers refer to factors which are imposed by the social environment such as gender discrimination (Ahuja, 2002). As such, the IT workplace is not gender neutral, with reasons ranging from inflexible work practices to lack of recognition of the value of a gender diverse workforce (Panteli et al., 2001). This ultimately contributes to women being dissatisfied with the monetary and promotion opportunities they receive (Kowal and Roztocki, 2016).

It is within this sectoral context of low representation of women, low retention rates and limited career development opportunities that we position the present study on inclusive pedagogy in IS education using the case method.

### **3.0 Case Method for Inclusive Education**

The case method, the teaching of problem-based scenarios (of varied length and detail), has become highly popular within management education. Research on the use of the case method argued that its significance lies in the fact that the case presented becomes a proxy for experience and therefore prepares students for employability (Pegg, et al. 2012), therefore making the learning gained from case discussion purposeful. Cases can be used in HE education to provide opportunities for decision making and conflict resolution, in (real or hypothetical) business scenarios. Research has found that the case method is the most effective teaching method in terms of self-awareness when comparing it to simulation exercises and lectures (Farashahi and Tadjehin, 2018). Evidence exists of the collaborative potentials of the case method and opportunities for enhancing learning (Khosa and Volet, 2013), whilst existing literature shows that the case method stimulates class discussion, enables students to develop critical ability and critical thinking skills and prepares them for the ambiguous and complex organisational world (Booth et al, 2000). With the use of authentic assessment high on the agenda for many business educators (Montano et al, 2023), the case method lends itself particularly well to authenticity as a tool for teaching and assessment.

It is important noting that educators play an important role in the choice of cases they adopt to support their teaching. By choosing case authors of different genders and of diverse backgrounds, as well as by using case studies that represent and showcase a variety of regions, sectors and scholarly perspectives that might be otherwise under-represented in the discipline, the case method could serve as an inclusive pedagogy (Sanger, 2020). To feel as if they belong,

students must be able to envision the value of taught content in their own lives and in their own future lives (Yeager and Walton, 2011).

Further to these, and with our view that inclusive education should be an embedded component of IS education and not just an afterthought, we posit that when considering using the case method in their teaching, educators should adapt the case from a content perspective to make it inclusive.

## **4.0 Study Design**

For the purpose of the current paper, we present how the case method was used to showcase inclusive education in IS education and some of the results that show its effectiveness. Below, we explain the case selection and how this was adapted to inclusive pedagogy.

### **4.1 Case (re)-writing**

As a way for promoting inclusive IS education, a case study that was written by the first author for use in a session on ‘Digital Transformation’ was revisited and parts of it re-written. The case was initially written following a set of primary data that was collected in a specific organisation during the pre- and post-Covid-19 pandemic. The case itself presented the experiences of a female chief digital officer (CDO) in a professional services organisation and the challenges she experiences in getting recognition of the strategic importance of her role in the organisation. The case on the topic of ‘Leading Digital Transformation’ sought to explore the challenges that digital leaders may experience in their efforts to promote digital transformation in their organisation. It was written with the aim to raise awareness among IS students that digital transformation may not always have a strategic orientation (despite existing research studies pointing to the significance of this) and the challenges that this may have to those leading digital transformation in the organisation and to the organisation as a whole.

Following discussions with and specific recommendations made by the second and third authors, two particular aspects of the case were modified: First, the female protagonist was presented as belonging to an ethnic minority group; as such her name was given an Indian origin, with her undergraduate education completed in her native country before continuing her postgraduate studies and beginning employment in the UK. Second, alternative employment options other than the traditional permanent and full-time posts for leaders within the IT

profession was introduced in the case. The case was re-written as a way for encouraging students to rethink of alternative, yet fulfilling careers (e.g. independent contractors, see Panteli and Urquhart, 2021) especially in male-dominated sectors such as IT.

#### **4.2 Applied Case Method Teaching and Assessment**

We invited students to take part in specially designed case method sessions which took place in January 2023. There were 40 students who were split into two groups. In recognition of their time, we offered £30 Amazon voucher to each student. Participants consisted of both undergraduate students and postgraduate students, though the vast majority belonged to the former category. Students were primarily from the discipline of business and management but there was representation from other disciplines too including psychology, computer science, philosophy and law.

At the end of each session, students were asked to take a short quiz testing their understanding of the subject discussed. In order to gain greater insights into students' experience in the case method teaching participants were also invited to take part in focus groups which took place a week later. They were solicited through a volunteer sign-up sheet. In total, there were 27 students who participated, 14 male and 13 female students, 24 undergraduate students and 3 postgraduate students. The group was also widely diverse representing different ethnicities: Asian (2); British White (3); British Indian (4); Black (2); Chinese (1); Indian (7); White/White other (6) and Mixed Other (2).

The focus groups contributed to insights on the effectiveness of the case method for inclusive education and showcased the sense of belongingness that students felt in the case method session. Analysis of the focus group discussions pointed to evidence of two types of belongingness. First, the case method contributed to an increased sense of belongingness within the class itself, with students noting that they felt more actively involved with the group:

*“The case study did give a sense of belonging socially, because you are always sharing ideas, maybe others have the same idea as you or similar understanding”.*

*“I think the case study definitely harboured a better sense of belonging, because it gave everyone a chance to make their own contribution to the discussion and what was being talked*



*about. Whereas in the lecture, although I do think it was a bit more useful, it was just the lecture mostly speaking.”*

*“So being forced to talk to other people, we can kind of create a friendship and then with that familiarity you feel like you belong a bit more, instead of like constantly walking into place and being the only person who can't relate to anyone”*

The second type of belongingness related to the connection that students felt with the case scenario and in particular the case protagonist: A female student for example specifically talked about the fact that the case protagonist was female:

*“I think it was a bit more obviously relatable, like it's kind of you put yourself in their shoes and being a woman, it is nice to be like ‘okay’ this is a real life situation. I could be in that situation ... it could be me and what would I do and would that be the best choice to do or what not and how would it benefit me later on? So it was nice in that sense”.*

## **5.0 Implications**

The aim of this study has been to examine inclusive education within IS education by taking a focus on the case method teaching. By drawing on a case that was re-written with the purpose of promoting inclusive pedagogy and in particular to encourage more female students not only to consider IT as a career but also to cultivate in them an interest in taking on leadership roles in IT and influence in this way the shape and impact of digitalisation and digital transformation in organisations of different sizes and sectors. Through the re-writing of a case study on digital transformation with a female protagonist from an ethnic minority, our study suggests that the case method with its potential to make students develop a sense of belonging and a connection to the class and case scenario, female students can develop confidence, interest and passion in taking on roles and develop their career in ways that perhaps they had not considered previously.

Case method teaching enables students to learn with purpose, goals, and meanings. As such, the findings are of interest to IS educators contemplating using the case method to deliver IS, digital business and related teaching. Importantly, the study reveals how the main feature of inclusive education namely sense of belonging can be achieved during our daily teaching practice. This will extend our role as educators beyond just delivering contents, but also by

writing teaching material and teaching cases that showcase characteristics of inclusivity and diversity, such as reference to female protagonists and other case participants from diverse groups. On the contrary to the traditional lecture, the case method teaching applies a constructivist teaching method (Biggs, 2001) and better engages students with interactive discussion time. Furthermore, educators have more control in requesting students join pre-set allocated groups to complete the case study questions which can further facilitate group analysis and discussion.

The study opens up the agenda on inclusive education within IS education. It is our strong position that it is not enough to admit students into our programmes from diverse backgrounds, but we also need to make adaptations to our teaching material to encompass inclusivity and diversity and increase female representation in IT employment and leadership in particular. Authors of IS textbooks can also play a critical role in this process by adapting teaching content, cases and exercises to the characteristics of inclusive education. In doing so, we as educators should actively take on the role for promoting inclusive and responsible IS education.

Future use of the specific case presented in this paper could be accompanied by discussions on students' careers preferences following inclusive pedagogy in IS education and examine the impact that these practices have on students' employability. Furthermore, we encourage IS researchers to undertake pedagogical research to promote inclusivity and diversity in IS education. For example, among others, research is needed to strengthen the results of this study by examining students' performance in inclusive case method classes and contrast these to non-inclusive case classes.

### **Acknowledgements:**

*The work was supported by a grant awarded to the authors by Royal Holloway, University of London for research into inclusive education. The authors are grateful to the University for the financial support and to the students who took part in this study. The authors would like to thank two doctorate students Haiqa Mushtaq and Ibrahim Bin Mibrad for assisting with the focus groups and data analysis.*

## References

Ahuja, M. K., (2002), Women in the Information Technology Profession: A Literature Review, Synthesis and Research Agenda, *European Journal of Information Systems*, 11, 20-34.

Biggs, J. (2001). The reflective institution: Assuring and enhancing the quality of teaching and learning. *Higher education*, 41, 221-238.

Booth, C., Bowie, S., Jordan, J., & Rippin, A. (2000). The use of the case method in large and diverse undergraduate business programmes: problems and issues. *International Journal of Management Education*, 1(1), 62-75.

Farashahi, M., & Tajedin, M. (2018) Effectiveness of teaching methods in business education: A comparison study on the learning outcomes of lectures, case studies and simulations. *International Journal of Management in Education*, 16, 131e142. <https://doi.org/10.1016/j.ijme.2018.01.003>

Florian, L. & Black-Hawkins, K. (2011) Exploring inclusive pedagogy. *British educational research journal*, 37(5), pp.813-828. <https://doi.org/10.1080/01411926.2010.501096>.

Harvey Nash Tech (2021), Digital Leadership Report, [Nash Squared Digital Leadership Report 2021 | NashTech \(nashtechglobal.com\)](#) (accessed online on 12 September 2023)

International Bureau of Education. (2016). Reaching out to all learners: A resource pack for supporting inclusive education. Available online <https://unesdoc.unesco.org/ark:/48223/pf0000243279> (accessed 27 January, 2023).

Kirton, G. and Robertson M. (2018), Sustaining and advancing IT careers: Women's experiences in a UK-based IT company, *The Journal of Strategic Information Systems*, 27 (2), Pages 157-169, [doi.org/10.1016/j.jsis.2018.01.001](https://doi.org/10.1016/j.jsis.2018.01.001).

Khosa, D. K., & Volet, S. E. (2013). Promoting effective collaborative case-based learning at university: A metacognitive intervention. *Studies in Higher Education*, 38(6), 870-889.

Kowal, J. and Roztocki, N., (2016), Gender and job satisfaction of information technology professionals in Poland. In *System Sciences (HICSS)*, 2016 49th Hawaii International Conference on (. 3625-3634). IEEE.

Lang, Michael; Freeman, Mark; Kiely, Gaye; and Woszczynski, Amy B. (2022) "Special Issue Editorial: Equality, Diversity, and Inclusion in IS Education," *Journal of Information Systems Education*: Vol. 33 : Iss. 1 , 1-6.  
Available at: <https://aisel.aisnet.org/jise/vol33/iss1/1>

McDonald, C (2023), More than two million now work in UK tech, says CompTIA, *ComputerWeekly*, 7 September 2023, <https://www.computerweekly.com/news/366551292> (accessed September 8, 2023)

Montano, S., Gill-Simmen, L., Lee, D., Walsh, L., Duffy, D. and Newman, N., 2023. Assessing authentically—learnings from marketing educators. *Journal of Marketing Management*, pp.1-33.

Morina, A., & Biagiotti, G. (2022). Inclusion at university, transition to employment and employability of graduates with disabilities: A systematic review. *International Journal of Educational Development*, 93, 102647.

Panteli, N., Stack, J., & Ramsey, H. (2001). Gendered patterns in computing work in the late 1990s. *New technology, work and employment*, 16(1), 3-17.

Panteli, N., & Urquhart, C. (2022). Job crafting for female contractors in a male-dominated profession. *New Technology, Work and Employment*, 37(1), 102-123.

Pegg, A., Waldock, J., Hendy-Isaac, S., & Lawton, R. (2012) *Pedagogy for employability*. York, UK: Higher Education Academy. Available online <https://www.advance-he.ac.uk/knowledge-hub/pedagogy-employability-2012> (Accessed: 30 August 2023).

Sanger, C.S. (2020) Inclusive pedagogy and universal design approaches for diverse learning environments. In *Diversity and inclusion in global higher education* (pp. 31-71). Palgrave Macmillan, Singapore.

Von Hellens, L., Nielsen, S. H., & Beekhuyzen, J. (2004). An exploration of dualisms in female perceptions of IT work. *Journal of Information Technology Education: Research*, 3(1), 103-116.

Yeager, D.S. and Walton, G.M., 2011. Social-psychological interventions in education: They're not magic. *Review of educational Research*, 81(2), pp.267-301. <https://doi.org/10.3102/003465431140599>

# EMOTIONAL REGULATION FOR IMPROVED STUDENT LEARNING: CAN SMARTWATCHES BE UTILIZED TO NAVIGATE AN INCREASINGLY DIGITAL STUDENT CONTEXT?

**Authors**

## **Abstract**

*This paper reports from a qualitative intervention study of nine participants over a six-week period on the role of digital tools such as Smartwatches to enhance emotion regulation for improved learning and performance in a digital study context. We found that students initially had little awareness of their physiological sensations and emotions finding it hard to see links between their experiences and learning environments, as well as learning outcomes. However, aided by their Smartwatches they were able to see patterns in their daily life that allowed them to become more proactive in setting the ground for more productive learning sessions. This awareness also helped them plan and structure their study sessions, sometimes in conjunction with various other digital tools.*

**Keywords:** *Smartwatches, learning, emotions, emotion regulation*

## **1.0 Introduction**

Technologies such as smartphones and social media are sources of stress (Tarafdar et al., 2019) and distraction (Dontre 2020), but can also be utilized in innovative and creative ways to navigate these experiences in a learning context (Haleem et al., 2022). The current study explores the role of digital tools such as wearables (smartwatches) and applications (Microsoft Viva Insights) for emotion regulation to facilitate improved learning and performance in an increasingly digital learning context. We understand emotions to occur as the result of a sensemaking process where people use their past knowledge to make sense of their surroundings, as well as their bodily sensations in the present moment, to predict what is going to happen, and to act in appropriate ways (Barrett, 2016). However, emotions also need to be regulated when they are not helpful or appropriate in the situation in which they occur. We refer to emotion regulation as the ways in which individuals influence the types of emotions they experience; when they experience them; and how these emotions are expressed and experienced on an everyday basis (Gross, 1999). Thus, it is assumed that individuals regulate their emotions in ways that they find helpful depending upon their bodily sensations,

their subjective experiences, and the situation in which they find themselves. Ultimately emotion regulation is about controlling the physiology and expression of positive and negative emotions and is a way for individuals to relate to themselves and their emotions for improved wellbeing and functioning. Finally, there is a temporal element to emotion regulation. This implies that individuals have various pressure points to touch upon depending on where they are at in the emotion generation process as they evaluate what is good for them or not. Early in the emotion generation process individuals may actively choose the environment and situations they are exposed based on the emotions they anticipate they will experience, or they may attempt to change the situation they are in. Later in the emotion generation process, as they cannot necessarily choose or change the situation, individuals may still actively choose to shift their attention or adopt a certain mindset their interpretation of their bodily sensations and the environment or situation they find themselves. Finally, late in the day in the emotion generation process individuals are left with the strategy of modulating their emotional responses.

We know from the literature that emotions play a crucial, but neglected, role in the learning process (Ben-Eliyahu, 2019) and for performance (Ashkanasy, 2004; Ashkanasy & Dorris, 2017, Hökkä et. al., 2020). Emotions are an important antecedent to learning (e.g. motivation), a byproduct of learning (e.g. sense of achievement) and are evoked during the learning process (e.g. in the experience of flow during learning). Of relevance to learning there has been some studies on creativity in the organizational psychology domain. Staw and Barsade (1993, p. 304) did for instance put forward the “sadder-but-wiser versus happier-and-smarter” hypothesis. Here, they found that positive emotions, such as happiness, led to improved productivity, but did not necessarily lead to the best decisions. It has also been suggested that negative emotional experiences makes individuals less susceptible to bias and less likely to be swayed by persuasion (Forgas and George, 2001). To et al. (2015) similarly showed that negative emotional experiences can provide motivational and cognitive resources helpful for solving problems.

There are numerous studies on the role of emotional intelligence for performance in the academic learning context (for a review see e.g. Quílez-Robres, 2023). However, although emotional intelligence and emotion regulation are related constructs, the former refers to a set of dispositions and abilities, while the latter is a set of behaviours. It has therefore been suggested that it may be particularly important to cultivate emotion regulation, e.g. as a

targeted intervention, to help students navigate stress, temptations, distractions and performance anxiety and to facilitate motivation in the educational context (Jacobs and Gross, 2014). It has for instance been pointed to the importance of training students to adopt a malleable mindset for learning, focusing on social connection to decrease feelings of social threat in student groups. Furthermore, to facilitate reflection through expressive writing to improve test anxiety. However, we know from this literature that less is known about the role of technology in this domain. Furthermore, we need more research studies adopting technologies that measure physiological data in real time on the within-person level combined with self-report data. Digital technologies allow for a fine-grained assessment of contexts under which individuals are employing regulation strategies, the ways in which they employ or modify these strategies, and the effectiveness of these strategies in various situations and over time (Bettis et al., 2021). Studies from the work setting has pointed to the promising role of wearables to regulate stress (Tarafdar et al. 2019) and the utilization of the Microsoft Viva insights application to navigate the boundaries between work and non-work (John et al., 2022). At the same time organizations and mental health practitioners have been cautioned that digital tools, such as wearables, may be detrimental to performance and raise a number of ethical considerations and issues (Plester et al., 2022; Bettis et al., 2021; Maltseva, 2020).

Drawing on these studies, we aim to examine the role digital tools, focusing particularly on wearables in the form of Smartwatches, for emotion regulation with implications for learning and performance. To this effect, we present an intervention study where students were using digital tools to facilitate emotion regulation with the aim of helping them plan and organize learning situations.

## **2.0 Research Design and data collection**

In this study we utilized a qualitative intervention project to explore how students experienced the adoption of digital tools that aimed to support their emotion regulation. The project took place over a six week period between May and June 2023. Data was collected from workshops with students, group interviews and student diaries. In line with the literature it was important to design an intervention that stimulated (1) a sense of community in the group studied, (2) reflection among participants and (3) development on the intra-individual level over time.

Nine undergraduate information system (IS) students (20-25 years) were recruited from a Norwegian university. The study took place prior to and during the exam period after the teaching period had ended. Initially the students attended a kick-off workshop. Here, students were equipped with Smartwatches and informed about the project and their tasks and responsibilities. A Teams site was also set up to enable interaction with, and between, the students throughout the project period. Initially, students were asked to reflect around their user experiences in a study setting of their choice in open-ended diaries twice a week over a two-week period. Then, the students attended a planning workshop where they were allowed to share their experiences with their fellow students, and with the team of researchers, from the first two weeks of the project. Here, they also learned about the role of emotions and emotion regulation in the study setting and ways in which smartwatches could be utilized to support emotion regulation. In this workshop students also agreed on two emotion regulation strategies from Microsoft Viva Insights (mindfulness and body scan) they were to implement in the following four weeks. In addition, they were encouraged to test two additional strategies of their own choice (e.g. the Pomodoro technique for increased focus). This was followed by twice weekly diary study over a period of four weeks. Here students were asked to describe one study session and emotion regulation strategies implemented and to reflect around how they had experienced the activity while implementing these strategies. They were also probed to describe their physical and digital surroundings, their emotional experiences, and physiological data from their smartwatch before and after the session. Finally, students attended an evaluation workshop and focus group interviews. In total we collected 108 diaries and four post-workshop focus group interviews that were recorded and transcribed. We also recorded and transcribed parts of the discussion in the planning workshop, and took detailed observation notes after all three workshops. Finally, all participants submitted a reflection note after the project had ended, including a suggestion of five study habits that they would like to share with other students. Figure 1 below presents an outline of the data collection.



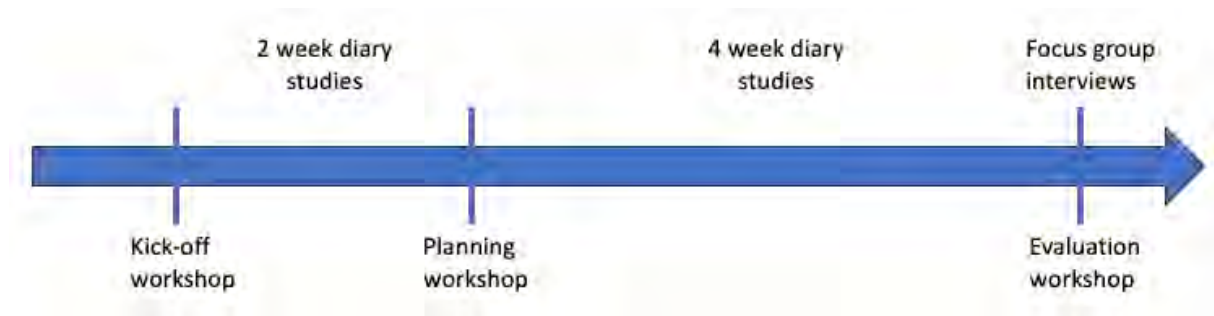


Figure 1. Project and data collection

## 2.1 Data analysis

The data was analyzed through theory led thematic analysis (Braun and Clarke, 2021). This meant that although we remained close to the data we specifically searched for instances of emotions and the ways in which access to physiological data through wearables informed emotion regulation behaviours. We also continued to return to the literature throughout the analytical process. Finally, we focused particularly on the ways in which emotion regulation developed on the intra-individual level over the course of the intervention period. The data was analysed in five steps: data familiarization; initial coding generation; generation of themes based on the initial coding; review of themes; and theme definition and labelling.

## 3.0 Preliminary findings

It was evident that initially students experienced having little awareness of their physiological sensations and emotional experiences. Furthermore, they found it hard to see the link between these sensations and experiences and their physical and digital learning environment, as well as the link to their learning outcomes:

*“I struggled reflecting about myself and my environment, both physical and digital“.*

Thus, the first part of the intervention was crucial to stimulate this awareness through experimenting with various digital tools and to establish the links between sensations and emotional experiences and the learning process:

*“I am starting to see that how I feel that day affects how much of what I read I am able to digest”.*

Over time, through reflecting on the links between their emotional experiences, emotion regulation and various learning experiences and outcomes they were able to see some patterns where they could intervene. For instance, they reported becoming much more proactive in setting the ground for productive learning sessions, e.g. through using their smartwatches to monitor their sleep and activity levels in conjunction to study activities. For instance, they reported that due to their increased awareness, they made sure that they were rested before important study sessions or postponed high-concentration activities if they were not feeling well and took breaks during study sessions when they were not properly rested. They also became much more proactive in using digital tools such as Pomodoro and Hold to avoid interruptions, to structure study sessions and/or to be reminded to stay focused and to take breaks.

Overall, our preliminary data analysis showed that students

- **Used wearables to promote holistic balance and to adjust to study situations.** What you do the day, evening and night before matters. Smartwatches helped students monitor sleep, the effects of alcohol, the impact of various food choices and exercise and see the links to physical sensations (e.g., feeling agitated and stressed) and emotional experiences (e.g., anxiety) in the moment.
- **Structured the digital environment and interruptions through various apps on smartphones/laptop.** In conjunction with information from the Smartwatches students structured their digital environment to fit with current physical and mental state. E.g., improved awareness allowed them to see more clearly the links between e.g., lack of sleep and poor concentration. Something in which in turn enabled them to adjust digital surroundings accordingly, to better fit with their needs in the situation, such as using the Hold app when feeling unfocused.
- **Planned and structured physical and social surroundings.** E.g., choosing a quiet study space when not being properly rested, or choosing to work with others when feeling more energized and in a good mood.
- **Planned and structured work sessions.** E.g., doing a mindfulness exercise or body scan when feeling stressed to prepare for a study session that required increased focus (e.g., prior to an exam).

## **4.0 Tentative Implications and Conclusions**

With the increasing use of wearables opportunities arise for informing improved emotion regulation, with important implications for learning and performance in the educational context. Our intervention study focused on students in Higher Education and sought to examine the impact of digital tools such as Smartwatches on students' learning outcomes through enhancing their emotion regulation skills. The study showed that the use of smartwatches can increase students' emotional awareness enabling them to adopt appropriate practices to maintain their focus on their learning and advance their studies. We recognize that the analysis is still underway and by the time of the conference we will have a more complete appreciation of the findings and the implications on students learning. We acknowledge that the sample size of our study is relatively small, but considering that this was an inductive qualitative study statistical generalization was not a goal. However, considering the amount and richness of the data we still believe that our findings are applicable to broader study population as well as other and similar contexts, such as employees in an organizational context.

## References

- Ashkanasy, N. M. (2004) Emotion and Performance, *Human Performance*, 17:2, 137-144.
- Ashkanasy, N. M. and Dorris, A. D. (2017). Emotions in the workplace, *Annual Review of Organizational Psychology and Organizational Behavior*, 4 (1), 67-90.
- Barrett, L. F. (2016) The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, 12 (1): 1-23.
- Ben-Eliyahu, A. (2019) Academic Emotional Learning: A Critical Component of Self-Regulated Learning in the Emotional Learning Cycle, *Educational Psychologist*, 54 (2), 84-105, <https://doi.org/10.1080/00461520.2019.1582345>
- Bettis A.H., Burke T.A., Nesi J., Liu R.T. (2022). Digital Technologies for Emotion-Regulation Assessment and Intervention: A Conceptual Review. *Clinical Psychological Science*, 10 (1), 3-26. <https://doi.org/10.1177/21677026211011982>
- Braun, V., & Clarke, V. (2021). *Thematic Analysis: A Practical Guide*. London: Sage.
- Dontre, A. J. (2021). The influence of technology on academic distraction: A review. *Human Behavior & Emerging Technologies*, 379-390.
- Gross, J. J. (1999). Emotion regulation: past, present, future, *Cognition and Emotion*, 13 (5): 551-573.
- Haleem, A., Javaid, M., Qadri, M. A. & Suman, R. (2022). Understanding the role of digital technologies in education: A review, *Sustainable Operations and Computers*, 3, 275-285, <https://doi.org/10.1016/j.susoc.2022.05.004>
- Hökkä, P., Vähäsantanen, K. & Paloniemi, S. (2020). Emotions in Learning at Work: a Literature Review. *Vocations and Learning* 13, 1–25.
- Jacobs, S. E., & Gross, J. J. (2014). Emotion regulation in education. *International handbook of emotions in education*, 183-217.
- John, B., Alsamarrai, Z. & Panteli, N. (2023). Enhancing Employee Experience in the Era of Hybrid Work: The Case of Microsoft Viva. *IEEE Software*, 40(2), 70–79. <https://doi.org/10.1109/ms.2022.3229956>
- Forgas, J. P. and George J. M. (2001). Affective influences on judgments and behavior in organizations: an information processing perspective. *Organizational Behaviour and Human Decision Processes*, 86, 3–34.

Maltseva, K. (2020). Wearables in the workplace: The brave new world of employee engagement, *Business Horizons*, 63 (4), 493-505.

Plester, B., Sayers, J., & Keen, C. (2022). Health and wellness but at what cost? Technology media justifications for wearable technology use in organizations. *Organization*. <https://doi.org/10.1177/13505084221115841>

Staw B. M. and Barsade S. G. (1993). Affect and managerial performance: a test of the sadder-but-wiser versus happier-and-smarter hypotheses. *Administrative Science Quarterly*, 38, 304–28

Tarafdar, M., Cooper, C. L. & Stich, J. (2019). The technostress trifecta - techno eustress, techno distress and design: Theoretical directions and an agenda for research. *Information Systems Journal*, 29(1), 6–42. <https://doi.org/10.1111/isj.12169>

To, M. L., Tse, H. M., Ashkanasy N. M. (2015). A multilevel model of transformational leadership, affect, and creative process behavior in teams. *The Leadership Quarterly*, 26, 543–56

Quílez-Robres, A., Usán, P., Lozano-Blasco, R., & Salavera, C. (2023). Emotional intelligence and academic performance: A systematic review and meta-analysis, *Thinking Skills and Creativity*, 49, <https://doi.org/10.1016/j.tsc.2023.101355>.

# Artificial Intelligence and Blockchain Technology

## Abstract

*Recent rapid developments in Artificial Intelligence (AI) have led many observers to believe we are on the cusp of a revolution, with AI poised to have an enormous impact upon societies and economies. However, many challenges must be met before AI can safely and fairly fulfil its potential. Blockchain is a set of inter-related, interconnected technologies that allow for the development of a range of socio-technical constructs such as data markets and prediction markets which have unique attributes and capabilities such as data immutability and designable anonymity and privacy. This research explores how these attributes and capabilities of these systems could be leveraged to address some of the challenges in AI development.*

**Keywords:** Artificial Intelligence, Blockchain, Prediction Markets, Data Markets

## Introduction

Artificial Intelligence (AI) is currently experiencing a massive boom in attention from both academics and practitioners. The confluence of a number of trends including increasingly powerful hardware, the increasing availability of data in digital formats and methodological breakthroughs in areas such as deep learning and reinforcement learning has led to the deployment of several eye-catching AI applications such as ChatGPT. These developments have led many observers to believe we are on the cusp of an AI revolution.

The excitement generated by recent developments notwithstanding, there are many fundamental challenges that must be addressed before AI can fulfil its touted potential. Some are technical in nature – for example, how can we manage and validate the vast amounts of data modern AI systems require? Others are social or political in nature – how can society integrate AI decision-making systems into our businesses and societies in a manner that is efficient, just and ethical? Such questions are intimidating in scope and will undoubtedly require contributions from a wide range of fields and disciplines.

Blockchain is a set of inter-related, interconnected technologies that allow for the development of a range of socio-technical constructs such as data markets and prediction markets which have unique attributes and capabilities such as data

immutability and designable anonymity and privacy. This research aims to explore how these attributes and capabilities could be leveraged to address some of the challenges in AI development.

## **Literature Review**

In this section, a brief overview of the two fields to be synthesized is provided. First, a high level introduction to AI is provided. The particular purpose of this section is to identify the key high level challenges that are faced by academics and practitioners seeking to design, develop and deploy AI systems. The second field surveyed is blockchain technology. In this section, focus is given to elucidating the specific capabilities offered by blockchain technology. This catalogue is used to analyse how blockchain technology can offer potential solutions to some of the key challenges slowing the progress of AI.

## **Artificial Intelligence**

The development of machines that can mimic or surpass human intelligence has been predicted since before the dawn of digital computing. Issac Asimov's description of sentient robots and the "Three Laws of Robotics" first appeared in print in 1942 (Asimov 1950). The Dartmouth Summer Research Project on Artificial Intelligence conference organized in 1956 is generally seen as the origin of academic AI research (Nilsson 2009). The field is notoriously volatile, with sentiment in the field oscillating between euphoric over-expectations (Crevier 1993) followed by "AI winters" of pessimistic retrenchment (Nilsson 2009).

Today, AI research is experiencing a period of sustained interest and optimism. New techniques such as genetic algorithms have been developed, while techniques with a longer pedigree such as neural networks have been reinvigorated by innovative approaches such as deep learning. Along with these theoretical advances, increasingly powerful computing platforms and the immense data sets generated by the internet have allowed AI powered systems to make advances in areas such as voice assistants and self-driving cars (Badue et al. 2021, Hoy 2018) Other fields where AI has had a significant impact include finance, medical decision support systems, recommender

systems, face recognition and machine translation (Marr 2019). AI is often described by the mainstream media as the dominant technology of the future.

Making specific forecast in such a dynamic field is notoriously difficult. A commonly selected target for such forecasts is that of an AI system demonstrating human-level general intelligence (Baum, Goertzel, and Goertzel 2011). Presenting an aggregated summary of several surveys of AI expert communities, Bostrom (2016) provides the following median estimates: A 10% probability of Artificial General Intelligence (AGI) by 2022, a 50% probability by 2040 and a 90% probability by 2075.

However, despite positive forecasts, there is still a vast degree of uncertainty about the future developmental trajectory and impact of AI (Mindell and Reynolds 2022). Some experts foresee AI systems being tasked with building even more advanced AI systems leading to a “Cambrian explosion” of intelligence (Muehlhauser and Salamon 2012). Forecasts of this nature often see AI systems with a level of intelligence comparable to human beings as being a temporary milestone along the road to systems which dramatically exceed the intellectual capacity of human beings (Bostrom 2016). However, such a perspective is far from universal (Mindell and Reynolds 2022). While acknowledging progress, they suggest the path to artificial intelligence may be far more difficult than cheerleaders suppose (Penrose and Gardner 2002). Some researchers believe that intelligence is fundamentally non-algorithmic and deterministic Turing machines will never be able to replicate intelligence (Penrose and Gardner 2002). Another, more philosophical issue is whether the concept of intelligence as an attribute associated with a singular entity is fundamentally flawed (Clark 2005). Instead, both consciousness and intelligence may be properties embedded in a larger cultural feedback loop. From this perspective, consciousness and intelligence are properties embedded in society and cannot be created absent a larger social context (Dennett 2017).

Uncertainty also dominates prognostications about the impact of AI on society. Optimists forecast that the impact of Artificial Intelligence to be positive (Kurzweil 2005). Cognitively superior AI will turbocharge the development of solutions to challenges such as resource depletion and climate change. AI systems and robots will perform the physical and cognitive tasks required to produce goods and services. Freed from the necessity of labour. Individual humans will have far more choice in



how they spend their time, be that in consuming entertainment, creative endeavours, or more traditional economically focused activities.

Pessimists also proffer potential futures where the development of AI has negative impacts. For example, some researchers suggest that an inferior intelligence will be unable to control either the capabilities or motivations of a superior one (Bostrom 2016). In the same way that, for example, a dog or cat is unable to even conceive of human motivations, the inferior intelligence of humans will be utterly unable to understand, much less control, AI's that advance beyond a certain level of cognitive capability. In this situation, some fear a future where humans become an endangered or extinct species (Joy 2000). Others fear the diminution and eventual destruction of human agency by the practical and philosophical superiority of AI systems (Harari 2016).

Even in a scenario where AI do not advance beyond a cognitive horizon that renders them beyond human control, pessimists raise serious concerns about the spread of AI (Arntz, Gregory, and Zierahn 2016). On the face of it, predictions that AI systems and robots will perform the majority of all of the labour required to meet human needs seem benign. However, even such an eventuality raises considerable questions. The decline in skills such as navigation and map-reading due to satellite navigation can be seen as an example of systems that start as question-answering "oracles" before evolving into authoritative "sovereigns", which can lead to learned helplessness (Bostrom 2016). More prosaically, in a context where economic activity is managed by AI systems, social and political power will reside with those who control the AI systems (Autor and Dorn 2013). An extrapolation of current trends which suggests increasing inequality and a future where societal power is vested in a small group of elite actors, while the majority of humanity has little or no real agency is not unreasonable (Harari 2016).

The above survey demonstrates briefly the difficulty of making accurate predictions with regard to AI, and the range of the possibilities that the development of the technology invokes. However, these challenges notwithstanding, the consensus in academic and wider society today is that AI will have a significant and increasing effect on societies and economies for the foreseeable future.

## **Challenges in Developing AI Systems and Models**

Building AI models and systems is a technically difficult task. Despite the enormous amounts of time, money and effort which is being expended by a wide variety of powerful and well financed actors, there are many outstanding challenges in developing and deploying AI. In this section, the literature is used to identify some of these challenges at a relatively abstract level. This catalogue will be subsequently used to structure an analysis of where the unique capabilities of blockchain based technologies offers potential solutions.

### **Data Quality and Quantity**

As such, the accuracy, effectiveness and efficiency of AI systems trained using approaches such as deep learning is determined to a large degree by the data that is used to train them. Accurate, reliable data allows these systems to learn efficiently. A common feature of most methodologies being used to develop AI systems today is that they require vast amounts of high quality data in order to allow them to be trained (Sun et al. 2017; Halevy, Norvig, and Pereira 2009).

This requirement for large amounts of high-quality data about the real world presents a number of major challenges (Busch 2014). First, obtaining and curating such data can be expensive and time-consuming. Second, biased, incorrect or incomplete data can lead to biased or inaccurate models (Zhang, Lemoine, and Mitchell 2018). AI models often inherit the biases present in their training data, which present significant challenges to their use, particularly in sensitive applications in the areas of finance, hiring and criminal justice. Unvetted training data can render AI models vulnerable to attacks where adversaries specifically craft the data presented during the training phase to cause the model to misclassify (Papernot et al. 2016).

Currently, most AI models are trained on datasets that have been scraped from the open Internet. This is because the Internet represents one of the few sources of datasets large enough to train LLM's and other AI models. However, to a large degree this data is posted by unknown or anonymous individuals whose knowledge and motives in posting are unknown and unknowable. The unvetted and unverified nature of the vast majority of this data raises significant challenges.

## **Regulation and Compliance**

As the capability, scale and influence of AI systems increases in the economy and society at large, so will the need for a regulatory and compliance regime that can oversee the design, development and deployment of these systems. Governments, NGO's and transnational regulatory bodies will certainly move to insist on standards and accountability, particularly when AI systems become deployed in sensitive areas such as transportation, health, industrial relations and justice. What these regulatory regimes will govern, who will design them and how they will be implemented are all open questions of enormous import both to the developers of AI systems and the societies they operate in.

## **Data and Model Distribution**

One of the major sources of anxiety with regard to the development of AI is the fear that because of the expense involved in training AI models such as ChatGPT or Google Bard, only a few actors will be able to access these models. As such, society at large has an interest in ensuring that the power associated with these models is not concentrated in the hands of few actors. Ensuring at least the possibility of widespread access to AI models involves ensuring the widespread distribution of at least two categories of data.

The first category is training data. Any actor seeking to create their own AI model will need access to large volumes of data to feed into an appropriate algorithm. Amassing this volume of training data is an expensive pursuit, both in terms of time and money. In addition, as mentioned previously, the volume of data by itself is insufficient unless consumers can be assured as to the accuracy and reliability of the data.

The second category of data that could be distributed is informational representation of AI models themselves. In this case, the widespread dissemination of AI technology is enabled by sharing the data that represents the trained model. Here, the need for computationally expensive training is removed, and the model can essentially be run as is.

In both cases, if the decision is made to make the AI widely available, then a significant question is how can we ensure the integrity of the data being shared. The distribution of large, verified and verifiable data sets has been a continuing challenge of the digital era, and is only rendered more pressing by the requirements brought about by the rise of AI (Philip Chen and Zhang 2014).

### **Continuous Learning and Adaption**

In general, AI models are trained with large data sets. These datasets can only contain information up until the point where they were created. This means that AI models can only “know” about information up to the point in time when their dataset was created. This limitation can be clearly seen in interactions with ChatGPT where asking about, for example, events which happened after 2021 will only produce “hallucinations” (Kumar et al. 2023). Actors wishing to maintain the efficiency of their AI models face a pressing need to continually correct, update and expand their training data sets. For models which capture and scrape training data from the Internet, this does not present a significant problem, since the Internet is constantly being added to. However, data captured from the Internet has other problems. Given the pseudo anonymous open nature of the public Internet, it seems unlikely that a situation where guarantees about the validity and reliability of data gathered from it can be made. If AI models are to be built using reliable data, it seems that some mechanism from capturing and validating data will be required. This will in turn require some mechanism for rewarding not only the creation of these datasets, but also specific incentives around for truthfully validating the data and penalties for problematic data.

### **Ethical Considerations and Oversight**

One of the biggest challenges to the widespread deployment of AI systems is significant social and political concerns about their use in automated decision-making. AI systems are already being used to make critical decisions about individuals that have ethical implications, in areas as diverse as healthcare diagnoses, autonomous vehicles, public surveillance and criminal sentencing. Moreover, most of these systems are notoriously opaque. The majority of systems, particularly those trained using deep learning methods, are essentially “black boxes” where virtually no information is provided as to how an AI system reached a particular conclusion.

Of course, it is important to note that oversight of decision processes and the accountability of decision makers are perpetual challenges, and our current systems which are largely dependent on individual humans are far from perfect. However, many of the approaches that we currently deploy to address these challenges in our current paradigms, for example mandated transparency or legal liability, seem to be ill-suited to being applied to AI systems. Ensuring ethical behaviour and accountability in AI is a complex and evolving challenge.

### **AI Interaction**

As AI systems become more powerful and ubiquitous, there will be an increasing demand for such systems to become autonomous. It is easy to imagine AI systems tasked with performing business functions that will move beyond giving recommendations to humans and beginning to act without human oversight, if for no other reason to take advantage of the speed advantages they will have over human operators.

Such actions may be AI systems requesting information from other AI systems. Or it may take the form of an AI requesting a service from a more traditional system, such as an AI tasked with inventory control making an order with a supplier. In both cases, these interactions will raise the need for an AI system to be able to exchange value with a partner in an automated manner. Of course, such interactions are already possible, with companies, for example, providing API's to access their systems, and traditional currency being used to exchange value. However, current systems have their limitations. Traditional financial payment systems have significant limitations. They impose transactions costs, which may rise exponentially in an environment where actors are interacting at the speeds associated with digital technology. They are generally poor at handling micro-transactions, which may become increasingly common when interacting systems must pay for the computing power required for AI systems to execute. Moreover, such systems generally depend on a trust relationship existing between parties prior to transactions. All these inefficiencies may serve to diminish the productivity gains that many expect to arise from AI systems into business.

## **Blockchain Technology and Cryptocurrencies**

A blockchain is a set of data storage units usually referred to as blocks that is stored on a list in the order in which they were created (Gorkhali, Li, and Shrestha 2020). A blockchain can be distributed, which is a storage model where copies of the blockchain are stored and synchronised across multiple computing nodes. A distributed blockchain can be used to create an unalterable database. When this database stores transactions, it allows for the creation and secure transfer of digital assets. Amongst other purposes, these digital tokens can be used to exchange value between actors, serving as what are commonly referred to as cryptocurrencies. The most famous implementation of this model to date is Bitcoin, which is a digital currency that enables users to transfer currency pseudo-anonymously without the need for a central authority regulating the transactions. Bitcoin's white paper (Nakamoto 2008) has been used as the basis of many other blockchain-based technologies.

Since the original development of the suite of technologies referred to blockchain, there has been a steady stream of theoretical and practical developments. Of particular note is the development of blockchain platforms, which seek to move beyond storing data to providing a decentralised distributed computing platform. The oldest and most prominent example of this trend, Ethereum, is a multipurpose blockchain platform. A particular feature of Ethereum is that developers can write small fragments of code, called smart contracts, which can execute on a distributed virtual machine called the Ethereum Virtual Machine (EVM) (Wood 2014). Smart contracts offer a way of digitizing and automating the execution of trustless agreements between parties (Szabo 1997). Blockchain can be viewed as a set of related, interlocking and rapidly evolving technologies that provide a set of capabilities to actors that use them. In the following section, we categorise these capabilities.

### **Crypto-economic Primitives**

Cryptocurrencies and other applications such as NFTs are built from a suite of crypto-economic primitives, including a shared, tamperproof database or ledger, digital assets and a set of protocols that dictate how actors can interact via those primitives. There are obvious parallels between these crypto-economic primitives and the components required to create more traditional markets. The exchange of tokens or

cryptocurrencies of value can serve the same purpose as fiat currencies did in traditionally constructed markets. Similarly, the blockchain, an itemised, ever-increasing list of transactions can be trivially re-purposed into a list of exchanges made by participants, providing traceability and transparency with regard to transactions. Taken as a whole these crypto-economic primitives allow for the creation of both traditional and novel markets structure. They enable these constructions with the significantly lower overheads and efficiency improvements associated with digital environments, while also delivering the additional benefits of being shared and immutable, an important consideration in establishing and building trust in a trading environment.

### **Decentralisation**

Blockchains such as Bitcoin or Ethereum are permissionless and public. The associated blockchain can be downloaded by anyone in the world and anybody can add records to the public blockchain. However, other models of ledger construction are possible. With permissioned blockchains only nodes that have been granted permission to access the network can download the blockchain and add records. Prominent examples of such networks include HyperLedger and Ripple. Such networks may still be decentralised, in the sense that many nodes from many different organisations in many locations may participate, and no node has a veto on adding transactions to the ledger etc. In these cases, the degree of decentralisation is a design decision in the hands of the access permission granting authority.

In generally, decentralisation lends two important characteristics to blockchain based constructs. The first is fault tolerance. The distribution of data and computing across many computers means the system as a whole has fewer points of failure. This fault tolerance is a function of the degree of decentralisation across the network as a whole. A permissionless public network like Bitcoin is essentially as resilient as the Internet itself, while, for contrast, a private blockchain consisting of nodes inside a single organisation is vulnerable to any failure that affects the entire organisation.

The second major characteristic of such systems is that the data stored on the blockchain is generally considered to be immutable, in that no single party can arbitrarily change a record once it has been added. Of course, this immutability is not

absolute. Attacks such as a 51% attack, whereby a group of malicious nodes acting together can conspire to alter the blockchain are theoretically possible. From a practical perspective however, such attacks are extremely difficult, and are again a function of the degree of decentralisation of the network, in that the more nodes that store a copy of the ledger, the harder it is to mount such attacks.

### **Designable Anonymity and Privacy**

In the public mind, cryptocurrencies and by extension blockchain technology is often associated with anonymous and therefore legally dubious financial transactions. This is a too crude representation of the situation. In reality, blockchain technology offers a palette of design choices. This can be considered along two dimensions, that of anonymity (can the participants in a market be tied to a specific “real world” identity) and privacy (can the modifications made to data stored on a blockchain be tied to a particular participant).

At one extreme, public, permissionless blockchains like Monero essentially allows completely anonymous and private participation in a blockchain based system. Both the identity of the participant and the information they add is provably untraceable. On the other hand, blockchains may also be designed in such a way that all the transactions undertaken by a particular account are publicly visible. In this case, participants have anonymity, but not privacy. There are numerous examples of blockchains operating thusly, with Bitcoin itself being the most famous.

Other configurations are also possible. A permissioned blockchain by definition requires that participants identify themselves to a gatekeeper before they can use the blockchain and participate in the network. A permissioned blockchain can be constructed in a decentralised manner, retaining the advantages of decentralisation, while at the same time insisting that participants prove their identity. In many situations, this management of participants is a legal or regulatory necessity. However, it is also possible in this situation to construct the blockchain in such a way that transactions cannot be tied back to a particular participant. This allows for the construction of markets which are not anonymous, but are private, thereby allowing participants to add information to the blockchain without fear of social or power dynamics which can often be an impediment to truthful information revelation.



## Oracles

Within their own context, blockchains are used to create an immutable ledger of irreversible transactions. These guarantees allow them to be used to exchange value in the form of Bitcoins and other cryptocurrencies. However, these guarantees only extend to data that is directly recorded on the blockchain ledger. One of the major challenge for creating blockchain and decentralised applications is that they will often require information from the “real world”. For example, to implement a simple futures contracts, two participants may agree to a smart contract that will automatically pay the second participant funds from the first participant's account if a particular stock price exceeds a particular value. The stumbling block is providing the smart contract with the stock price in the real world. Both of the participants in the smart contract have an obvious vested interest in misleading the smart contract. These misincentives can affect any third party providing information to a smart contract. This is referred to as the Oracle problem and can be simply described as the problem of gathering verified, reliable information about the real world.

This challenge is being address in a number of ways using blockchain technology. A number of approaches are being investigated. The first, and simplest, is that an independent third party is appointed as arbitrar and provider of information. This approach has the virtue of simplicity, and given a suitable third party, it is a plausible, pragmatic solution to the problem. However, it does not ultimately resolve the challenge of incentive misalignment and is contrary to the animating spirit of decentralisation. Moreover, if such a system requires human judgement, scalability will inevitably become a problem.

Other approaches seek to use the principles of decentralisation and incentive alignment. Voting is one simple solution. First is simple voting. In this model, after data has been added to the blockchain, participants are asked to vote to confirm the validity of the information. Crypto-economic primitive are used to construct additional safeguards. In order to vote, participants must stake their own cryptocurrency or equivalent digital assets on the accuracy of their vote. Participants who vote with the majority receive their own stake back, plus a percentage of the combined stakes of the participants who voted for a different evaluation of the data.

A second model is based on the notion of allowing participants to challenge an Oracle. In this case, an Oracle adds data to the blockchain. As part of adding the data, the Oracle must stake its own digital assets on the veracity of the data. After a period of time has elapsed, if no dispute is raised, then the data is confirmed and the Oracle receives a percentage fee from all blockchain participants, as payment for the information they provided. In that period of time, other participants can challenge the Oracle, by staking their own assets to contest the veracity of the provided data. If the value of the assets staked against a veracity claim exceeds a limit determined by pre-determined mathematical formula, a voting process commences, and if the Oracles outcome is rejected, the Oracles entire stake is deemed forfeit and distributed amongst the dissenters. On the other hand, if the Oracles outcome is upheld, the dissenters stakes are forfeit. This approach attempts to avoid the temporal overhead associated with simple voting, while ensuring that incentivised collective oversight applies.

## **Research Question**

Artificial intelligence and AI systems are exciting enormous interest at this time, with both national governments and the world's largest corporations spending enormous amounts of time and money on promoting research. There are significant challenges in developing AI systems. Some of these challenges are technological in nature, but many are more concerned with the potential social, economic and political impact of AI. The breakneck speed of technological advances in this space makes the necessity of designing and developing ways of addressing these challenges all the more urgent. As such, this imperative necessitates a broad effort to draw on solutions and ideas from a wide range of disciplines and perspectives.

This paper aims to explore the question, “What challenges in building socially and economically beneficial AI systems can potentially be addressed by the application of blockchain technology?” This research is exploratory in nature. Blockchain technology allows for the creation of socio-economic artefacts that have unique properties. For example, Data Markets can be created which allow a user or participant to have certainty about attributes of a data set either stored or reference in the data market, without needing a trust based relationship with other participants in

the market. The objective of this paper is identify specific artefacts that can be created using blockchain technology that may be used to address some of the challenges raised by the development of AI.

The major intellectual work in this paper is the synthesis of two distinct disciplines, namely blockchain technology and AI, with a view to enabling new theoretical solutions and perspectives to emerge (Torraco, 2005). One of the ways that this work can be conducted is an integrative literature review (Snyder, 2019). Integrative literature reviews aim to synthesise existing mature topics in order to generate novel frameworks and new theoretical models that may advance the state of the art.

## **Analysis**

Blockchain technology can be used to build at least three types of socio-economic constructs that have specific features and capabilities that mean they could be used to address the challenges outlined. These constructs have particular features or attributes that mean they can address some of the challenges associated with the development of AI. Such constructs can be designed to meet specific requirements. For example, because all systems built using blockchain technology allows for designable privacy and anonymity, the socio-economic constructs described can be tailored to the needs of the context.

In the following subsections, we describe three types of constructs that can be built using blockchain technology and crypto-currencies. For each type of construct, we describe how it can be used to address some of the challenges that are associated with the development and deployment of AI systems. Where appropriate, we discuss the choices available to designers that would allow them to better match systems to the socio-economic requirements.

## **Data Markets**

The first potential application of blockchain technology to address some of the challenges associated with the development of AI systems is using blockchain technology to create data markets. In their simplest form, these would be blockchains that would store either the data used to train AI models, or the actual trained models

themselves. In either case, the data stored would have the same guarantees around immutability that are normally conferred by blockchain technology. The blockchain could be designed to match the particular balance of anonymity and privacy required. One possible concern is that the size of the datasets, particularly training datasets might be too large to be distributed in a permissionless environment. This could be addressed by either only allowing access to participants who can meet the computational requirements of storing and distributing large blockchains, or the blockchain might only store the hashes of data, rather than the data itself.

From the perspective of addressing challenges around data quality in AI development here, the attribute of a data market is that rather than imagining a solitary actor responsible for determining the accuracy or inaccuracy of information, we instead imagine an eco-system where many evaluators, potentially both human and AI interact to evaluate the accuracy of information, with successful agents being rewarded with digital assets (which in turn would have the effect of increasing their impact on future evaluations), and unsuccessful agents being penalised.

The problem of evaluating the provenience of information is similar to the challenge of constructing an Oracle that can provide access to validated real world information to a smart contract. Researchers and practitioners have developed several models on how to guarantee the integrity of data used in smart contracts on a blockchain, and these models can be applied to the problem of verifying and validating data. Broadly speaking, these models can be broken into types, those which use trusted third parties to provide data, and those which use a consensus mechanism to arrive at a evaluation of the data provided. Trusted third party models have the advantage of simplicity, but essentially serve to re-situate the validation problem. As described in the section on Oracles, consensus models attempt to use the attributes of blockchain technology to create systems where participants are incentivised to search for and reveal the most accurate evaluation of an information source they can provide.

Further mechanisms could be used to improve the evaluation of information. Evaluators could be linked to their real world identities. In this case, evaluators reliability could be tied to their skills and reputation. A person in the real world who

has an advanced qualification in Maths may be seen as a more reliable evaluator of mathematical information than someone who doesn't have a qualification.

More scalably, an alternative model would see evaluators ranked based on the combination of their history of validated evaluations and the weighting they give to their evaluation. In this case, it is easy to build AI systems who are designed to evaluate the integrity and trustworthiness of data that is presented to them. Similarly, it is easy to understand how these AI systems could interact on a market. In this model, the evaluation of information is being performed by AI's, with their advantages in speed and scale. However, rather than depending on one "black box" AI to be accurate, in this framework the accuracy of evaluation is based on a diverse group of agents who have an incentive to compete.

Using a data market to store and verify data also provides a technological foundation for building a regulatory and compliance regime. By using blockchain technology to store data, you are creating a publicly available and thus publicly reviewable data sets that can be used for training or instantiating models. By providing this information in a public, verifiable and immutable form, at least one of the pre-requisites that will be required to create robust regulatory and compliance regimes can be met.

Enabling secure, verifiable data and model distribution is another challenge associated with the development of AI systems. In the case of either the training dataset or the model itself being made available, it is necessary to be able to guarantee the integrity of the shared data. Otherwise, the significant risk is that an adversary would be able to corrupt the data with malicious intent, allowing them to, for example, degrade the performance of an AI system, or alter the data in such a way as to allow them to select or predict the output of the AI given certain inputs.

One of the core capabilities of blockchain technology is its ability to guarantee the integrity of data. This capability has obvious applications in the context described above. Blockchain technology can be used to verify the integrity of the both training data or model. The widespread availability of training data/models should serve as a prompt to innovation. Moreover, the distribution of training data/models could also ameliorate the potential environmental impact of the development of AI. Capturing

and storing the large data sets required for training is expensive in terms of computational power. Training models is often exponentially more expensive. In both cases, a massive amount of duplicated can be avoided, assuming state and corporate actors are willing to work collaboratively.

### **Prediction/Decision markets**

A second potential application of blockchain technology to the address the challenges of AI is the use of blockchain based prediction markets. Prediction markets are “markets that are designed and run for the primary purpose of mining and aggregating information scattered among traders and subsequently using this information in the form of market values in order to make predictions about specific future events” (Tziralis and Tatsiopoulos 2007). This definition emphasises their use of a market mechanism to aggregate the information held by a group of participants regarding future uncertain events (Buckley 2016). It also distinguishes them from other markets, such as those whose primary purpose is investment, the hedging of risk or enjoyment (Wolfers and Zitzewitz 2004).

Since their origin in the 1980’s, they have been the subject of small but steady stream of academic research. Proponents suggest that they have a number of advantages over comparable information aggregation mechanisms such as polls or expert groups. First, prediction markets encourage information revelation (Hahn and Tetlock 2006b; Hall 2010). Second, they reward participants for searching for relevant information (Berg & Rietz, 2003; Hahn & Tetlock, 2006a; Sunstein, 2006). Third, they automatically communicates and aggregate information through the use of a market (Garvey and Buckley 2010) . Another fourth benefit is that the market provides an inherent weighting mechanism for the information provided. If participants are more confident of their beliefs in a particular topic, they will be willing to buy more of the relevant contracts, and vice versa (Berg and Rietz 2006; Graefe and Weinhardt 2008; Hahn and Tetlock 2006a). Fifth, markets, particularly those implemented using information technology can scale to very large groups (Hahn and Tetlock 2006c) Rather than providing point estimates like polls, prediction markets can operate in real-time over an extended period of time (Spann and Skiera 2003). Traditionally prediction markets have been implemented using traditional computing platforms, but the advent of blockchain technology has excited new interest in prediction markets as the

characteristics of this technology has particular resonances with prediction markets. Prediction markets can trivially be converted to what are called decision markets. In this case, rather than select from a range of possible forecasts, the market is asked to select from a range of possible decisions.

Decision markets can be a solution to address the issue of oversight and ethical considerations with autonomous AI systems making decisions. From the outset, it is important to note that this is a challenge to all forms of decision-making in modern society. Corporations, governments and individual experts make poor decisions every day, and often the effects of these decisions are borne by individuals. An individual who, for example, suffers an unwarranted incarceration due to a biased decision doesn't care whether the decision is made by an AI system or a human. Poor decisions are not solely the purview of AI systems. Nonetheless, we should always seek ways to constrain the ability of individual agents, be they human, corporate or AI to make poor, malicious or short-sighted decisions.

As with Data Markets, the underlying principle here is not to seek a perfect AI decision maker free from biases or imperfections, but instead to use crypto-economic primitives to create decision markets that amalgamate the decisions of many interacting AI agents. Many of the traditional limitations that affect decision markets do not apply here. AI systems can be directed to have an opinion, and so the problem of non-participants is resolved. AI systems can interact at computational speed, and so decisions can be reached practically immediately, removing the time issues that bedevil markets that require coordination and communication in human time scales. Decision markets implicitly reward or punish participants. Other possibilities present themselves. There is no obvious technological impediment to humans participating in decision-making, allowing for human input into the decision-making process. The key point here is that what is required is a diversity of AI systems. Rather than depending on the validity and good intentions of one model and one model making actor, we are depending on the wisdom of the crowd (in this case a crowd of interacting actors, many of whom may be AI's) and a market mechanism to reward the best decision makers over time.

## **Smart Contracts**

A third major application of blockchain technology to some of the challenges of AI development is the use of smart contracts. More advanced blockchain platforms such as Ethereum offer participants the ability to interact with smart contracts. These smart contracts are essentially programming code that represent business logic. This code executes in the context of the blockchain. They have the ability to create, store and transfer digital assets stored on the blockchain. They are guaranteed to execute in accordance with their code, and provide a way of allowing participants on a blockchain to interact in a more advanced and customised way than the simple exchange of digital assets and cryptocurrencies.

As AI systems move into the operational aspects of businesses, many expect the advantages they possess to offer increased productivity. However, these productivity gains are dependent not just on the effectiveness of the AI system themselves, but also on the systems these AIs interact with. Gains in speed and efficiency offered by an AI system may be quickly swallowed up by inefficiencies in other parts of the supply chain. In this context, the limitations of the existing financial systems in terms of handling high velocity, low value transactions amongst trustless entities may be a major impediment to the gains many expect to gain from AI use.

In this context, blockchain technology may again offer a supporting technology that can ameliorate a particular challenge to AI systems. Many blockchains are explicitly designed to support the low costs, high velocity exchange of value in a trustless environment. There is no doubt that currently operational blockchains have limitations in terms of the velocity and volume of transactions they can support, but innovations such as sharding and chaining are being actively developed to address these limitations. In addition, advanced blockchains such as Ethereum offer smart contracts, which allow for the execution of business logic securely in a trustless environment. These offer the ability to build automated marketplaces that move beyond the exchange of value. Smart contracts can be used to provide financial services such as payment processing, loans and insurance. They can also be used to provide information services and the management of digital assets. These blockchain platforms offer the potential to remove many of the inefficiencies of traditional financial networks.



## **Limitations and Future Research**

This paper presents an integrative literature review that synthesises the research in two topics to identify spaces where blockchain technology can be used to address some of the challenges associated with AI. As such, the research is exploratory in nature, and suffers from the limitation associated with that work. The operationalisation of any of the concepts derived in this paper would require a significant body of further work, both theoretical and empirical. From a theoretical perspective, many of the concepts outlined in this paper require a more detailed examination within the proposed context. Empirically, many of the properties attributed to blockchain based socio-economic constructs, e.g. the accuracy of prediction market forecasts require would require empirical validation. The research presented here aim to serve as signposts and suggestions for where research efforts might be useful focussed in the future.

## **Conclusions**

In this paper, we have discussed how blockchain technology can be used to build constructs that can address some of the challenges that affect the development and deployment of AI systems. The technological capabilities offered by Blockchain only represent part of the solution to these challenges. For example, blockchain based information markets theoretically allow for large numbers of actors to contribute to the creation of large sets of training data in an untrustworthy environment. This would have the effect of reducing the amount of duplicate effort actors would otherwise have to go to in order to create large data sets individually. However, for this benefit to accrue would require actors to active in collaboration. A decision market allowing AI agents to interact to arrive at a consensus necessarily requires the participation of actors who accept that their AI agents may be flawed.

More generally, many of the potential benefits of using Blockchain technology described above are predicated on actors be willing to act collaboratively. In a market economy, this is far from a given. It seems very likely that many of the actors in the AI development will be willing to forgo efficiency benefits in the name of capturing a technological edge over their competitors. Such an issue is just one of the many where the development of AI models moves into the realm of political, economic and social

considerations. The question thus becomes one of political desire for egalitarian development of AI technology and political will to enforce it. Blockchain technology is important because it provides a technological foundation that can be used to build a more egalitarian version of AI development, but it will remain dependent on a political class desiring and if necessary, forcing these more egalitarian development paths.

## References

- Arntz, Melanie, Terry Gregory, and Ulrich Zierahn. 2016. "The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis," May. <https://doi.org/10.1787/5jlz9h56dvq7-en>.
- Asimov, Isaac. 1950. *I, Robot*. Garden City, N.Y.: Doubleday.
- Autor, David H., and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review* 103 (5): 1553–97. <https://doi.org/10.1257/aer.103.5.1553>.
- Badue, Claudine, Rânik Guidolini, Raphael Vivacqua Carneiro, Pedro Azevedo, Vinicius B. Cardoso, Avelino Forechi, Luan Jesus, Rodrigo Berriel, Thiago M. Paixao, and Filipe Mutz. 2021. "Self-Driving Cars: A Survey." *Expert Systems with Applications* 165: 113816.
- Baum, Seth D., Ben Goertzel, and Ted G. Goertzel. 2011. "How Long until Human-Level AI? Results from an Expert Assessment." *Technological Forecasting and Social Change* 78 (1): 185–95.
- Berg, Joyce E., and Thomas A. Rietz. 2003. "Prediction Markets as Decision Support Systems." *Information Systems Frontiers* 5 (1): 79–93.
- . 2006. "The Iowa Electronic Markets: Stylized Facts and Open Issues." In *Information Markets: A New Way of Making Decisions*, edited by Robert W Hahn and Paul C Tetlock, 142–69. Washington D.C: AEI-Brookings Joint Center for Regulatory Studies.
- Bostrom, Nick. 2016. *Superintelligence: Paths, Dangers, Strategies*. Reprint edition. Oxford, United Kingdom ; New York, NY: Oxford University Press.
- Buckley, Patrick. 2016. "Harnessing the Wisdom of Crowds: Decision Spaces for Prediction Markets." *Business Horizons* 59 (1): 85–94. <https://doi.org/10.1016/j.bushor.2015.09.003>.
- Busch, Lawrence. 2014. "Big Data, Big Questions| A Dozen Ways to Get Lost in Translation: Inherent Challenges in Large Scale Data Sets." *International Journal of Communication* 8 (0): 18.

- Clark, Andy. 2005. "Intrinsic Content, Active Memory and the Extended Mind." *Analysis* 65 (1): 1–11.
- Crevier, Daniel. 1993. *AI: The Tumultuous History of the Search for Artificial Intelligence*. Basic Books.
- Dennett, Daniel C. 2017. *From Bacteria to Bach and Back: The Evolution of Minds*. 1 edition. New York: W. W. Norton & Company.
- Garvey, John, and Patrick Buckley. 2010. "Implementing Control Mutuality Using Prediction Markets: A New Mechanism for Risk Communication." *Journal of Risk Research* 13 (7): 951–60. <https://doi.org/10.1080/13669877.2010.488742>.
- Gorkhali, Anjee, Ling Li, and Asim Shrestha. 2020. "Blockchain: A Literature Review." *Journal of Management Analytics* 7 (3): 321–43. <https://doi.org/10.1080/23270012.2020.1801529>.
- Graefe, Andreas, and Christof Weinhardt. 2008. "Long-Term Forecasting with Prediction Markets A Field Experiment on Applicability and Expert Confidence." *The Journal of Prediction Markets* 2 (2): 71–91.
- Hahn, Robert W, and Paul C. Tetlock. 2006a. "A New Tool for Promoting Economic Development." In *Information Markets: A New Way of Making Decisions*, edited by Robert W. Hahn and Paul C Tetlock, 170–94. Washington D.C: AEI-Brookings Joint Center for Regulatory Studies.
- Hahn, Robert W., and Paul C. Tetlock. 2006b. *Information Markets: A New Way of Making Decisions*. Washington D.C: AEI-Brookings Joint Center for Regulatory Studies.
- . 2006c. "Introduction to Information Markets." In *Information Markets: A New Way of Making Decisions*, 1–12. Washington D.C: AEI-Brookings Joint Center for Regulatory Studies.
- Halevy, Alon, Peter Norvig, and Fernando Pereira. 2009. "The Unreasonable Effectiveness of Data." *IEEE Intelligent Systems* 24 (2): 8–12. <https://doi.org/10.1109/MIS.2009.36>.
- Hall, Caitlin. 2010. "Prediction Markets: Issues and Applications." *The Journal of Prediction Markets* 4 (1): 27–58.
- Harari, Yuval Noah. 2016. *Homo Deus: A Brief History of Tomorrow*. Random House.
- Hoy, Matthew B. 2018. "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants." *Medical Reference Services Quarterly* 37 (1): 81–88.
- Joy, Bill. 2000. "Why the Future Doesn't Need Us." *Wired Magazine* 8 (4): 238–62.

- Kumar, Mukesh, Utsav Anand Mani, Pranjali Tripathi, Mohd Saalim, Sneha Roy, and Sneha Roy Sr. 2023. "Artificial Hallucinations by Google Bard: Think Before You Leap." *Cureus* 15 (8).
- Kurzweil, Ray. 2005. *The Singularity Is near: When Humans Transcend Biology*. Penguin.
- Marr, Bernard. 2019. *Artificial Intelligence in Practice: How 50 Successful Companies Used AI and Machine Learning to Solve Problems*. John Wiley & Sons.
- Mindell, David A., and Elisabeth Reynolds. 2022. *The Work of the Future: Building Better Jobs in an Age of Intelligent Machines*. MIT Press.
- Muehlhauser, Luke, and Anna Salamon. 2012. "Intelligence Explosion: Evidence and Import." In *Singularity Hypotheses: A Scientific and Philosophical Assessment*, edited by Amnon H. Eden, James H. Moor, Johnny H. Søraker, and Eric Steinhart, 15–42. The Frontiers Collection. Berlin, Heidelberg: Springer. [https://doi.org/10.1007/978-3-642-32560-1\\_2](https://doi.org/10.1007/978-3-642-32560-1_2).
- Nakamoto, Satoshi. 2008. "Bitcoin." *A Peer-to-Peer Electronic Cash System*.
- Nilsson, Nils J. 2009. *The Quest for Artificial Intelligence*. 1 edition. Cambridge ; New York: Cambridge University Press.
- Papernot, N., P. McDaniel, S. Jha, M. Fredrikson, Z.B. Celik, and A. Swami. 2016. "The Limitations of Deep Learning in Adversarial Settings." In , 372–87. <https://doi.org/10.1109/EuroSP.2016.36>.
- Penrose, Roger, and Martin Gardner. 2002. *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. 1 edition. Oxford: Oxford University Press.
- Philip Chen, C.L., and C.-Y. Zhang. 2014. "Data-Intensive Applications, Challenges, Techniques and Technologies: A Survey on Big Data." *Information Sciences* 275: 314–47. <https://doi.org/10.1016/j.ins.2014.01.015>.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Spann, Martin, and Bernd Skiera. 2003. "Internet-Based Virtual Stock Markets for Business Forecasting." *Management Science* 49 (10): 1310–26.
- Sun, C., A. Shrivastava, S. Singh, and A. Gupta. 2017. "Revisiting Unreasonable Effectiveness of Data in Deep Learning Era." In , 2017-October:843–52. <https://doi.org/10.1109/ICCV.2017.97>.

- Sunstein, Cass R. 2006. "Deliberating Groups vs. Prediction Markets (or Hayek's Challenge to Habermas)." *Episteme: A Journal of Social Epistemology* 3 (3): 192–213. <https://doi.org/10.1353/epi.2007.0007>.
- Szabo, Nick. 1997. "Formalizing and Securing Relationships on Public Networks." *First Monday*.
- Tziralis, Georgios, and Ilias Tatsiopoulos. 2007. "Prediction Markets: An Extended Literature Review." *The Journal of Prediction Markets* 1 (1): 75–91.
- Torraco, R. J. (2005). Writing Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review*, 4(3), 356–367. <https://doi.org/10.1177/1534484305278283>
- Wolfers, Justin, and Eric Zitzewitz. 2004. "Prediction Markets." *The Journal of Economic Perspectives* 18 (2): 107–126.
- Wood, Gavin. 2014. "Ethereum: A Secure Decentralised Generalised Transaction Ledger." *Ethereum Project Yellow Paper* 151 (2014): 1–32.
- Zhang, B.H., B. Lemoine, and M. Mitchell. 2018. "Mitigating Unwanted Biases with Adversarial Learning." In , 335–40. <https://doi.org/10.1145/3278721.3278779>.

# Driving Student Success through a Data-Driven Approach in Higher Education

*Development Paper*

**Dr Yun Chen and Dr Kate Han**

*Salford Business School, The University of Salford, UK*

## **Abstract**

*The project aims to explore a data-driven approach to enhance student engagement and achievement in Higher Education (HE), with the ultimate goal of promoting student success. The project utilised a case study approach at the University of Salford, employing data analysis and Machine Learning (ML) techniques to understand correlation between students' engagement and academic performance to support strategy of students support in their learning process. Being a project in progress, this paper delves into the initial phase of our research findings. This phase focuses on the data collected from a pilot module, specifically pertaining to student engagement and progression data. Additionally, the paper presents a prototype of a ML algorithm that aims to facilitate decision-making in the realm of student support. Moving forward, the next stage of the project aims to automate the entire process, spanning from data analysis to student intervention. It aims to use this automation to drive student success throughout their HE journey.*

**Keywords:** Higher Education (HE), student engagement, data-driven approach, student support strategy, data analysis, Machine Learning (ML) algorithm

## **1.0 Introduction**

The concept of "Student Engagement" encompasses the degree of enthusiasm, attentiveness, and dedication that students invest in their educational experiences (Kahn, 2014). Strong student engagement within higher education (HE) can yield numerous positive outcomes, benefiting not only individual students but also the wider academic community (Bryson, 2014). These advantages transcend the confines of the classroom and contribute to personal, academic, and institutional achievements.

It's essential to recognise that student engagement is a shared responsibility, not solely borne by students themselves. Educational institutions also play a pivotal role in fostering and sustaining an environment conducive to student engagement (Bond & Bedenlier, 2019). This involves employing effective teaching techniques, implementing mentoring programs, establishing monitoring mechanisms, and providing ongoing support. When institutions prioritise and invest in strategies to enhance student engagement, the rewards can be substantial for all stakeholders within the realm of HE.

It is believed that there is generally a positive correlation between student engagement and academic performance, but the strength and nature of this correlation can be based on several factors, such as interventions by academic tutors (Fredricks et al., 2019). UK universities are increasingly acknowledging the pivotal role of personal tutoring in bolstering student engagement within the HE milieu and promoting future success. It is widely acknowledged that personal tutors play a crucial role in cultivating a sense of belonging, which is fundamental to students' development of a learner identity and sustained engagement (Lochtie & Walker, 2022; Ross et al., 2014)

At the University of Salford, a progression framework has been established since 2016, identifying three key facilitators vital for student engagement: fostering a sense of belonging, offering support, advice, and guidance, and building academic confidence. A dedicated group of academic staff members, known as Academic Progress Tutors (APTs), has been appointed to champion these initiatives. APTs have access to university systems that provide data for a comprehensive understanding of student engagement and achievement. They use this information to organise APT meetings with relevant students, offering academic references throughout their program and after graduation. These meetings are designed to facilitate robust personal academic development and professional growth through active engagement. They place a significant emphasis on increasing learner autonomy and challenging students to expand beyond their current capabilities.

This developmental paper introduces a continuing research project conducted at Salford Business School. The primary objective of this research is to tackle the obstacles associated with implementing a data-driven approach to enhance student engagement. In this paper, we will begin by examining the challenges we have encountered, then proceed to explore the potential solutions we have developed, and finally outline the next steps for this project.

## **2.0 Students Support Challenges in HE**

### **2.1 Data Integration**

The pandemic has expedited the adoption of a variety of information systems and learning platforms within HE institutions. These platforms provide invaluable data insights by leveraging existing data resources. However, a significant challenge arises due to the disparate storage of data across various information systems within universities (e.g., QlikView, Jigsaw, Blackboard). Institutions encounter a significant challenge when attempting to integrate various systems for monitoring student engagement, including absence data, academic alerts related to tutorials, and case management. This fragmented data landscape leads to issues of data integration, resulting in unclean, irrelevant, and redundant data. Furthermore, when these systems fail to communicate seamlessly, gaining a comprehensive overview of students at risk becomes a formidable task. Consequently, this poses a substantial challenge for academic tutors who rely on this data to identify low-engaged students and implement interventions to enhance their performance (Aldowah, H., Al-Samarraie, H. and Fauzy, W.M., 2019).

A prevalent research focus aims to tackle this challenge through initiatives like the development of information systems for monitoring student engagement (JISC 2022; StDREAM, 2023). Another approach involves proposing models that aid decision-makers in identifying crucial factors contributing to elevated graduation rates, as discussed by Addison and Williams in 2023.

## **2.2 Complexity in Student Support**

Although information systems and machine learning have enhanced prediction based on data analytics to enable early intervention in instances of academic risks, the most current models often focus solely on subjective student factors without examine the external environmental and objective elements. (Qin, et al., 2023). The COVID-19 global pandemic has introduced substantial disruptions to HEIs teaching and learning, posing challenges that existing models struggle to address. It enhances the importance of the role and operational procedures of personal tutors within universities. In current HE, monitoring student engagement across multiple systems has the potential to impose an increased workload and add barriers for the APT team. Considering the substantial number of students in Higher Education, APT team members are now tasked with overseeing student behaviour and activities across a multitude of systems and platforms,



with consideration on external factors who might be the reason to impact students' performance. (Gajewski, E. M., 2023) These systems and platforms generate distinct sets of data, complicating the work of the APT team as they must develop proficiency in comprehending and interpreting these diverse data sources and insights. Furthermore, they are required to amalgamate data from these platforms before making informed decisions, rather than relying on a single system. These additional demands place barriers in the path of the APT team, requiring not only data analysis skills but also strong communication skills for effective student support.

In addition to the two aforementioned challenges, diverse HEIs face unique situations, adding further complexity. This variability encompasses aspects such as dataset capture, adopted information systems, and factors influencing student performance. Consequently, a pilot case study was conducted specifically for Salford Business School, which will be discussed in the following session.

### **3.0 Pilot Case Study**

To address the aforementioned challenges, an internally funded project initiated by the University of Salford has been underway since February 2023. Its primary aim is to explore the integration of data from diverse learning platforms and the meticulous cleansing of this data to eliminate inaccuracies. This process is intended to provide a clearer understanding of the correlation between student engagement and their academic performance. Once the data pipeline is established, it opens the door for the adoption of ML algorithms, ultimately influencing decision-making and strengthening responsive student support.

The project started with interviews involving the APT team at Salford Business School in May 2023. Two interviews took place—one with the Head of the APT team and another with the Academic Student Success Lead. These interviews served to validate the data integration and process complexity issues discussed in the prior session. Furthermore, both interviewees expressed concerns about the time-consuming nature of consistently assessing and monitoring the advancement of disengaged students. They also highlighted the challenge of determining the optimal timing for intervention.

During the interviews, another focal point was to pinpoint the essential attributes that the project should utilise for predicting students' performance, particularly for Salford Business School. This will be elaborated upon in the later part of this session.

After conducting interviews, we selected a pilot module, i.e., Level 7 in Information Systems and Digital Transformation, to develop our machine learning algorithm for data analysis and predicting students' performance. This work will establish the groundwork for upcoming phases.

### **3.1 Establishing the Data Pipeline**

The establishment of this data pipeline forms the bedrock of our research and sets a noteworthy precedent for the consistent management of extensive datasets in HE. Operating as a pilot dataset, we gathered data from both QlikView and VLE (i.e., Blackboard). The objective was to integrate data from these two systems while filtering out extraneous information from the raw data. This effort was aimed at uncovering correlations between student engagement and academic performance data with consideration of their personal backgrounds. The selection of relevant features was guided by input from the APT team during interviews, taking into account aspects such as gap awards and Equality Diversity and Inclusion (EDI) considerations.

As indicated in Table 1 below, the integrated dataset for the pilot module consists of selected features. The Blackboard dataset encompasses data on student weekly active time and assignment performance scores, whereas the QlikView dataset contains student personal information and registration data. Both of these initial raw data have undergone a data cleaning process, which involved feature selection and integration. This process was carried out to create a dataset that is well-suited for ML classification purposes which is explained in section 3.2.

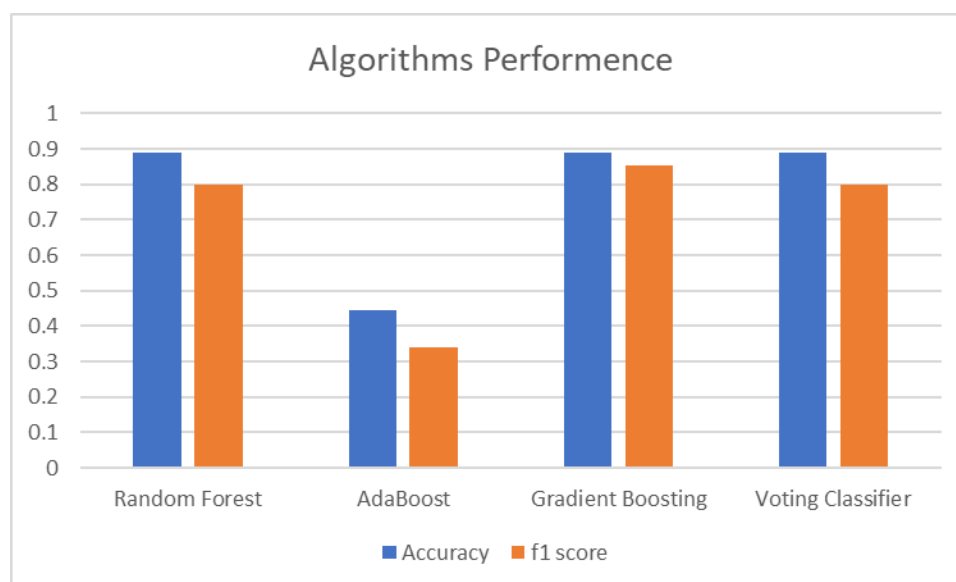
<b>Data features</b>	<b>Data source</b>
Assignment Score	Blackboard Data
Last Access of Blackboard	
Activities during w/c 08/11/2021	
Activities during w/c 15/11/2021	
Activities during w/c 22/11/2021	
Activities during w/c 29/11/2021	
Activities during w/c 06/12/2021	
Activities during w/c 13/12/2021	
Degree	QlikView data
Directorate	
Disability Description	
Domicile	
Ethnicity	
Gender	
Greater Manchester Indicator	
HQE (previous degree)	
Nationality	
Postcode on Entry	
Previous Institution Name	
Program Title	
Region	
Residency Summary	
UK Region	
Year of Program	

**Table 1. Pilot Dataset Features**

### **3.2 Developing Machine Learning Algorithms**

Following the establishment of the data pipeline, we progressed to the subsequent stage of creating training, testing, and validation datasets, with the objective of applying ML algorithms for classification. The target is the final assessment score, with features encompassing student personal information and Blackboard weekly active time.

Once the dataset was prepared, we applied various classic state-of-the-art machine learning classification algorithms, including Random Forest (Rigatti, S.J., 2017), AdaBoost (Schapire, R.E., 2013), Gradient Boosting (Natekin, A. and Knoll, A., 2013), and the Voting Classifier (Ruta, D. and Gabrys, B., 2005). For the selected module with a small dataset, we achieved promising results (see Figure 1).



**Figure 1. 2Algorithms Performance Comparison**

Figure 1 shows the performance of the four ML algorithms that we employed on the pilot dataset. Among these algorithms, Random Forest, Gradient Boosting, and Voting Classifier all achieved an Accuracy score of 0.8888 within the range of 0 to 1. However, AdaBoost fall behind with an Accuracy score of 0.4444 within the range of 0 to 1. In terms of the F1 score within the range of 0 to 1, Gradient Boosting outperformed the others, attaining the highest value of 0.8518. This showcases the promising potential of the developed ML algorithm, which has room for improvement through further training with a diverse range of datasets in future research.

#### **4.0 Conclusion and Future Work**

In this research, we successfully accomplished the integration of raw data from QlikView and VLE (i.e. Blackboard) used at the University of Salford. Additionally, we extract information APT interview results to facilitate feature selection for data pipeline and achieved promising outcomes through the application of cutting-edge ML algorithms. The next step of this project entails the acquisition of a more comprehensive

dataset to enhance the training of the ML algorithm. This extended dataset will encompass modules ranging from Level 4 to Level 7 to refine the accuracy of the developed model. Building upon a more precise ML algorithm, the next project phase also seeks to automate various functions within the APT team's workflow. This includes automating interventions based on data analysis, progress monitoring, and generating feedback reports.

The future work will be implemented in the following three phases:

### **1. Establishment of Dataset Benchmark**

Considering the results obtained in this project, it is clear that the expansion of the data pipeline to establish a benchmark dataset is not only imperative but also holds significant potential. This benchmark dataset will include modules of different sizes and levels of complexity. In addition to the current module-specific active time tracking, we will also gather general student engagement data, including library learning time and email activity for benchmark dataset set up in Academic Year 2023-24. This diversification is vital as it guarantees that the research findings derived from the benchmark will maintain their robustness and adaptability in the face of ever-changing real-world scenarios. Additionally, it is planned to make this dataset available to the wider community and potentially organise a competition to stimulate research interest within the field of ML.

### **2. Development of Ensemble Learning Algorithms**

As a natural progression, following the establishment of the benchmark dataset, the following phase involves delving into the application of ensemble learning algorithms. These algorithms will be specifically geared towards generating classification and prediction outcomes in scenarios involving incomplete data, a frequent occurrence in real-world situations. It is expected these algorithms can facilitate swift decision-making by providing timely prediction results.

### **3. Creation of APT Process Automation**

After engaging in preliminary conversations with the APT team, we have acquired valuable insights into their operational workflow and the difficulties they encounter. As the last stage of this project, we aim to provide assistance to by initially replicating

their processes and subsequently enhancing them through the application of the enhanced ML algorithms. Additionally, we intend to work in close partnership with the APT team, employing an agile approach to continually assess and propose process automation process.

## **References:**

Addison, L. and Williams, D. (2023). Predicting student retention in higher education institutions (HEIs), in *Higher Education, Skills and Work-Based learning*, Vol. 13 No. 5, pp. 865-885. Available at: <https://doi.org/10.1108/HESWBL-12-2022-0257>

Aldowah, H., Al-Samarraie, H. and Fauzy, W.M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37, pp.13-49.

Bond, M., & Bedenlier, S. (2019). Facilitating Student Engagement Through Educational Technology: Towards a Conceptual Framework. *Journal of Interactive Media in Education : JiME*, 2019(1). Available at: <https://doi.org/10.5334/jime.528>

Bryson, C. (2014). *Understanding and developing student engagement*. Routledge. Available at: <https://doi.org/10.4324/9781315813691>

Fredricks, J. A., Reschly, A. L., & Christenson, S. (Eds.). (2019). *Handbook of student engagement interventions : working with disengaged students*. Elsevier.

Gajewski, E. M. (2023). The Effects of the Covid-19 Pandemic on Students in the School Setting: The Adaptations Needed for Tiered Supports. PhD Thesis, National-Louis University.

JISC (2022). *Enhancing student engagement using technological solutions*. Available at: <https://www.jisc.ac.uk/reports/enhancing-student-engagement-using-technological-solutions>

Kahn. (2014). Theorising student engagement in higher education. *British Educational Research Journal*, 40(6), 1005–1018. Available at: <https://doi.org/10.1002/berj.3121>

Lochtie, Stork, A., & Walker, B. W. (2022). *The Higher Education Personal Tutor's and Advisor's Companion: Translating Theory into Practice to Improve Student Success (1st ed.)*. Critical Publishing.

Natekin, A. and Knoll, A., (2013). *Gradient boosting machines, a tutorial*. *Frontiers in neurorobotics*, 7, p.21.

Qin, K., Xie, X., He, Q. and Deng, G. (2023). Early Warning of Student Performance with Integration of Subjective and Objective Elements, in *IEEE Access*, vol. 11, pp. 72601-72617, 2023.

Rigatti, S.J. (2017). *Random forest*. *Journal of Insurance Medicine*, 47(1), pp.31-39.

Ross, Head, K., King, L., Perry, P. M., & Smith, S. (2014). The personal development tutor role: An exploration of student and lecturer experiences and perceptions of that relationship. *Nurse Education Today*, 34(9), 1207–1213. Available at: <https://doi.org/10.1016/j.nedt.2014.01.001>

Ruta, D. and Gabrys, B., (2005). Classifier selection for majority voting. *Information fusion*, 6(1), pp.63-81.

Schapire, R.E., (2013). Explaining adaboost. In *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik* (pp. 37-52). Berlin, Heidelberg: Springer Berlin Heidelberg.

StREAM (2023). Website: <https://www.solutionpath.co.uk>

## **Business models for digital sustainability: The case of Fintech startups**

### **ABSTRACT**

This study examines the business models of financial technology (Fintech) startups to create, deliver, and capture value while attaining digital sustainability. Adopting a qualitative method using three case studies of Fintech startups based in Taiwan, the findings indicate that firms leverage digital technologies and data to create complementary value propositions. The firms also deliver their value to their users and partners through co-supporting, co-designing, and co-engaging mechanisms. In terms of value capture, our findings show that the firms derive revenues from complementary revenue streams. This study further elaborates how these business models contribute to digital sustainability, including by enhancing the efficiency of financial service provision (economic sustainability), fostering the inclusion of underserved segments (social sustainability), and reducing the environmental impact through reduced consumption of natural resources (environmental sustainability). This study contributes to the existing literature on digital sustainability and business models by examining and analyzing business models to achieve digital sustainability, in the context of Fintech startups.

*Keywords: digital sustainability, business models, value creation, value delivery, value capture, Fintech, qualitative*

### **INTRODUCTION**

The financial services industry has transformed as a result of digitalization, characterized by the utilization of digital technologies to facilitate organizational changes and business model (BMs) innovation (Lähteenmäki, Nätti, & Saraniemi, 2022). Within this context, financial technology (Fintech) innovations have played a crucial role in reshaping the financial landscape, allowing for the provision of a much wider range of services than those offered by traditional banks (Gomber, Kauffman, Parker, & Weber, 2018; Hornuf & Haddad, 2019). These offerings by Fintech companies range from personal loans to general insurance, and more recently, financial advisory services that have traditionally been seen as more complex services (Gomber et al., 2018; Lee & Shin, 2018). These offerings have not only diversified the range of financial services available but have the potential to advance the United Nations Sustainable Development Goals (SDGs) (United Nations Environment Programme, 2016).

In this regard, Fintech solutions play a crucial role in driving the progress of digital sustainability (Bohnsack, Bidmon, & Pinkse, 2022; George & Schillebeeckx, 2021), which is defined as the “organizational activities that seek to advance the sustainable development goals through creative deployment of technologies that create,



use, transmit, or source electronic data” (George, Merrill, & Schillebeeckx, 2021: 1000). Because of their digital nature, digital sustainability activities are less constrained by geographic boundaries and exhibit high scalability across ecosystems, thereby increasing their impact. In addition, these activities are primarily guided by the objective of incorporating socioecological value creation as an integral part of an economic proposition, eliminating the need for a tradeoff between profit and purpose.

The implementation of Fintech’s technology-driven activities and solutions not only generates economic gains but also mitigates the negative societal and environmental impacts associated with traditional financial practices. However, there still remains a considerable knowledge gap regarding how Fintech businesses achieve digital sustainability. This gap arises from the relative novelty of digital sustainability as a concept, and the fact that the existing literature on sustainability and digital/technology literature has remained largely separate (Höllner, Shinkle, George, Mair, Pan, & Tim, 2023; Merello, Barberá, & De la Poza, 2022), indicating the need for further exploration of this topic. As Guandalini (2022: 456) points out, “while sustainability is undeniably one of the most growing phenomena, it is still an inadequately discussed field of application for digital technology.”

In the digital/technology literature, an increasing number of studies has also looked at how businesses create, deliver, and capture value as part of their business models (BMs) (e.g., Lähteenmäki, et al., 2022; Tidhar & Eisenhardt, 2020). For example, Fintech companies like Coinbase, a trading platform for various cryptocurrencies, have expanded their offerings to include a range of financial products and services. As an illustration, Coinbase provides customers with a credit card facility that can be loaded with cryptocurrencies for making purchases. Coinbase has also recently partnered with Google Cloud infrastructure to facilitate the storage and management of its blockchain data and data-related applications, as well as to enable Google’s customers to pay for cloud services using cryptocurrencies (Novet, 2022). This study contends that a firm’s design of its BM to create, deliver, and capture value is important, but little research has been done thus far to examine how firms utilize their BMs to achieve digital sustainability.

This study expands upon research on digital sustainability (George et al., 2021) through the utilization of a technology- and data-driven BM approach (Subramaniam, 2022). This focus on three aspects of digital sustainability: economic, social, and environmental (Lee, Che-Ha, & Syed Alwi, 2021). Fintech firms heavily rely on extensive data within their digital platforms, necessitating a critical focus on both technology and data for value creation. Furthermore, Subramaniam (2022) pointed out that the future of competitive strategy has shifted from generating income through products to generating revenue through data, underlining the significance of technology

and data in the provision of goods and services in the digital ecosystems. Specifically, the study is guided by the following research question: *How do Fintech startups' business models of value creation, value delivery, and value capture help firms in achieving digital sustainability?* We draw on three case studies of Fintech startups and our findings indicate the aspect of complementarity in value creation, delivery, and capture, while attaining digital sustainability.

This study contributes to the existing body of knowledge in both BMs and digital sustainability literatures. First, we address the research gap surrounding digital sustainability (George et al., 2021; Höllerer et al., 2023; Merello et al., 2022; Pan, Carter, Tim, & Seshadrinath, 2022; Pan & Zhang, 2020) by shedding light on how businesses, specifically Fintech startups, can leverage their BMs to achieve digital sustainability (Bencsik, Palmié, Parida, Wincent, & Gassmann, 2023). While technologies such as blockchain, AI and IoT (AIoT), have been recognized as important to achieving digital sustainability (George et al., 2021), we are still unclear how firms do this and what these digital sustainability outcomes are. This research focuses on how Fintech businesses utilize and leverage digital technologies and data to achieve digital sustainability by: (i) creating complementary value propositions with users and partners, (ii) delivering value through complementary mechanisms of co-supporting, co-designing, and co-engaging mechanisms, and (iii) capturing value through complementary revenue streams.

Second, this study adds to the understanding of the multidimensional nature of a BM in terms of value creation, delivery, and capture. Existing literature on BMs often examines certain elements, such as solely focusing on value creation (e.g., Achtenhagen, Melin, & Naldi, 2013; Freudenreich, Lüdeke-Freund, & Schaltegger, 2020; Priem, Wenzel, & Koch, 2018), or considering both value creation and value capture (e.g., Franco, Presenza, Messeni Petruzzelli, & Peruffo, 2022; Rietveld, 2018; Sjödin et al., 2020). Although there are exceptions where studies have examined all three elements of a BM (e.g., Codini, Abbate, & Messeni Petruzzelli, 2023; Saebi, Lien, & Foss, 2017), they remain relatively scarce. Zott, Amit, & Massa (2011: 1028) noted that “the business model is not a value proposition, a revenue model, or a network of relationships by itself; it is all of these elements together.” Foss and Saebi (2017: 215) further pointed out that there is still a need for “defining and dimensionalizing the BMI construct”. Accordingly, our study contributes by focusing on technologies and data to dimensionalize the three elements of a BM, namely value creation (identification of value propositions), value delivery (identification of target customers and delivery mechanisms), and value capture (identification of revenue streams) (Codini et al., 2023). Specifically, we focus on the aspects of value creation, delivery, and capture that arise

from external relationships with users and partners that span across the firm's boundary.

Finally, our context, Fintech startups, constitutes a rich setting for exploring the implications of digitalization for digital sustainability. The industry, characterized by ongoing drive and pressure to develop digital offerings (Lähteenmäki et al., 2022), plays a pivotal role in promoting financial inclusion and sustainable development (United Nations Environment Programme, 2016). Consequently, our context presents a unique opportunity to examine how Fintech startups create, deliver, and capture value while also attaining digital sustainability.

## **THEORETICAL BACKGROUND**

### **Business models**

A BM is defined as the underpinning logic, architecture, or design by which an organization creates, delivers, and captures value (Saebi et al., 2017; Teece, 2010; Zott & Amit, 2010). A firm's BM is developed to clearly articulate a firm's value propositions in a particular business context (Frankenberger & Sauer, 2019) with the goal of achieving better firm performance (Latifi, Nikou, & Bouwman, 2021). A BM is essentially used in the identification of a firm's target markets, understanding of market segmentation, provision of products and services, and generation of revenues (Chatterjee, 2013; Latifi et al., 2021).

To put it another way, a company's BM outlines how it creates value through a series of interactions with its stakeholders, such as suppliers, partners, or customers, and how it plans to deliver and profit from those interactions (Teece, 2010; Zott & Amit, 2010). In relation to how a firm creates value, it needs to be able to define their products/services and how they create such value from their offerings. With regard to value delivery of a firm's BM, this is about how and by what means firms deliver value along the supply chain using their suppliers and/or external partners (Achtenhagen et al., 2013). Value delivery is also about understanding the target customer, for example which target market they are serving or whether the firm is looking at an entirely new customer segment. Finally, value capture is described as the process by which value propositions are transformed into revenue streams (Teece, 2010).

There are many studies exploring BMs and the impact of BMs. Studies have looked at the conceptualizations of BMs (e.g., Massa, Tucci, & Afuah, 2017; Osterwalder & Pigneur, 2010), how to design a BM (e.g., Casadesus-Masanell & Ricart, 2011) and the evolution of BMs (e.g., Cozzolino, Verona, & Rothaermel, 2018; Demil & Lecocq, 2010). Others have examined the antecedents that led to the creation of BMs and the modifications to them (e.g., Amit & Zott, 2015; Frankenberger & Sauer, 2019; Osiyevskyy & Dewald, 2015), as well as the impact of BMs (e.g., Latifi et al., 2021). Schrauder, Kock, Baccarella, & Voigt (2018), for example, highlighted that BM

evaluation at the early stages of innovation has a positive impact on front-end effectiveness and efficiency. Snihur and Bocken (2022) emphasized how BM innovation affects ecosystems, society, and the environment. Practical examples and scholarly research have also shown that firms are increasingly shifting their attention away from value creation, delivery, and capture within the firm's boundary to incorporating BMs that span across the firm's boundary. For example, Aversa, Haefliger, Hueller, and Reza (2021) looked at how Amazon's BMs complement one another with these models targeting similar customer groups (such as online shop customers) and creating demand across their platform of services.

The adaptation of a firm's BM is imperative in response to digitalization to effectively capitalize on the benefits offered by these technological advancements (Caputo, Pizzi, Pellegrini, & Dabić, 2021; Langley, van Doorn, Ng, Stieglitz, Lazovik, & Boonstra, 2021). Digitalization has triggered the rise and expansion of businesses employing diverse BMs (Ritter & Lettl, 2018; Trabucchi, Talenti, & Buganza, 2019) and compelling incumbent businesses to adapt their BMs to cope with the disruptions (Cozzolino et al., 2018). New BMs are also increasingly leveraging digital technologies to facilitate interactions and exchanges among businesses, consumers, and intelligent objects (Langley et al., 2020). For instance, Broekhuizen, Emrich, Gijzenberg, Broekhuis, Donkers, and Sloot (2020) emphasized the interconnections between digital platforms and suppliers, customers, service providers, product categories, and even distribution channels.

In the case of Fintech, digitalized services, utilizing cutting-edge technologies like blockchain, mobile phones, artificial intelligence (AI), machine learning (ML), cloud computing, and the Internet of Things (IoT), have been introduced as alternatives to conventional banking services (Gomber et al., 2018; Hornuf & Haddad, 2019). Lee and Shin (2018) outlined six categories of Fintech BMs: (i) payment BM (e.g., peer-to-peer (P2P) mobile payment, mobile payment services linked to digital wallets like Google Wallet, Apple Pay, and Samsung Pay), (ii) wealth management BM (e.g., automated wealth management or rob-advisors, online budgeting and financial planning), (iii) crowdfunding BM (e.g., rewards-based crowdfunding platforms like Kickstarter), (iv) lending BM (e.g., online-only loan brokers and websites), (v) capital market BM (e.g., trading Fintech platforms connecting investors and traders for commodities and foreign currency transactions), and (vi) insurance services BM (e.g., online-only insurance providers, insurance premium comparison websites). By eliminating the need for physical branches and automated teller machines (ATMs) and utilizing technologies to lower transaction costs, alleviate information asymmetry, and reduce inequality (Demir, Pesquè-Cela, Altunbas, & Murinde, 2022; Song, Han, Liu, & Ganguly, 2022), Fintech disrupts the financial services industry (KPMG, 2019;

OECD, 2021). The focus of this study is to expand research on BMs in the context of Fintech firms and examine how these firms create, deliver, and capture value while achieving digital sustainability.

### **Digital sustainability and business models**

Digital sustainability has been defined in the literature as an organization's ability to advance Sustainable Development Goals (SDGs) through the efficient and effective use of digitally enabled resources and solutions (George et al., 2021; Merello et al., 2022; Pan et al., 2022; Pan & Zhang, 2020). In essence, digital sustainability is based on both digital transformation and sustainability as mutually reinforcing elements of a business, in that these two aspects are intertwined (George et al., 2021; Pan & Zhang, 2020). Despite the fact that there is still no consensus on the measures of sustainability (Hutchins, Richter, Henry, & Sutherland, 2019), there is broad agreement that sustainability has three components: economic, social, and environmental sustainability, also known as the triple bottom line of sustainability focusing on profit, people, and the planet (Elkington, 1998; Liute & De Giacomo, 2022).

Economic sustainability refers to an organization's financial returns or profits, and the effects on the economic ecosystems (Brenner & Hartl, 2021). The impact of digitalization on economic sustainability is undeniable. For instance, Broccardo, Truant, and Dana (2023) found that digitalization supports sustainability, and consequently, profitability. Other studies have focused on the impact of digitalization on new economic models - the sharing economy (e.g., Geissinger, Laurell, Öberg, Sandström, & Suseno, 2020; Schiavone, Mancini, Leone, & Lavorato, 2021) and the circular economy (e.g., Schwanholz & Leipold, 2020), exploring the potential of digital technologies and platforms to facilitate sharing, leasing, recycling, and reusing existing products, thereby fostering long-term sustainability.

In terms of social sustainability, this often refers to the social benefits and ramifications of an organization's operations and how they add value to the communities within which it operates (Lee & Jung, 2019; Missimer, Robèrt, & Broman, 2017). For example, the study by Castaldi, Wilhelm, Beugelsdijk, and van der Vaart (2023) explores global value chains governance, revealing their effectiveness in ensuring the CSR efforts of international buyers toward enhancing the social sustainability of their suppliers in emerging economies. Focusing on the link between digitalization and sustainability (Broccardo et al., 2023), Di Vaio, Palladino, Hassan, and Escobar (2020) further examined the role of AI in fostering the development of sustainable BMs that align with the SDGs. Oderanti, Li, Cubric, & Shi (2021) highlighted the BMs that underpin the market development of more sustainable eHealth innovations catering to an increasingly ageing population. Digitalization also contributes to social sustainability by altering the conventional methods of service

provision and delivery, and ensuring the accessibility, efficiency, and affordability of products and services (Del Rio Castro, Gonzales Fernandez, and Uruburu Colso, 2021).

In relation to environmental sustainability, this pertains to corporate behaviors, strategies, processes, and policies aimed at mitigating and minimizing environmental impact, through measures such as reducing energy consumption and waste, and utilizing ecologically sustainable resources (Walls, Phan & Berrone, 2011). For instance, Manninen, Koskela, Antikainen, Bocken, Dahlbo, and Aminoff (2018) outlined a framework that assesses the environmental value propositions of circular economy BMs. Research has also investigated the role of key corporate governance actors, such as board of directors and top management teams, in achieving environmental sustainability outcomes (Aguilera, Aragón-Correa, Marano, & Tashman, 2021). However, the findings from Liute and De Giacomo's (2022) study suggest that their sample of UK-based B Corp companies perform better in terms of social sustainability compared to environmental sustainability. The study by Parmentola, Petrillo, Tutore, and De Felice (2022) further highlights the potential of blockchain technology in enhancing environmental sustainability, with many studies focusing on energy and utilities (17.3% of the articles they reviewed), logistics and supply chain (11.8%), agriculture/agri-food (8.7%), and Fintech/cryptocurrency (8.2%).

Additionally, there has been a growing body of research focusing on how BMs stemming from digital innovations can be used to promote sustainability (Bencsik et al., 2023; Oderanti et al., 2021). However, the challenge lies in how firms identify and implement digital technologies and innovations to support environmental, social, and governance sustainability (Guandalini, 2022). In recent times, scholars have shown increased interest in sustainable BMs (e.g., Baldassarre, Calabretta, Bocken, & Jaskiewicz, 2017; Bocken, Boons, & Baldassarre, 2019). Firms develop their sustainable BMs through various ways, such as by (i) increasing efficiency, (ii) exploring novel ways to foster business sustainability, (iii) placing greater emphasis on society and/or the environment, and (iv) being inherently embodying sustainability principles (Mignon & Bankel, 2022).

In the context of Fintech, recent studies have looked at how Fintech can improve financial innovation and sustainable development (Buckley, Zetsche, Arner, & Veidt, 2021). In order to achieve the SDGs, the United Nations has emphasized the significance of digital financial inclusion through services like mobile money, electronic payments, insurance, and credit that reach individuals who were previously excluded (United Nations Environment Programme, 2016). Arner, Buckley, Zetsche, and Veidt (2020), for instance, highlighted the link between sustainability, Fintech and financial inclusion. Fintech solutions further contribute to financial inclusion through

which Fintech reduces income inequality, particularly in higher-income countries (Demir et al., 2022). Merello et al. (2022) found that the market value of Fintech companies is determined by its sustainability, with Zhou, Zhu, and Luo (2022) demonstrating that Fintech innovation also positively affects an economy's green growth through green finance. This study expands on previous research on how Fintech can contribute to digital sustainability from a BM lens of creating, delivering, and capturing value.

## **METHODOLOGY**

This study used a case study approach (Siggelkow, 2007) and this was primarily motivated by two factors. First, given the exploratory focus of the study exploring a phenomenon that demands in-depth understanding, the case study was chosen as it enables a holistic exploration of events, interactions, and actions of various actors within the Fintech context. This is because the case study technique enables rich data collection "with many layered, intricate, detailed, nuanced to reach saturation with content validity" (Fusch & Ness, 2015: 1409).

Second, we used the case study approach since it helps us understand the subtleties of the phenomenon. In our case studies, they helped us to provide a better understanding of how the digital technologies lower the entry barriers for users to obtain essential financial services, and how multiple BMs are developed as cooperative endeavors with a variety of partners are initiated. To further understand the phenomenon, the case study method also enables the use of a variety of data collection sources (Strauss, 1987), and for this study, we gathered data using participant observations, internal meetings, and interviews that were conducted over a period of time to learn about value creation, delivery, and capture as part of BMs and how they are used to achieve digital sustainability. As a result, the likelihood of bias is reduced, and rich data that capture the intricate social and organizational settings can be gathered and analyzed (Yin, 2011).

We selected three exemplary Fintech startups (Firms A, B, and C) based in Taiwan to illustrate how digital technology and data are used to create value, and how these firms' BMs consequently help to achieve digital sustainability. The banking industry in Taiwan is worth trillions of TWD and is regulated by Taiwan's government. It accounts for 6.56% of Taiwan's GDP (\$762.67 billion USD) and 7.4% of all jobs in Taiwan. As of December 2022, Taiwan has 39 domestic banks with 3,433 domestic branches and 464 overseas branches or representative offices. The average distance to a bank branch and ATM is 3.7 and 1.32 km, respectively. For a small island with only 23 million people, Taiwan's banking industry is highly competitive.

Taiwan's Finance Supervisory Commission (FSC) announced the "Creating the Digital Finance Environment 3.0 Project" in January 2015, which relaxed restrictions

on online banking, particularly online applications (FSC, 2015a, 2015b, 2015c). The first stage of the project required banks to offer online financial services by the end of 2015 (Financial Supervisory Committee 2015a). Second, the FSC announced 11 big data application projects, including government open data (after deidentification) projects that involved data sets pertaining to real estate credit evaluation, stock market transaction data, personal credit card transactions, fraud statistics, and more than 900 other finance-related data sets (Data. Gov. Tw, 2015). Third, the shareholding ratio for investments into fintech companies by banks was relaxed from 5% to 100% (FSC, 2015c). Fourth, the FSC set up a fintech office, allocated funds for promoting fintech, and established a startup base (Financial Supervisory Committee 2016). Finally, the FSC published a fintech white paper in 2016 (Financial Supervisory Committee 2016). Taiwan's government aims to attract 5 billion TWD in overall funding for the fintech industry and facilitate the establishment of at least 30 fintech startups.

Firm A focuses on lending (P2P lending), Firm B focuses on insurance (user-based insurance or UBI), and Firm C focuses on investment (robo-advisory) services. These three businesses, which were established between 2015 and 2017, grew significantly within the first five years of their existence. The three case studies are based in Taipei, which is the primary hub for fintech startups in Taiwan. Two of them were incubated in a startup base supervised by Taiwan's Financial Supervisory Committee.

The three cases were selected because these businesses use digital technologies to evaluate investment risk, insurance ratings, and borrower creditworthiness rather than traditional financial risk management. Then, these digital technologies are utilized to collect crucial information, such as their customers' investment preferences and even driving habits, which may be used to create evidence-based profiles for their product and/or service offerings. Finally, these firms have expanded their range of offerings by partnering with other companies in order to not only achieve economic gains but also attain sustainability for the greater good.

### **Data collection and analysis**

In this study, we used a variety of data collection methods. We interviewed the CEOs of the three firms as our case studies to learn about their business milestones and how digital technologies have been used to lower the entry barriers for their users, particularly those who were previously unable to access such financial services. We also interviewed 6 senior managers, two from each company, in order to further understand how digital technologies, particularly AI, are used to assess personal risk in credit, insurance, and investment portfolios. In order to learn more about how these organizations collaborate, we further conducted interviews with 17 partner managers of businesses with which these companies work together in order to learn how they



collaborate. We then triangulated and validated our data from these different data collection methods.

In addition, we have three other researchers who did their summer internship in each selected Fintech start-up in 2022 to learn how to assess risk for a variety of demographics. These researchers also took on the roles of consumers to understand the automated process of using specific financial services. We also used a variety of data collection sources to better comprehend the three cases. In the case of Firm A, we also joined the Dcard online forum for college students to discover more about their stories and experiences with P2P lending, in their hopes to get money they required in real time. In the case of Firm B, we interviewed 22 car drivers to understand the benefits and disadvantages of user-based insurance (UBI). In the case of Firm C, we interviewed 25 of their clients to learn about their investing habits and to get their opinions on how robot advisors differ from other financial advisory tools. We also conducted expert meetings and workshops to further clarify our contributions to digital sustainability. In summary, the cross-case analysis and comparative case study involving the use of multiple sources enriched our understanding of digital sustainability. In total, we accumulated 205 hours of data collection, as shown in Table 1.

-----  
Insert Table 1 about here  
-----

Throughout the data collection period, we documented all our interviews, observations, and meetings using field notes. The interviews were conducted in Mandarin Chinese, and all transcripts of them were recorded and afterwards translated into English. The transcripts, both in Mandarin Chinese and English, were then shared and reviewed by the CEOs of the three Fintech startups to verify the accuracy and, consequently, the reliability of the data.

Our data analysis was performed as a three-step thematization procedure (Miles et al., 2014) involving data condensation, data display, and conclusion drawing. In the data condensation phase, we examined each transcript to identify the key data driven by digital technologies. We then examined how each fintech company made good use of their data to innovate their financial services.

Lichtman (2013) highlighted the existence of two types of codes, namely predetermined (a priori) codes and codes that emerge from data. With existing business model constructs serving as a theoretical foundation, we developed emergent codes to illustrate value creation, value communication, and value capture of business model innovation by fintech companies. We then iteratively grouped the codes that were based on the same notion into first-order concepts (Gioia et al., 2013), enabling us to reduce the initial categories to a more manageable set of five categories (Figure 1).

In the second phase, namely data display, we identified the connections between and among the first-order concepts to create second-order themes (higher-order themes). These second-order themes represent broader theoretical categories that align with the theoretical foundation and embody the various elements of business model that were evident in our data. For example, the first-order concepts of incorporating AIoT technologies into cars and collaborating with car manufacturers, insurance companies, and fleet management service providers are concepts that exemplify the theoretical foundation of value creation for business model innovation. We continued to analyze the connections between the first-order concepts and eventually identified three second-order themes, which depicted the relationships between organizing practices and the constructs of business model innovation in the context of fintech.

In the third phase, namely conclusion drawing, we compared and added further details on the basis of our observations as participants (i.e., summer internships). We then systematically combined the second-order theoretical categories into aggregate dimensions (Corley & Gioia, 2004; Gioia et al., 2013). Notably, we observed the creation of value for digital sustainability in the areas of economic sustainability, social sustainability, and environmental sustainability. The coding process involved a recursive analytic procedure (Locke, 1996) instead of a linear approach, and it was performed until we acquired a clear understanding of the emerging theoretical relationships and until no new themes could be yielded from additional interviews. Figure 1 presents the coding results, namely the first-order concepts, second-order themes, and aggregate theoretical dimensions pertaining to how fintech startups innovate their business model for digital sustainability.

-----  
Insert Figure 1 about here  
-----

## **FINDINGS**

Table 2 provides an overview of how value is created, delivered, and captured, as well as the outcomes pertaining to economic, social, and environmental sustainability by Fintech firms. The findings are aggregated into three key components of a BM, centered around: (i) the creation of complementary value propositions, (ii) complementary value delivery mechanisms in terms of co-supporting, co-designing, and co-engaging mechanisms to users and partners, and (iii) the integration of complementary revenue streams for value capture. Each of the examined cases also presents compelling evidence highlighting how these elements of value creation, delivery, and capture contribute to the attainment of digital sustainability.

-----  
Insert Table 2 about here  
-----

The dimension of economic sustainability encompasses various themes that contribute to enhancing the efficiency and effectiveness of financial service provision. These themes include streamlining transaction processes, reducing transaction costs, providing accurate assessments, mitigating information asymmetry, ensuring information accuracy, and facilitating the provision of customized product and service offerings. The dimension of social sustainability focuses on fostering inclusive financing of underserved segments in terms of accessibility and the provision of affordable financial service offerings, promoting social empowerment, ensuring safety through digital financial services, promoting a sense of responsibility, supporting health and wellbeing, and reducing inequality. Additionally, our findings indicate that firms also achieve environmental sustainability through several means. These include reducing the consumption of natural resources, promoting sustainable transportation practices, reducing air pollution, and increasing investments in environmental, social, governance (ESG)-related stocks or exchange-traded funds (ETFs).

#### **Firm A**

Fintech Firm A was founded in 2017 and it uses digital technologies combined with Intellectual Property (IP) to offer a P2P platform, providing credit risk assessment and automated lending services. The AI technologies introduced by Firm A enable college students and other underrepresented groups that were previously neglected by traditional banks obtain loans within 24 hours. It is also a firm that provides loans to SMEs quickly and more affordably.

***Complementary value propositions as value creation.*** Firm A effectively integrates and harnesses digital technologies, including AI technologies, automation, and mobile device connectivity, to facilitate credit risk assessment and the provision of its automated lending services. Through the implementation of these technologies, Firm A overcomes the barriers that impede consumers and lenders from accessing essential financial services. The P2P platform operated by Firm A acts as a vital link connecting potential lenders and borrowers, thereby creating opportunities for individuals and enterprises that previously faced challenges or were unable to obtain loans. The presence of a greater number of lenders on the platform also provides users with more options, thus extending services to those who were previously excluded or encountered financial disadvantages. In this way, Firm A creates value propositions in collaboration with its users and partners, not only in terms of achieving economic sustainability but also by reducing transaction costs to foster social sustainability.

Firm A utilizes and leverages its technology to also enhance process efficiency, thereby contributing to economic sustainability. The company's innovative technologies facilitate swift funding for customers through its P2P platform, with funds typically being disbursed within 24 hours. Additionally, the technology in place also improves the accuracy and effectiveness of loan processing by enhancing the overall loan evaluation process. Firm A has also implemented a comprehensive risk assessment system using technologies to detect fraud, facilitate person registration, and identify anti-laundering occurrences, all integrated into automated workflows for credit evaluations to achieve economic sustainability.

The data captured through these processes is further utilized to derive additional benefits. Indeed, during the initial three years following its establishment, Firm A focused significantly on data collection and aggregation methods to support the credit rating process and evaluation. It employs technology for risk assessments and basic credit lines evaluation for its customers, and it also analyzes various types of data gathered from diverse sources to evaluate loan applications and mitigate the risk of default. For every application, data related to an applicant's occupation, qualifications, current employment, and other relevant details, is gathered to precisely evaluate an applicant's profile. Firm A also acquires non-financial data, such as criminal records or social behaviours, to enhance the credit evaluation process and ensure a more comprehensive dataset. Based on the assessed risk level, they are then charged interest rates ranging from 2% to 15%. This approach enhances the accuracy and efficiency of the lending process while reducing the likelihood of default across different risk levels. In this regard, Firm A demonstrates responsible practices by targeting customer segments that were previously ignored by traditional banks and offering inclusive financing to them. The company also establishes stringent procedures to mitigate adverse societal effects, such as the inability to repay loans. As a senior manager of Firm A explained:

*Our P2P lending platform has 13 risk grades. We collect personal ID data, together with their shopping and payment bills, their living standards, and other social behaviors to make the risk assessment process more accurate and efficient. In these days, we also develop the warning system for potential bad debts, such as the length of lending periods, particular months to lend, and social identity provided. All these data-driven works (from deep learning) help us to make our risk grades more accurate.*

In addition to attaining economic and social sustainability objectives, Firm A leverages its technologies to minimize bias in risk awareness and evaluation systems by reducing human involvement in loan processing and evaluation. This utilization of technology

enhances the overall the effectiveness of the system. The firm also employs advanced technology to collect and analyze a wide range of data, including applicants' personal information sourced from social media, news outlets, and publicly accessible platforms. This comprehensive data assessment assists in determining applicants' overall creditworthiness, thus contributing to the achievement of economic sustainability. The CEO provided the following explanation of the novel application of digital technologies in automating the loan process:

*The first and the most challenging part of automated lending process is the authenticity of personal data. We require borrowers to take several pictures of themselves of their own free will. It means that you are not threatened or cheated to borrow money! We have IP with automated assurance of personal pictures. Also, we have AI technologies to search for personal data, such as FB (Facebook), IG (Instagram), news and so on to triangulate the accuracy of personal data. In conclusion, we want to make sure that they are good customers without criminal records.*

The utilization of Firm A's P2P platform by a substantial number of lenders and borrowers presents opportunities to eliminate intermediaries and mitigate administrative inefficiencies associated with such services. Moreover, the platform's capacity to directly connect lenders and borrowers while leveraging technologies, big data, and real-time information enhances the potential for minimizing waste and environmental harm. One of the senior managers of Firm A's client companies noted that Firm A's technologies substantially reduce paper usage, thereby contributing to environmental conservation through the reduction of the natural resource consumption, as follows:

*The digital technology really helps to reduce the use of papers. We have estimated that all the paper accounts in our bank, when accumulated, to be as tall as Taipei 101 for the past 10 years.*

In essence, our findings reveal the complementary value propositions that effectively support economic sustainability. These value propositions attract a large number of lenders and borrowers to their platform by enhancing process efficiency, reducing transaction costs, improving the accuracy of credit risk assessment, and mitigating the risks of bad debts. At the same time, Firm A utilizes and leverages its technology and data assets to drive social sustainability by fostering inclusion among diverse lenders and borrowers and providing them with affordable financial services. Firm A also actively contributes to environmental sustainability by conserving resources as well

minimizing paper usage and waste through the strategic implementation its technologies and data-driven practices.

***Co-supporting value delivery mechanism.*** Firm A's provision of inclusive financing aims to enhance both process efficiency and accessibility to financial services for underserved customer segments by reducing the lending barriers and credit thresholds that previously hindered these segments from accessing such services. In the case of Firm A, the delivery of value occurs via a co-supporting mechanism that caters to the underserved customer segments. This mechanism is established between Firm A's P2P platform and diverse social community platforms, including Dcard, a widely utilized social community platform among Taiwanese college students to seek assistance and connect with fellow students, and PTT, an online forum for discussions on social issues and entrepreneurial endeavours.

The complementary mechanism between Firm A in partnerships with these social community platforms is used to deliver value to specific target customer segments, one of which is the college students who were previously not the focus of conventional banks. Firm A leverages its connection to Dcard, using it to forge relationships and advertises their product and service offerings to college students. Firm A also utilizes its technology-driven and data-driven P2P platform to evaluate 'stories' that are posted by college students on Dcard. These postings or 'stories' may be about the needs of some students needing to pay for their college tuition, take part in short-term exchange programs, enroll in supplementary courses to gain certificates in, for example, software or languages, or even seek medical care following, say, an accident. Firm A gathers information from these postings using its technology and big data analytics to assess whether these 'stories' are true. If they are, Firm A can then support these needs by giving short-term loans to college students. In this regard, Firm A utilizes its technologies to reduce the individuals' or even communities' financial struggles and poverty (social sustainability) by providing them with relevant financial product/service offerings. The CEO of the firm further explained why college students are a crucial market for the firm:

*College students are potential markets for financial services. As over 70% students have stable income from part-time jobs or parents' sponsorship, with \$20,000-25,000 NT dollars as monthly income on average, we consider them as good customers, often with urgent financial needs. In the long run, we also try to build life-long relationships with them. As they grow up, they may need cars, housing, and furnishing offerings. Our platform hopes to become their first priority to provide them with essential financial services.*

In addition to college students, among the primary target markets of Firm A are individuals who work as engineers in their professions. This customer segment often has strong credit backgrounds; however, they may still encounter situations where they require quick loans to address unforeseen financial issues in their personal or professional lives. The conventional banking system, characterized by rigorous requirements and stringent loan evaluation and approval procedures, frequently leaves this segment feeling disempowered. Additionally, owing to their strong credit histories and potentially higher financial capabilities, individuals in this occupation also have the opportunity to supplement their income by acting as lenders on the platform. They do so by offering more affordable interest rates compared to traditional banks, effectively facilitating increased opportunities for loan provision and enabling borrowers to access more affordable financial services. This mechanism contributes to the promotion of social sustainability, as one of Firm A's senior managers observed:

*As compared to college students, engineers can get approved for loans quickly within 24 hours. They require shorter time and larger amounts which make them become popular to draw more investors to our P2P platform. They also sometimes become investors as well.*

Another customer segment that is targeted by Firm A is the SMEs, including startups, e-commerce stores, and even influencers that engage in direct selling. This customer segment holds significant importance for the firm due to their inclination towards establishing enduring relationships and their desire for a variety of product and service offerings. Although SMEs have often used traditional banks, their revenue base often is not significant enough for banks to pay attention to them. Firm A targets this market by not only meeting their financial demands but also helping them manage and keep track of their financial situation, as indicated by one of Firm A's senior managers:

*SMEs in Taiwan prefer to build longer relationship with suppliers, customers, and banks. They take it as comparative advantages to reduce transaction cost. However, they also need real-time monitoring system and we can use this to learn about their repayment capability in time. Our firm's AI technology can help us to check SMEs' monthly repayments, just like the Just-in-Time system in financial services.*

The findings indicate that Firm A utilizes its platform to deliver value by co-supporting social community platforms, as well as the partners and users engaged on the platform, with a specific emphasis on college students, engineers, and the SMEs. Through this mechanism, Firm A's offerings improve the efficiency and effectiveness of financial service provision while lowering transaction costs, thereby promoting economic

sustainability. It also contributes to social sustainability by fostering inclusive financing practices and empowering individuals and SMEs facing poverty and/or financial difficulties.

***Complementary revenue streams to capture value.*** Firms also need to have profits, i.e., capture value, from the delivery of their value propositions. Firm A captures its value from the P2P platform based on transaction fees. All participants on the P2P platform, including borrowers and lenders, must pay transaction fees. The borrowers are further charged transaction interest rates ranging from 1-3%. These interest rates are affordable and typically less expensive than those provided by traditional banks. This implies that Firm A is not just concerned with maximizing financial gains but also making sure that the target customer segments can afford the services offered.

The second revenue stream relies on contractual agreements involving IP licensing fees. Firm A extends its product and service offerings to traditional banks, assisting them in the development of automated risk assessment methods to enhance the efficiency and expediency of their internal lending and auditing processes. In return, Firm A levies IP licensing fees based on the terms outlined in the contracts. Additionally, Firm A engages in collaborative partnerships with banks as its clients across multiple service provisions, thereby enabling the capture of value from various projects. The CEO of the firm noted the following:

*It really takes time to help the elephant lose some weight and learn how to dance with AI! Besides, they also need to learn the new products development know-how with AI technologies. We have \$ 1-2million NT dollars of consultant fee at first year and also have other revenue sharing provisions in relevant projects.*

The provision of a range of products and services by Firm A enables it to capture value from its customers, including the SMEs, by giving them the flexibility to opt for instalment-based payment plans (BNPL) when purchasing products. By offering this payment option, Firm A captures value through the collection of instalment fees, enabling individuals and SMEs to acquire necessary products and continue operating their businesses without the burden or complexities associated with upfront charges. Essentially, this revenue stream simplifies the process for clients to access financial services by facilitating seamless rental and payment options, which are often more accessible, affordable, and convenient than those offered by traditional banking systems, hence fostering social sustainability.

Firm A's strategic collaborations encompass a diverse range of partnerships, which also serve as a means to leverage shared data. A notable example pertains to the firm's collaboration with various government agencies to acquire pertinent data that



could be used to evaluate the creditworthiness of SMEs and individuals, thereby expediting loan approval procedures. The data, in turn, facilitates the provision of relevant product and service offerings such as vehicle loans, low-interest instalment plans, and charitable donation services. The shared data is also utilized to evaluate one's borrowing terms, including loan amounts, loan terms, and interest rates, across a wide range of financial service offerings. Consequently, the complementary nature of the collected data progressively enhances Firm A's risk assessment accuracy, enabling it capture value through the streamlining of the process, safeguarding the lending process, and mitigating any payment issues.

In essence, Firm A's BM demonstrates the value propositions derived from the complementary utilization of technologies and data with users and partners through its P2P platform. It co-supports social community platforms and leverages the acquired data to provide affordable financial services, thereby addressing financial struggles and poverty concerns prevalent in society, in the hope of promoting economic growth. In terms of economic sustainability, Firm A improves the efficiency and effectiveness of financial services and reduce the bad debt rates with AI-driven risk assessment tools. For social sustainability, Firm A increases the accessibility of financial services for underserved customers and reduce the financial struggles and poverty. Additionally, the provision of financial services by Firm A aligns with environmental sustainability objectives, as it significantly reduces the consumption of natural resources and enhances the efficiency of credit risk assessment and lending processes.

### **Firm B**

Firm B, established in 2015, partners with car manufacturers, insurance companies, and fleet management service providers, to deliver integrated services aimed at capturing driving behaviour data and assessing accident rates. Leveraging its data assets, the Firm offers usage-based insurance (UBI). UBI premiums are determined based on the principle of 'how much you drive', shifting from fixed charges to variable fees that consider factors such as the kilometres/miles driven and other driving behaviour characteristics. By considering the frequency of vehicle usage throughout the insurance policy term and the driver's history of traffic incidents or violations, Firm B can accurately calculate discounts for cautious drivers while charging higher premiums to clients exhibiting more aggressive driving behaviours.

***Complementary value propositions to create value for Firm B.*** Firm B employs a range of novel technologies as their value propositions. The integration of AI within their UBI offering, combined with the IoT, is considered crucial by Firm B for evaluating different driving scenarios and accurately determining UBI premiums. These technologies, embedded within the UBI offerings, contribute to a unique service

experience for customers. For example, they facilitate the timely recording videos in the event of a car collision, aiding in the determination of who was at fault. The UBI value proposition, supported by various technologies, effectively enhances the accuracy of insurance premium calculations (economic sustainability), promotes safety and a sense of responsibility of drivers (social sustainability), and has the potential to alleviate traffic congestion (environmental sustainability), as mentioned by the firm's CEO:

*Car drivers get used to have video-recorders as the essential equipment for driving safety. But they do not have the advanced tools to help them monitor their driving behaviors and the road conditions ahead. Our products have proven to be the advanced devices for drivers.*

To create value, Firm B leverages a range of AIoT technologies in collaboration with car manufacturers and insurance companies. One such technology employed by Firm B is On-Board Diagnostics (OBD), which provides real-time data to ensure accurate information on vehicle problems and malfunctions. Another AIoT technology is the On-Board Unit (OBU), an electronic device integrated into vehicles and connected to roadside infrastructure and satellite navigation systems. Firm B also effectively leverages the Digital Video Recorder (DVR), which incorporates advanced telecommunications and location precision technologies such as Global Positioning Systems (GPS) and Geographic Information Technology (GIS) system. These technologies enable round-the-clock monitoring of vehicles, ensuring the security and safety of drivers and passengers while also providing reliable evidence in circumstances where such recording is required. These technologies are also used to enhance the scoring/rating system for car insurance premiums.

The application of these AIoT technologies thus reinforces Firm B's value proposition to assess driving behaviours accurately, reduce information asymmetry, ensure accurate accident recall for insurance claims through recorded data, and improve the accuracy of the overall claims evaluation process, thereby contributing to economic sustainability. It also fosters a sense of safety and responsibility through the digital financial services offered (social sustainability), as noted by one of the firm's senior managers:

*DVR has gradually become the standard configuration of public transportation such as taxis and buses. Serving as in-vehicle monitoring, it not only ensures the safety of passengers in the car, but also provides strong evidence for some cases. DVR also promotes the development of intelligent transportation.*

Firm B also integrates its technologies with Advanced Driver Assistance Solutions (ADAS) to cater specifically to high-end drivers. The ADAS system incorporates

algorithms that can accurately identify and report various traffic violations, including detecting vehicles that are wrongly parked, and those making wrong turns. To leverage the data gathered through this technology, Firm B partners with an established insurance company to offer a streamlined ‘one-click’ insurance service. This service allows for real-time video recording of car collisions within 15 seconds, facilitating real-time online insurance claim processing and contributing to economic sustainability.

Furthermore, Firm B’s embedded services also enhance traffic management and reduce congestion and air pollution by promptly activating real-time ‘rescue’ and car-towing services in the event of accidents. This improvement in service efficiency not only contributes to economic sustainability but also supports environmental sustainability efforts. Through emphasizing safety during driving or commuting, Firm B’s value propositions effectively contribute to social sustainability by providing a more reliable car journey experience, while also fostering a sense of responsibility among drivers. Overall, these complementary value propositions demonstrate Firm B’s commitment to creating value through its offerings, as explained by the CEO of the firm:

*The technologies and the more data we learned, the more visibility we saw, the more accurate the insurance scores are ... We learned a lot from the Tier 1 electronic company to leverage AIoTs. For example, Tesla launched a scoring system which is calculated by analyzing the customers’ real-time driving behaviors based on five criteria: instances of forward collision warnings, hard braking, turning corners aggressively, unsafe following, and forced autopilot disengagement. The scoring system was first used to determine whether a driver could access their cars’ full self-driving option, and it will now be used to underwrite Tesla drivers who buy its insurance.”*

In addition to embedding technologies into their service offerings, Firm B collaborates with diverse platforms to collect and utilize data that can be leveraged to calculate and assess a driver’s driving scores. The database contains comprehensive information regarding a driver’s behaviour and driving patterns, encompassing variables such as total distance travelled, average speed, frequency of driving during different times of the day, driving on weekdays versus weekends, instances of exceeding the speed limit, and even frequency of abrupt accelerations or decelerations, and the occurrence of sudden turns. These driving behaviour data can be subsequently utilized to generate a driving score and determine the appropriate insurance premium for individuals seeking car insurance. A senior manager from the company emphasizes the significance of such

data, as it ensures the accuracy of insurance premiums and facilitates the provision of relevant insurance products (economic sustainability).

*The maintenance data also provide great evidence to learn the drivers' behaviors. The condition of vehicle tires, the engines, and the overall conditions of the car all become key evidence to triangulate with our OBD to provide relevant insurance products.*

**Co-designing value delivery mechanism.** Firm B effectively delivers its value propositions by engaging in co-design processes with car manufacturers, insurance companies, and fleet management service providers. This complementary co-design approach enables Firm B to integrate its digital technologies directly into the standard equipment of vehicles, resulting in enhanced safety and convenience for drivers, ranging from occasional drivers to premium or high-end drivers, and those categorized as frequent drivers who spend substantial time on the road. The embedded technologies offered by Firm B assist in various aspects, including object detection, monitoring traffic conditions, and providing relevant traffic information. These technologies prioritize the safety of not only the drivers but also passengers, pedestrians, and other motorists, as noted by a senior manager of Firm B:

*This could be on drivers' drowsiness or parking assistance, safety alerts, blind spot detection, back over protection or back over prevention, etc. Together with ADAS, it would involve actuating the subsystems in case the initial warnings are ignored for some reasons. The prime example of this is system actuating brakes by itself to slow down a vehicle to avoid collision.*

Furthermore, Firm B co-designs its value delivery mechanism with insurance companies to offer more competitive and cost-effective insurance plans, along with efficient and accurate claim assessments and offering a 'pay-as-you-go' or on-demand insurance services. Firm B also engages in the co-design process with fleet management service providers, who serve as valuable partners in the delivery of its value propositions. Leveraging its vast array of data sources, Firm B facilitates access to comprehensive data for fleet management service providers. This includes crucial information pertaining to driving behaviours of fleet drivers, driver positioning, and historical data of roadside conditions. Firm B utilizes this data to then establish a fleet management call center service, with an aim to offer fleet management providers affordable yet comprehensive insurance packages that cater specifically to their needs. The firm also actively collaborates with small and micro car rental companies, co-designing service offerings to assist them in effectively managing their vehicle fleets

through bundled insurance contracts and supporting them with marketing campaigns aimed at supporting these small businesses.

Additionally, Firm B targets the younger generation as a customer segment, acknowledging that they may either lack the means or choose not to own a car. This customer segment also places a higher value on convenience over the hassles associated with car ownership. To deliver value to this segment, Firm B creates an online car rental app or platform, leveraging its technological expertise and data-driven competitive advantage. Using the partnerships with car rental companies and insurance providers, Firm B co-designs a comprehensive and streamlined rental experience for the users, providing them with a one-stop solution through this platform, as explained by the CEO of the firm:

*Instead of meeting insurance clerks, the Z generation (born after 1995) prefers to have on-line orderings with automated services at their convenience. As they get used to rent-as-you-drive services, they also expect to have UBI (Usage-based insurance) just in time as they rent a car.*

***Complementary revenue streams to capture value.*** To capture value, Firm B engages in partnerships with multiple partners to generate revenue from various channels. The predominant revenue stream stems from transactions, including the sale of its OBD products. As previously stated, Firm B collaborates with car manufacturers to integrate OBD equipment as a standard feature in recently purchased vehicles. Additionally, customers have the option to directly purchase the product from an e-commerce platform. In this case, the BM of Firm B heavily depends on transactional proceeds arising from the sale of this product, which is used to enhance driving safety. A senior management in Firm B commented on this:

*In the future, the OBD would be the standard accessories for cars. The OBD-inside would become our core business model. In order to increase our competitive advantage, we have to enhance our system stability while also upgrade our database and real-time alarming system as well.*

The second source of revenue involves generating income through contractual arrangements related to its UBI services. This includes subscription fees for fleet management – each vehicle is equipped with GPS navigation, real-time alerts, and insurance services, which serve to reduce information asymmetry and provide accurate assessments particularly when assessing insurance claims (to ensure economic sustainability), as well as contributing to the safety and wellbeing of drivers, passengers, and other commuters (to promote social sustainability). Firm B also captures value from these insurance contracts by collaborating with established insurance companies to

provide streamlined ‘one-click insurance claim’ services, leveraging the sharing of more precise data such as driving conditions, thereby capturing value from these service firms. This is highlighted by a senior manager from one of Firm B’s partner insurance companies, noting the importance of these technologies for drivers:

*The AIoT-inside services would become part of the car insurance contracts as more and more drivers are more conscious about themselves. At the same time, the logistic companies and the delivery services also take the AIoT-inside insurance as necessity.*

Another source of revenue stream stems from its car rental services offered through its platform. Firm B levies charges on both the renters utilizing the service and the car rental agencies, thereby capturing value from the intricate dynamics of supply and demand. In the case of the renters, Firm B captures revenues through the inclusion of services embedded within the UBI contracts. As for the car rental agencies, Firm B charges a subscription fee for the fleet management services it provides. The platform essentially serves as a comprehensive, one-stop shopping service for its customers and partners, facilitating a seamless and convenient experience, as a senior manager of the platform explained:

*It’s more like the ‘pay-as-you-go’ services for drivers to have short-term driving experiences in Taiwan. We selected cars within 5 years to have our T-BOX services and provide remote-control, video-recording, real-time warning, and positioning services for drivers. We believe that the car-sharing service will be the dominant model in the near future.*

In essence, Firm B creates, delivers, and captures value by means of co-designing partnerships with its partners, while offering a range of products and services tailored to a variety of customer segments. The firm captures revenue through complementary revenue streams based on transaction-based and contract-based arrangements, subscription fees, and rental service fees. Distinctive technologies are integrated into its product and service offerings, and the firm actively engages in co-designing additional offerings, such as streamlined ‘one-click insurance claim’ services. Firm B also leverages its technological capabilities and data resources to build a novel car rental platform, facilitating the collection of relevant driving behaviour data and enhancing the provision of relevant product and service offerings.

In terms of economic sustainability, Firm B reduces the information asymmetry, improves the efficiency of UBI financial services and provides accurate assessments with accurate driving data. For social sustainability, Firm B ensures the driving safety through digital services, promoting a sense of responsibility for safe driving, and

supporting drivers' health and well-beings with the AI-driven co-pilot services. It also enacts inclusive finance by increasing the accessibility and affordability with appropriate UBI and car maintenance packages to small and micro rental companies and to those who are not able or not willing to own a car. For environmental sustainability, it reduces air pollution with less traffic congestions.

### **Firm C**

Established in 2017, Firm C specializes in providing robo-advisory (RA) services, which encompass an automated investment platform that utilizes algorithms to manage investors' portfolio. The RA platform provides a comprehensive range of services, including personalized risk assessment, investment planning, portfolio design, investment management, portfolio monitoring, and dynamic withdrawal capabilities. These offerings collectively contribute to mitigating the barriers faced by inexperienced (and even experienced) investors when investing in local and global stocks. Firm C also extends its expertise using its AI-driven technology as a provider of investment advisory services, catering to business-to-business (B2B) and business-to-consumer (B2C) markets and offering tailored investment plans.

***Complementary value propositions to create value for Firm C.*** Firm C utilizes and leverages RA to assist investors in assessing their investment risk preference and identifying and managing their investment portfolios. These RA services encompass various services such as determining investor profiles, allocating assets, implementing investment strategies, rebalancing portfolios, and reviewing and reporting performance. Compared to conventional human portfolio management systems, the utilization of RAs by Firm C confers strategic benefits, including reduced transaction costs, lower minimum investment thresholds, and more efficient and transparent workflow. Through the utilization of such digital technologies, Firm C not only enhances the efficiency of the process but also expands accessibility and affordability to a broader range of individuals, including lowering barriers to entry for novice investors, as noted by the CEO of the Firm:

*We try to lower the entry barriers for local and global investments. With robot advisors (RA), you just need to fill in the risk assessment list, make the investment plans, and the RA will help you construct your own investment portfolio automatically.*

Firm C creates complementary value propositions that encompass both economic sustainability and social sustainability by reducing the entry barriers to investment opportunities. By providing a range of services in the form of service 'modules', Firm C delivers comprehensive investment portfolio management without bias. In contrast,

financial consultants working with clients may lack the requisite knowledge to provide a wide range of services, thereby risking the provision of erroneous advice with some even occasionally exaggerating promises of returns when offering recommendations. Leveraging their technology, Firm C swiftly acquires pertinent data to evaluate and analyze the risk profile of potential investors within minutes and provides clients with more precise and impartial recommendations, facilitated by algorithms and the consideration of their risk profiles. This streamlined process enhances efficiency and promotes economic sustainability, as highlighted by one of the senior managers of Firm C:

*In order to profit more from the stock markets or other financial markets, financial consultants may encourage the HNWI (high net-worth individuals) to invest in some unfamiliar financial assets, such as the hedge funds or derivatives, and claim to have 20-30% profit return in 3-6 months. On the contrary, the robo-advisors never encourage investors to achieve high profits. Instead, we encourage stable and long-term profits with 5-15% returns.*

Furthermore, the technology is utilized to provide guidance for portfolio reviews, advising investors regarding their investment positions either on a seasonal or annual basis as part of the investment evaluation process. Investors are also presented with a series of other questions, including topics such as planned annual family vacations, retirement plans with specific monthly payments, or aspiration to own a home within a five-year period, helping them in long-term financial planning and contribute to the reduction of income disparities. Furthermore, investors lacking knowledge about investment trends and strategies are supported throughout the process in a hope to reduce inequalities, particularly in relation to the levels of knowledge disparities. Consequently, the streamlined process of accessing investment advice and managing investment portfolios, facilitated by Firm C's services, enhances accessibility and inclusion of individuals who were previously excluded from having the opportunity to engage in investment activities.

***Co-engaging value delivery mechanism.*** Firm C delivers value through a co-creation mechanism, which involves fostering collaborations between Firm C and various platforms that cater to the target segments. One of these platforms is a well-known 'mom-platform', dedicated to addressing maternal concerns relating to pregnancy, postpartum recovery, and hiring babysitters/childcare services by offering comprehensive solutions. The partnership with the mom-platform helps Firm C to leverage the platform's customer relationship management (CRM) to gain greater



insights into user behaviour, and co-design relevant product and service offerings, as noted by one of the top executives of the platform:

*We are the Agoda of the mom-and-baby market. We help to solve mothers' problems relating to pregnancy, postpartum confinement, and hiring babysitters. All these information together with [Firm C] tells a lot of how mothers deal with their expenses.*

The platform currently boasts a substantial user base, and Firm C is utilizing the data gathered from this platform to offer RA services tailored to mothers and expectant mothers seeking information regarding their children's and families' financial matters. The offerings are co-designed and embedded within the product and service provisions offered on the platform. For example, expectant mothers frequently rely on the platform to assist them in organizing their postpartum hospital visits and doctor appointments. To enable the technology to assist them effectively, these expectant mothers are required to provide their maternal health books containing personal details such as their name, date of birth, identification number, weight, blood pressure, and expected delivery date. Drawing upon the submitted data, the platform can subsequently offer suitable appointments and personalized financial packages catering to various stages of family life. Within a span of 7 minutes, a report featuring graphics and illustrations can be promptly generated and made accessible to these mothers, highlighting the efficiency and customization of the service provision (economic sustainability).

Furthermore, Firm C co-designed with the platform to provide 'value packages' through the facilitation of expert-led classes, including seminars on parent-child relationships, skincare, and meditation. In this context, Firm C would provide financial advisory services, while the platform would offer nonfinancial services, with both services being tailored specifically to meet the needs of mothers and parents. Thus, Firm C facilitates not only the efficiency of the process (economic sustainability) but also fosters social empowerment among mothers and expectant mothers. It empowers them to build their financial literacy, helping them in long-term planning to mitigate any income disparities (social sustainability). Furthermore, the automated process facilitated by the platform reduces the consumption of natural resources by minimizing the need for paper printouts. This practice significantly contributes to environmental sustainability.

Firm C also engages in co-designing its offerings in collaboration with a renowned stock tracking platform. Through this partnership, Firm C acquires data, wherein the platform's users are provided with access to financial indices and news. Subsequently, Firm C enhances the significance of such information by providing advisory services, utilizing tools such as cost-benefit analysis to assist investors in

effectively managing their investment portfolios, thereby enriching the user experience. Notably, these services encompass the inclusion of environmental, social, and governance (ESG)-related stocks and EFTs as investment recommendations, underscoring the focus of Firm C in sustainability investment strategy (environmental sustainability). This co-designing approach to value delivery improves process efficiency and transparency (economic sustainability). It also serves to bridge knowledge gaps regarding investments, enabling investors to quickly acquire knowledge about investments with the assistance of technology, thereby contributing to the long-term reduction of inequality in investment knowledge and practices (social sustainability).

In essence, Firm C's product and service provision involves a thorough understanding of their target segments, encompassing diverse life cycle stages and risk preferences. This understanding allows for the co-creation of customized investment packages in collaboration with platform partners, ensuring a prompt and efficient delivery without reliance on paper-based processes. Firm C introduces innovative RA services designed to cater to a wide range of investors, including both novice and experienced individuals. These services assist investors in bridging knowledge gaps, managing their investment portfolios, and providing recommendations for sustainability-focused investment strategies.

***Complementary revenue streams to capture value.*** Firm C employs a strategy characterized by the utilization of partnerships with multiple platforms to capture complementary revenue streams. Firstly, the firm captures value from the transaction fees from its RA platform, where members pay monthly charges to utilize the RA services. Secondly, Firm C captures revenue through contractual agreements with an investment/stock trading platform, where members subscribe to receive investment advisory services alongside the offerings provided by the stock trading platform. Firm C's RA technologies are used to complement the services provided by the investors' platform to provide advisory services, thereby strengthening the relationship between the platform provider and its users. The incorporation of Firm C's technologies allows for the provision of comprehensive investment insights, presented through graphics and illustrations, which surpass the capabilities of the platform provider alone. Consequently, Firm C benefits from a revenue stream through fee sharing, whereby it receives a percentage ranging from 20% to 30% when customers subscribe for bundled services. A senior manager from an investment advisory company on the investment platform noted the following:

*Traditionally, we are labor-intensive consultants who provide personal investment services to our customers. We build tight and long-term relationships with our customers. However, we also need professional*

*investment reports with rich graphics and illustrations to convince our customers. The robo-advisors with cost-benefit analysis are excellent tools to enhance our proficiency.*

Thirdly, Firm C engages in partnerships with advisory and asset management firms, enabling the company to provide specialized asset management services that encompass services such as estate and heritage planning, tax arrangements, and investment consultancy. By offering these specialized services, Firm C captures value from consultancy fees, effectively capturing value from both the B2C and B2B markets. This expansion of services enables the company to assist other businesses in reaching untapped demographics that may not have been targeted previously. As the CEO of the firm explained:

*We have to find our ways to survive within 5 years. From B2C to B2B, we learn to create diversified revenues and recombine our investment modules in more effective ways ... As a new start-up, we have to be very flexible and dynamic to shape our business models. As our partners increase, we manage to have flexible business models with multiple revenue streams by providing appropriate technologies.*

Additionally, Firm C captures value by forging partnerships with street vendors, small restaurants, hair salons and other establishments, with a specific focus on targeting college students (Generation Z). Recognizing that the younger generation may have limited financial resources to embark on their initial investments, and that delaying investment activities can result in disadvantages in terms of long-term financial planning, Firm C has recently undertaken an innovative initiative called the ‘RA investment with consumption’ project. Collaborating with establishments that appeal to a younger customer base, Firm C has introduced a system wherein students receive 20% reward points from their spending, which are subsequently used as seed capital for investment purposes. This approach alleviates the financial burden on college students, enabling them to participate in investment opportunities, while simultaneously empowering them to take control of their financial planning and potentially reducing inequality among the younger generation (social sustainability). The community engagement with the street vendors, small restaurants, hair salons and other establishments not only facilitates value capture for Firm C, but, more importantly, empowers these establishments to secure income and sustain their operations effectively, thereby reducing inequality in society (social sustainability). Furthermore, these transactions are conducted electronically, facilitating the tracking of transactions and the accumulation of reward points. This electronic mode of transaction not only

minimizes paper waste but also has the potential to reduce other waste. For instance, if restaurants are unable to sell all their food products, they can sell these products via Firm C's platform, thereby mitigating food waste. In this regard, the value captured through partnerships with other establishments not only supports economic sustainability but also contributes to the attainment of social and environmental sustainability.

In summary, Firm C leverages its RA technology and data assets to evaluate investors' risk preferences and effectively manage their investment portfolios. Through partnerships with diverse platforms including the mom-platform, investment platform, and asset management companies, Firm C creates complementary value propositions for the reduction of transaction costs, resulting in lower minimum investment thresholds and enhancing accessibility and affordability for individuals who were previously unable to afford investment services (social sustainability). The service also improves the efficiency and accuracy of the investment process (economic sustainability). Firm C also co-engages its partners and users, and captures value from its offerings with these platforms, such as by providing 'value packages' and customized investment opportunities focused on sustainability-related stocks. The firm also partners with street vendors, small restaurants, and hair salons via electronic services, thereby contributing to environmental sustainability as well as fostering social empowerment and contributing to the reduction of inequality (social sustainability).

## **DISCUSSION AND CONCLUSION**

The aim of this study was to gain a deeper insight into how BMs are designed to effectively harness digital technologies and data, enabling the creation, delivery, and capture of value, while ensuring the achievement of digital sustainability. Fintech solutions have emerged as pivotal drivers in advancing digital sustainability (Arner et al., 2022; Buckley et al., 2021; Zhou et al., 2022). Nevertheless, a significant knowledge gap persists regarding how Fintech firms achieve digital sustainability. The findings of our qualitative study present several contributions to the current understanding of the concept of digital sustainability (George et al., 2021; Merello et al., 2022; Pan et al., 2022; Pan & Zhang, 2020) from a BM lens.

First, building on research on digital sustainability (George et al., 2021), this study emphasizes the significance of examining technology- and data-driven BMs, as they are not only facilitating firms in the creation, delivery, and capture of value, but also serve as a means to achieve digital sustainability. Within the existing body of BM research, the majority of studies have predominantly centered on a purely commercial perspective (Achtenhagen et al., 2013; Casadesus-Masanell & Ricart, 2011; Latifi et al., 2021). While a growing number of studies have started to explore sustainable BMs (Baldassarre et al., 2017; Bocken et al., 2019), the literature integrating sustainability

with digital/technology is still relatively limited (Guandalini, 2022; Höllerer et al., 2023; Merello et al., 2022). Building on the different dimensions of digital sustainability - economic, social, and environmental (Lee, Che-Ha, & Syed Alwi, 2021), we contribute theoretically and empirically to shed light into what these digital sustainability outcomes are in the context of Fintech businesses, including by enhancing the efficiency of financial service provision (economic sustainability), fostering the inclusion of underserved segments (social sustainability), and reducing the environmental impact through reduced natural resource consumption (environmental sustainability).

Second, we expand the perspective of BMs in relation to how firms achieve digital sustainability (Bencsik et al., 2023). Our findings highlight the importance of (i) creating complementary value propositions with users and partners, (ii) delivering value through complementary mechanisms, and (iii) capturing value through complementary revenue streams. While previous studies emphasize the significance of BMs for digital sustainability (Bencsik et al., 2023), our study extends this understanding by highlighting the complementary roles of users and partners in the processes of value creation, delivery, and capture that contribute to the attainment of digital sustainability. Specifically, users and partners leverage digital technologies not only to fulfill their financial requirements, but also to contribute to the betterment of the environment and society. Within the cases examined, users assume diverse roles, encompassing economic roles as lenders, borrowers, drivers, or investors, as well as social and environmental roles. These latter roles involve facilitating access to loans for individuals who were previously marginalized, mitigating air pollution, and directing investments towards endeavors that adhere to ESG principles.

Our findings further illustrate how partners are inspired to deliver value to their users through complementary mechanisms of co-supporting, co-designing, and co-engaging. With Fintech firms increasingly adopting embedded strategies and forming partnerships with multiple platforms to acquire complementary assets, including data assets, this research further highlights the mechanisms through which Fintech firms collaborate with various platforms to deliver value. For instance, in the case of P2P platform, the technology and data-driven financial innovation focuses on credit risk assessments. However, given that a significant portion of financial data is housed within traditional financial institutions, such as banks, government agencies and credit bureaus, firms need to work together with (co-support) these established institutions to access essential data. For the UBI platform, co-designing plays a vital role in value delivery – the UBI platform prioritizes accident rate assessment as a pivotal Fintech innovation, aiming to revamp driving behaviours and mitigate car accidents. The AIoT technologies enabled through Fintech enhances driver safety, enabling them to be more aware of

their driving behaviours and prevailing traffic conditions. This facilitates a redesign of the UBI system in a manner that is not only more risk-conscious but also encourages safety and a sense of responsibility among drivers. For the RA investment platform, it leverages the technologies and data to co-engage with its numerous partners, customized investment packages that are specifically designed to meet the individual requirements of users. To achieve this, the platform relies on the integration of technology and data to assist investors in understanding their unique risk preferences, thereby enabling the creation of customized investment portfolios.

Our research further highlights how FinTech firms capture value by integrating complementary revenue streams through collaborations with various partners. Instead of solely relying on interest-based revenues, the firms adapt by diversifying their income streams through transaction fees and contractual revenues. The integration of complementary revenue streams emphasizes the significance of both transactional and relational relationships with multiple partners to effectively serve their customers. This approach to value capture not only renders their revenue streams more flexible and diversified but also establishes sustainable BMs through promoting economic, social, and environmental sustainability.

Each firm in this study demonstrates a commitment to not only pursue economic sustainability but also contribute to social and environmental sustainability through their BMs. While there are differences in the approaches taken, all firms exhibit similarities in achieving sustainability across these dimensions. From an economic sustainability perspective, Fintech firms utilize digital technologies and data to enhance the efficiency and effectiveness of financial services, through for example significantly reducing the processing time of lending procedures, offering more flexible insurance packages, or providing personalized investment portfolios. Furthermore, these firms leverage their digital technologies to reduce the consumption of natural resources by transitioning to paperless operations and investing in sustainability-related projects, thereby promoting environmental sustainability. The financial solutions provided by Fintech firms also foster financial inclusion by providing access to financial services for the underserved or disadvantaged groups in a more affordable manner. These financial innovations not only improve the availability of financial services but also empower users to develop a heightened awareness and capability in managing their financial assets, thereby promoting social sustainability.

This study had further developed the principle of business model innovation for digital sustainability. The first principle for digital sustainability is “using digital technologies to reduce the entry barrier for users in inclusive financing.” Fintech enterprises differ from incumbent financial institutions mainly in their accessibility and availability for inexperienced users, including college students, novice drivers, and

inexperienced investors. Being inclusive is a key characteristic of fintech startups, aligning them with social sustainability when they incrementally expand their markets to accommodate an increasingly older and more experienced target user base.

The second principle for fintech enterprises is “developing in sync with the life cycles of users to build partnerships”; this principle is crucial because fintech enterprises must learn to build partnerships to fulfill the increasing demand of their target users. As a borrower, a college student may grow up to become an office worker or an engineer and need more money to study abroad, travel, purchase a home, or start a business. As a driver, an individual may require more advanced UBI for security when they start to drive extensively; conversely, they may decide to use renting services on car-renting platforms. As an investor, an individual may require more investment advice for their children, travel plans, and retirement plans. In conclusion, the first principle of using digital technologies to reduce entry barrier for users and the second principle of developing in sync with the life cycles of users both contribute to the creation of complementary value propositions with users and partners through the acquisition of more authentic and sufficient data (first construct of digital sustainability).

The third principle is empowering users and partners to co-support, co-design, or co-create digital services. The key function of digital technology is empowering users and partners to contribute more data. Data with varying velocity, volume, and variety are collected from borrowers; such data include their lifestyle information, education backgrounds, academic scores, shopping behaviors, and social community connections; for drivers and investors, data pertaining to their driving behaviors and risk preferences, respectively, are collected. When users become more proactive and involved in the data collection process, they can create more value for digital services. Furthermore, partners are empowered by digital technologies because such technologies enable them to co-support more services for borrowers, co-design driving security services for drivers who drive extensively, and co-create consulting services for investors.

The fourth principle is extending one’s revenue stream by expanding from financial services to nonfinancial services. When fintech startups learn to cooperate with more partners, they learn to develop more diversified revenue streams by introducing nonfinancial services. In our case studies, subscription fees were used to generate contract-based revenue, and transaction fees for financial and other services were used to enrich revenue stream portfolios. Through these revenue streams, fintech companies can achieve economic sustainability. Furthermore, they reduce the entry barriers for users, contributing to the attainment of social sustainability by promoting inclusivity from the inception of business. Finally, when fintech companies provide financial services through digital technologies and reduce the consumption of natural resources (e.g., paper), they achieve environmental sustainability.

### **Managerial implications**

Our findings also offer managerial implications. In order to achieve digital sustainability, it is crucial for managers to gain a deeper understanding of the processes involved in creating, delivering, and capturing value propositions, not only from a commercial perspective but also considering how to achieve digital sustainability. Our analysis offers an overview of different ways through which firms can utilize and leverage their technologies and data to create complementary value propositions with their users and partners. We also emphasize the significance of partnerships in implementing different mechanisms for value delivery – from co-supporting, co-designing and co-engaging. In the context of Fintech firms, where financial data is primarily held by institutional banks, asset management firms or other institutional bodies, these Fintech firms leverage their technologies and data sharing to capture pertinent information, avoiding the need to reinvent the wheel. While the Fintech firms strive to enhance accessibility and affordability of their services, they work together with their partners to promote social and environmental sustainability. For instance, they adopt measures to minimize the consumption of natural resources while simultaneously increasing investments in projects aligned with ESG principles. In this regard, their BMs are focused on fostering relationships with users and partners that extend beyond the boundary of firm to not only ensure economic sustainability but also making significant contributions to social and environmental sustainability.

Second, managers learn of the significance of multiple revenue streams. The integration of multiple revenue streams enables businesses to leverage technologies and data to expand their reach and access to new markets. Our findings indicate that the multiple revenue streams also enable firms to enhance the efficiency and effectiveness of their financial service offerings, effectively reducing the consumption of natural resources through automated processes, and catering to previously underserved customer segments, including those who were previously excluded or unable to afford such financial services. Adopting a BM perspective on digital sustainability, our research provides insights for managers of their roles in creating, delivering, and capturing value to promote digital sustainability.

Third, the study provides implications for policymakers in relation to the inherent complexity of BMs and digital sustainability. In Taiwan, Fintech firms face regulatory limitations, and they can only engage in activities that align with established regulations. These firms are encouraged to experiment with their BMs within a sandbox environment, although this process often entails significant time, typically spanning 2-3 years before implementation. Notably, Firm B in our study participated in the sandbox program, which facilitated the redesign of its UBI offering. Given the rapid



advancement of digital technologies, time becomes a critical factor, and it requires policy makers and regulatory frameworks to keep pace with evolving changes. This underscores the importance of flexibility and timely adaptation in regulatory practices, enabling Fintech companies, and entrepreneurs more broadly, to thrive and scale their businesses. In addition, we highlighted the importance of digital sustainability in our study, and we urge policymakers to strengthen ESG information disclosure to encourage firms to support and promote digital sustainability.

### **Limitations and directions for future studies**

While this study offers significant insights into how firms create, deliver, and capture value, while attaining digital sustainability, it is important to acknowledge its limitations, which can serve as a guide for future research endeavors. First, the cases examined in this study focus on FinTech firms in their early phase to intermediate stages of business. The AI technologies are still in the emergent phase, and the potential implications and future prospects of leveraging AI advancements by these businesses remain unknown. Therefore, future research should consider conducting longitudinal studies to track the growth trajectories of these firms and gain an understanding of how they effectively utilize their technologies and data to explore how far they achieve digital sustainability.

Second, our study only focuses on the BM with respect to value creation, delivery, and capture, specifically pertaining to the achievement of digital sustainability. However, our research does not delve into the financial performance of firms or the strategic management of partnerships. Further research is needed to develop our understanding of the specific processes by which a BM archetype can either facilitate or impede digital sustainability endeavors. Scholars could also explore the underlying processes necessary to ensure the scalability of their BMs, for example through BM diversification, to achieve digital sustainability.

Third, our empirical study was limited to three cases in Taiwan, and as such, we do not claim the generalizability of our findings. The rich data of this study served as a starting point in exploring how firms design their BMs to promote digital sustainability, as well as the resulting economic, social, and environmental outcomes stemming from such endeavors. Future research can further examine BMs adopted by different types of digital businesses across diverse contexts.

Further, the problems associated with using data as a source of revenue must be considered. First, the challenge of data management must be addressed. FinTech managers, banks, and policymakers can ethically collect consumer data that can help them to improve their decision-making and product development processes. Data involving the use of mobile applications and transaction histories can be useful for improving a bank's products and services. Strict penalties should be imposed on fintech

companies for selling customer data, sending unsolicited messages, or gaining unnecessary access to customer contact lists (Sampat et al., 2023). Second, the problem of data bias or data discrimination should be addressed. Because fintech companies use a considerable amount of alternative data (e.g., data related to a customer's job, education, and social network) to evaluate customer creditworthiness, they may have to put in more effort to assess the authenticity and appropriateness of data for credit assessment (Di Maggio et al., 2022). Alternative data help fintech companies assess the risk that a customer will default on their loan repayments. Inaccurate data can result in considerable losses to the company. All these challenges of data protection and data discrimination need to be discussed in the future.

In conclusion, this research illustrates BMs for FinTech firms, enabling them to effectively leverage their technologies and data to create complementary value propositions with their users and partners, deliver their value through co-supporting, co-designing, and co-engaging mechanisms, and capture value from the integration of complementary revenue streams. We further elaborate how these BMs contribute to digital sustainability, including by enhancing the efficiency of financial service provision (economic sustainability), fostering the inclusion of underserved segments (social sustainability), and reducing the environmental impact through reduced natural resource consumption (environmental sustainability). Our study extends the relatively limited research on digital sustainability, with the hope of further investigations in this domain to better understand the design of BMs for digital sustainability.

## REFERENCES

- Achtenhagen, L., Melin, L., & Naldi, L. 2013. Dynamics of business models – Strategizing, critical capabilities and activities for sustained value creation. *Long Range Planning*, 46: 427-442.
- Aguilera, R. V., Aragón-Correa, J. A., Marano, V., & Tashman, P. A. 2021. The corporate governance of environmental sustainability: A review and proposal for more integrated research. *Journal of Management*, 47: 1468-1497.
- Amit, R., & Zott, C. 2015. Crafting business architecture: The antecedents of business model design. *Strategic Entrepreneurship Journal*, 9: 331-350.
- Arner, D. W., Buckley, R. P., Zetsche, D. A., & Veidt, R. 2020. Sustainability, FinTech and financial inclusion. *European Business Organization Law Review*, 21: 7-35.
- Aversa, P., Haefliger, S., Hueller, F., & Reza, D. S. 2021. Customer complementarity in the digital space: Exploring Amazon's business model diversification. *Long Range Planning*, 54: 101985.

- Baldassarre, B., Calabretta, G., Bocken, N., & Jaskiewicz, T. 2017. Bridging sustainable business model innovation and user-driven innovation: A process for sustainable value proposition design. *Journal of Cleaner Production*, 147: 175-186.
- Bencsik, B., Palmié, M., Parida, V., Wincent, J., & Gassmann, O. 2023. Business models for digital sustainability: Framework, microfoundations of value capture, and empirical evidence from 130 smart city services. *Journal of Business Research*, 160, <https://doi.org/10.1016/j.jbusres.2023.113757>
- Bocken, N., Boons, F., & Baldassarre, B. 2019. Sustainable business model experimentation by understanding ecologies of business models. *Journal of Cleaner Production*, 208: 1498-1512.
- Bohnsack, R., Bidmon, C. M., & Pinkse, J. 2022. Sustainability in the digital age: Intended and unintended consequences of digital technologies for sustainable development. *Business Strategy and the Environment*, 31: 599-683.
- Brenner, B., & Hartl, B. 2021. The perceived relationship between digitalization and ecological, economic, and social sustainability. *Journal of Cleaner Production*, 315: 128128.
- Broccardo, L., Truant, E., & Dana, L-P. 2023. The interlink between digitalization, sustainability, and performance: An Italian context. *Journal of Business Research*, 158: 113621.
- Broekhuizen, T. L. J., Emrich, O., Gijsenberg, M. J., Broekhuis, M., Donkers, B., & Sloot, L. M. 2020. Digital platform openness: Drivers, dimensions and outcomes. *Journal of Business Research*, 122: 902-914.
- Buckley, R. P., Zetsche, D. A., Arner, D. W., & Veidt, R. 2021. FinTech, financial inclusion and the UN Sustainable Development Goals. In I. Chiu and G. Deipenbrock (Eds.), *Routledge Handbook of Financial Technology and Law*, Routledge, UK.
- Caputo, A., Pizzi, S., Pellegrini, M. M., & Dabić, M. 2021. Digitalization and business models: Where are we going? A science map of the field. *Journal of Business Research*, 123: 489-501.
- Casadesus-Masanell, R., & Ricart, J. E. 2011. How to design a winning business model. *Harvard Business Review*, 89: 101-107.
- Castaldi, S., Wilhelm, M. M., Beugelsdijk, S., & van der Vaart, T. 2023. Extending social sustainability to suppliers: The role of GVC governance strategies and supplier country institutions. *Journal of Business Ethics*, 183: 123-146.
- Chatterjee, S. 2013. Simple rules for designing business models. *California Management Review*, 55: 97-124.

- Codini, A. P., Abbate, T., & Messeni Petruzzelli, A. 2023. Business model innovation and exaptation: A new way of innovating in SMEs. *Technovation*, 119: 102548.
- Cozzolino, A., Verona, G., & Rothaermel, F. T. T. 2018. Unpacking the disruption process: New technology, business models, and incumbent adaptation. *Journal of Management Studies*, 57: 1166-1202.
- Corley, K. G., & Gioia, D. A. 2004. Identity ambiguity and change in the wake of a corporate spinoff. *Administrative Science Quarterly*, 49(2): 173-208.
- Del Rio Castro, G., Gonzales Fernandez, M. C., & Uruburu Colso, A. 2021. Unleashing the convergence amid digitalization and sustainability towards pursuing the Sustainable Development Goals (SDGs): A holistic review. *Journal of Cleaner Production*, 280: 122204.
- Demil, B., & Lecocq, X. 2010. Business model evolution: In search of dynamic consistency. *Long Range Planning*, 43: 227–246.
- Demir, A., Pesqu e-Cela, V., Altunbas, Y., & Murinde, V. 2022. Fintech, financial inclusion and income inequality: A quantile regression approach. *The European Journal of Finance*, 28: 86-107.
- Di Maggio, M., Ratnadiwakara, D., & Carmichael, D. (2022). *Invisible primes: Fintech lending with alternative data* (No. w29840). National Bureau of Economic Research.
- Di Vaio, A., Palladino, R., Hassan, R., & Escobar, O. 2020. Artificial intelligence and business models in the sustainable development goals perspective: A systematic literature review. *Journal of Business Research*, 121: 283-314.
- Elkington, J. 1998. *Cannibals with forks: The triple bottom line of 21<sup>st</sup> century business* (2<sup>nd</sup> edition). Oxford: Capstone Publishing Ltd.
- Financial Supervisory Commission (2015a) Construct digital finance environments. <http://www.fsc.gov.tw/ch/home>.
- Financial Supervisory Committee (2015b) The act governing electronic payment institutions <http://law.fsc.gov.tw/law/>
- Financial Supervisory Committee (2015c) Regulations of bank and financial holding company investing FinTech firms <http://law.banking.gov.tw/Eng/FLAW/FLAWQRY03.aspx?lsid=FL078705&keyword=10460003280>.
- Financial Supervisory Committee (2016) Financial technology area <http://www.fsc.gov.tw/ch/home.jsp?id=478&parentpath=0,7>. Accessed 30 Sept 2016
- Foss, N. J., & Saebi, T. 2017. Fifteen years of research on business model innovation: How far have we come, and where should we go? *Journal of Management*, 43: 200-227.

- Franco, S., Presenza, A., Messeni Petruzzelli, A., & Peruffo, E. 2022. Tradition-driven business models at luxury companies: Revealing value-creation and value-capture activities. *Journal of Knowledge Management*, 26: 486-506.
- Frankenberger, K., & Sauer, R. 2019. Cognitive antecedents of business models: Exploring the link between attention and business model design over time. *Long Range Planning*, 52: 283-304.
- Freudenreich, B., Lüdeke-Freund, F., & Schaltegger, S. 2020. A stakeholder theory perspective on business models: Value creation for sustainability. *Journal of Business Ethics*, 166: 3-18.
- Fusch, P. I., & Ness, L. R. 2015. Are we there yet? Data saturation in qualitative research, *Qualitative Report*, 20: 1408-1416.
- Geissinger, A., Laurell, C., Öberg, C., Sandström, C., & Suseno, Y. 2022. The sharing economy and the transformation of work: Evidence from Foodora. *Personnel Review*, 51: 584-602.
- George, G., Merrill, R. K., & Schillebeeckx, S. J. D. 2021. Digital sustainability and entrepreneurship: How digital innovations are helping tackle climate change and sustainable development. *Entrepreneurship Theory and Practice*, 45: 999-1027.
- George, G., & Schillebeeckx, S. J. D. 2021. Digital sustainability and its implications for finance and climate change. *Macroeconomic Review*, 10: 103-108.
- Gioia, D., Corley, K., & Hamilton, A. 2013. Seeking qualitative rigor in inductive research: Notes on the Gioia Methodology. *Organizational Research Methods*, 16(1): 15-31.
- Gomber, P., Kauffman, R. J., Parker, C., & Weber, B. W. 2018. On the FinTech revolution: Interpreting the forces of innovation, disruption and transformation in financial services. *Journal of Management Information Systems*, 35: 220-265.
- Guandalini, I. 2022. Sustainability through digital transformation: A systematic literature review for research guidance. *Journal of Business Research*, 148: 456-471.
- Höllerer, M. A., Shinkle, G., George, G., Mair, J., Pan, S. L., & Tim, Y. 2023. Call for Papers: Digital sustainability: Addressing managerial, organizational, and institutional challenges. *Academy of Management Perspectives*, <https://aom.org/events/event-detail/2023/05/15/higher-logic-calendar/amp-call-for-special-issue-papers-digital-sustainability>
- Hornuf, L., & Haddad, C. 2019. The emergence of the global fintech market: Economic and technological determinants. *Small Business Economics*, 53(1): 81-105.

- Hutchins, M. J., Richter, J. S., Henry, M. L., & Sutherland, J. W. 2019. Development of indicators for the social dimension of sustainability in a U.S. business context. *Journal of Cleaner Production*, 212: 687-697.
- KPMG. 2019. Forging the future: How financial institutions are embracing fintech to evolve and grow. <https://assets.kpmg.com/content/dam/kpmg/ke/pdf/thought-leaderships/Forging-with%20bleeds.pdf>
- Lähtenmäki, I., Nätti, S., & Saraniemi, S. 2022. Digitalization-enabled evolution of customer value creation: An executive view in financial services. *Journal of Business Research*, 146: 504-517.
- Langley, D. J., van Doorn, J., Ng, I. C. L., Stieglitz, S., Lazovik, A., & Boonstra, A. 2021. The Internet of Everything: Smart things and their impact on business models. *Journal of Business Research*, 122: 853-863.
- Latifi, M-A., Nikou, S., & Bouwman, H. 2021. Business model innovation and firm performance: Exploring causal mechanisms in SMEs. *Technovation*, 107: 102274.
- Lee, C. M. J., Che-Ha, N., Syed Alwi, S. F. 2021. Service customer orientation and social sustainability: The case of small medium enterprises. *Journal of Business Research*, 122: 751-760.
- Lee, K., & Jung, H. 2019. Dynamic semantic network analysis for identifying the concept and scope of social sustainability. *Journal of Cleaner Production*, 233: 1510-1524.
- Lee, I., & Shin, Y. J. 2018. Fintech: Ecosystem, business models, investment decisions, and challenges. *Business Horizons*, 61: 35-46.
- Lichtman, M. 2013. *Qualitative research for social sciences*. London: Sage Publications.
- Liute, A., & De Giacomo, M. R. 2022. The environmental performance of UK-based B Corp companies: An analysis based on the triple bottom line approach. *Business Strategy and the Environment*, 31: 810-827.
- Locke, K. 1996. Rewriting the discovery of grounded theory after 25 years? *Journal of Management Inquiry*, 5(3): 239-245.
- Manninen, K., Koskela, S., Antikainen, R., Bocken, N., Dahlbo, H., & Aminoff, A. 2018. Do circular economy business models capture intended environmental value propositions? *Journal of Cleaner Production*, 171: 413-422.
- Massa, L., Tucci, C., & Afuah, A. 2017. A critical assessment of business model research. *Academy of Management Annals*, 11: 73– 104.
- Mignon, I., & Bankel, A. 2022. Sustainable business models and innovation strategies to realize them: A review of 87 empirical cases. *Business Strategy and the Environment*, <https://doi.org/10.1002/bse.3192>

- Miles, M. B., Huberman, A. M., & Saldaña, J. 2014. *Qualitative data Analysis: A methods sourcebook* (Third Edition). Thousand Oaks, CA: Sage Publications.
- Missimer, M., Robért, K-H., & Broman, G. 2017. A strategic approach to social sustainability – Part 2: A principle-based definition. *Journal of Cleaner Production*, 140: 42-52.
- Merello, P., Barberá, A., & De la Poza, E. 2022. Is the sustainability profile of FinTech companies a key driver of their value? *Technological Forecasting and Social Change*, 174: 121290.
- Novet, J. 2022. Google selects Coinbase to take cloud payments with cryptocurrencies and will use its custody tool. *CNBC*, October 11, <https://www.cnbc.com/2022/10/11/google-selects-coinbase-to-take-cloud-payments-with-cryptocurrencies.html>
- Oderanti, F. O., Li, F., Cubric, M., & Shi, X. 2021. Business models for sustainable commercialisation of digital healthcare (eHealth) innovations for an increasingly ageing population. *Technological Forecasting & Social Change*, 171: 120969.
- OECD. 2021. Artificial Intelligence, machine learning and big data in finance: Opportunities, challenges, and implications for policy makers, <https://www.oecd.org/finance/artificial-intelligence-machine-learningbig-data-in-finance.htm>.
- Osiyevskyy, O., & Dewald, J. 2015. Explorative versus exploitative business model change: The cognitive antecedents of firm-level responses to disruptive innovation. *Strategic Entrepreneurship Journal*, 9: 58-78.
- Osterwalder, A., & Pigneur, Y. 2010. *Business model generation*. Hoboken, NJ: John Wiley & Sons.
- Pan, S.L., L. Carter, Y. Tim, & S.M. Seshadrinath (2022). Digital sustainability: Technology solutions to climate change. *Business Think*, March 2022, <https://www.businessthink.unsw.edu.au/articles/digital-sustainability-solutions-climate-change>
- Pan, S. L., & Zhang, S. 2020. From fighting COVID-19 pandemic to tackling sustainable development goals: An opportunity for responsible information systems research. *International Journal of Information Management*, 55: 102196.
- Parmentola, A., Petrillo, A., Tutore, I., & De Felice, F. 2022. Is blockchain able to enhance environmental sustainability? A systematic review and research agenda from the perspective of Sustainable Development Goals (SDGs). *Business Strategy and the Environment*, 31: 194-217.

- Priem, R. L., Wenzel, M., & Koch, J. 2018. Demand-side strategy and business models: Putting value creation for consumers center stage. *Long Range Planning*, 51: 22-31.
- Rietveld, J. 2018. Creating and capturing value from freemium business models: A demand-side perspective. *Strategic Entrepreneurship Journal*, 12: 171-193.
- Ritter, T., & Lettl, C. 2018. The wider implications of business-model research. *Long Range Planning*, 51: 1–8.
- Saebi, T., Lien, L., & Foss, N. J. 2017. What drives business model adaptation? The impact of opportunities, threats and strategic orientation. *Long Range Planning*, 50: 567-581.
- Sampat, B., Mogaji, E., & Nguyen, N. P. (2023). The dark side of FinTech in financial services: a qualitative enquiry into FinTech developers' perspective. *International Journal of Bank Marketing*, .
- Schiavone, F., Mancini, D., Leone, D., & Lavorato, D. 2021. Digital business models and ridesharing for value co-creation in healthcare: A multi-stakeholder ecosystem analysis. *Technological Forecasting & Social Change*, 166: 120647.
- Schrauder, S., Kock, A., Baccarella, C. V., & Voigt, K-I. 2018. Takin' care of business models: The impact of business model evaluation on front-end success. *The Journal of Product Innovation Management*, 35: 410-426.
- Schwanholz, J., & Leipold, S. 2020. Sharing for a circular economy? An analysis of digital sharing platforms' principles and business models. *Journal of Cleaner Production*, 269: 122327.
- Siggelkow, N. 2007. Persuasion with case studies. *Academy of Management Journal*, 50: 20-24.
- Sjödin, D., Parida, V., Jovanovic, M., & Visnjic, I. 2020. Value creation and value capture alignment in business model innovation: A process view on outcome-based business models. *The Journal of Product Innovation Management*, 37: 158-183.
- Snihur, Y., & Bocken, N. 2022. A call for action: The impact of business model innovation on business ecosystems, society and planet. *Long Range Planning*, 55: 102182.
- Song, H., Han, S., Liu, W., & Ganguly, A. 2022. What role do FinTech companies play in supply chain finance? A signaling intermediary perspective. *Journal of Business & Industrial Marketing*, 38: 1279-1294.
- Strauss, A. L. 1987. *Qualitative analysis for social scientists*. Cambridge University Press.
- Subramaniam, M. 2022. *The future of competitive strategy: Unleashing the power of data and digital ecosystems*. Cambridge, MA: The MIT Press.



- Teece, D. 2010. Business models, business strategy and innovation. *Long Range Planning*, 43: 172-194.
- Tidhar, R., & Eisenhardt, K. M. 2020. Get rich or die trying ... finding revenue model fit using machine learning and multiple cases. *Strategic Management Journal*, 41: 1245-1273.
- Trabucchi, D., Talenti, L., & Buganza, T. 2019. How do Big Bang disruptors look like? A business model perspective. *Technological Forecasting and Social Change*, 141: 330-340.
- United Nations Environment Programme. 2016. Fintech and the Sustainable Development: Assessing the implications – Summary.  
<https://www.unep.org/resources/report/fintech-and-sustainable-development-assessing-implications-summary>
- Walls J. L., Phan P. H., Berrone P. 2011. Measuring environmental strategy: Construct development, reliability, and validity. *Business & Society*, 50: 71-115.
- Yin, R. K. 2011. *Applications of case study research*, 3<sup>rd</sup> edition. Sage Publications.
- Zhou, G., Zhu, J., & Luo, S. 2022. The impact of fintech innovation on green growth in China: Mediating effect of green finance. *Ecological Economics*, 193: 107308.
- Zott, C., & Amit, R. 2010. Business model design: An activity system perspective. *Long Range Planning*, 43: 216-226.
- Zott, C., Amit, R., & Massa, L. 2011. The business model: Recent developments and future research. *Journal of Management*, 37: 1019-1042.

**Table 1. Fieldwork investigation summary**

Research Methods	Informants	Responsibilities	No.	Frequency	Hours
Personal interviews	Case1: Firm A	The CEO and senior managers who are responsible for the cooperation with multiple partners	3	10	30
		Partners of Firm A including senior managers of banks	5	2	10
		Users: Interviews and secondary data collection on Dcards	20	--	15
	Case 2: Firm B	The CEO and project managers who are responsible for initiating	5	3	15

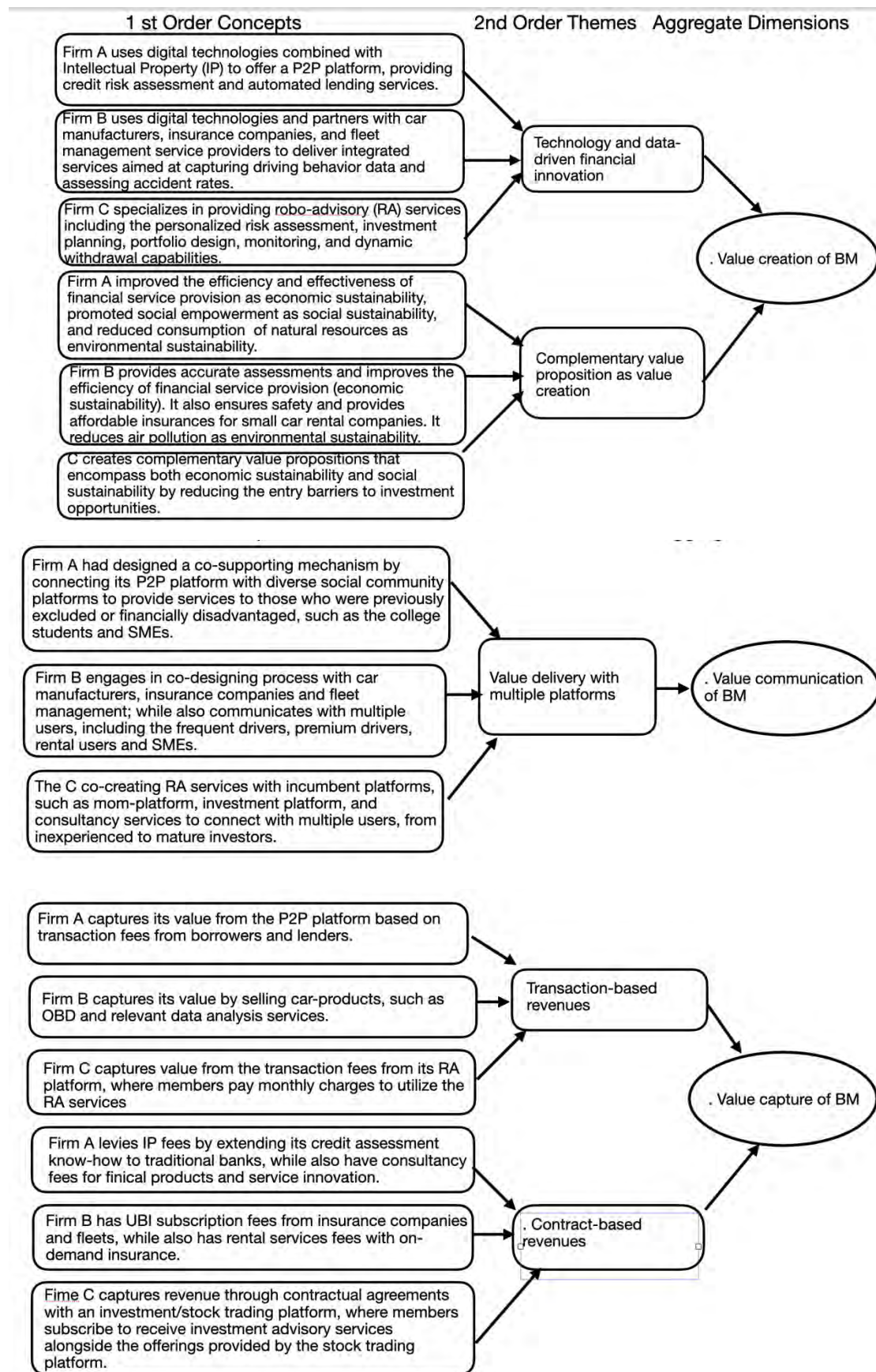
		new partnerships with cars, fleets, insurance companies			
		Partners of Firm B, including senior managers of insurance companies, cars, and fleet management service providers	6	2	15
		Users: Interviews and secondary data collection on e-commerce and other online forums	22	--	15
	Case 3: Firm C	The CEO and project managers who are responsible for initiating new partnerships with mom-platform, investment platform, and asset management companies	12	12	24
		Partners of Firm C, including senior managers of mom-platform, investment platform, and asset management companies	6	3	18
		Users: Interviews and secondary data collection on mom-platform, investment platform	25	--	15
Internal meetings	Researchers take summer interns to participate in the internal meetings to seek feedback from members and executives		3x10		30
Secondary data	The public media reports, and seminars to triangulate the data with multiple sources				10
Total Hours of Field Engagements					205

**Table 2. Business models for digital sustainability**

Case studies	Firm A	Firm B	Firm C
Technology- and data-driven financial innovation	Data for credit risk assessment through P2P platform	Driving behavior data, accident rate assessment and UBI	Investment risk preference assessment, manage investment portfolios
Complementary value propositions as value creation	P2P services, connecting lenders and borrowers, and providing services to those who were previously excluded or financially disadvantaged	Embedded services, incorporating AIoT technologies in cars, and working together with car manufacturers, insurance companies, fleet management service providers	Embedded services (mom-platform, investment platform, consultancy services)
Value delivery mechanisms	Co-supporting the P2P platform and social community platforms (such as Dcard, PTT) Multiple users and partners with P2P platform service: from junior to senior lenders, and college students, engineers, SMEs as borrowers	Co-designing with embedded platforms (car manufacturers, insurance companies, fleet management) Multiple users and partners: from those who occasionally drive to frequent drivers and premium drivers, fleet management, rental users, SMEs	Co-creating with embedded platform (mom-platform, investment platform, consultancy services ) Multiple users and partners: from inexperienced to mature investors, and mothers, senior stock investors, HNWI
Complementary revenue streams to capture value	Transaction-based revenue and interest fees Contract-based revenue: IP license fees, installment fees	Transaction-based revenue: product/service sales Contract-based revenue: UBI subscription fees, rental service fees ('as-you-go' or on-demand insurance)	Transaction-based revenue: transaction fees Contract-based revenue: subscription fees, investment consultancy fees

Digital sustainability	<p><b>Economic sustainability:</b></p> <ul style="list-style-type: none"> <li>- Improve the efficiency and effectiveness of financial service provision (e.g., reduction of bad debts)</li> <li>- Lower transaction costs</li> <li>- Provide accurate assessments (in credit risks)</li> </ul> <p><b>Social sustainability:</b></p> <ul style="list-style-type: none"> <li>- Inclusive financing (accessibility and affordable loan services)</li> <li>- Promote social empowerment (reduce financial struggles and poverty)</li> </ul> <p><b>Environmental sustainability:</b></p> <ul style="list-style-type: none"> <li>- Reduce consumption of natural resources (less paper waste)</li> </ul>	<p><b>Economic sustainability:</b></p> <ul style="list-style-type: none"> <li>- Reduce information asymmetry</li> <li>- Provide accurate assessments (in calculating insurance premiums)</li> <li>- Assess information accuracy (when assessing insurance claims)</li> <li>- Improve the efficiency of financial service provision</li> </ul> <p><b>Social sustainability:</b></p> <ul style="list-style-type: none"> <li>- Ensure safety through digital financial services</li> <li>- Promote a sense of responsibility</li> <li>- Support health and wellbeing (safe car journeys – peace of mind)</li> <li>- Inclusive financing (accessibility and affordable insurance package for small and micro car rental companies and those not able/willing to own a car)</li> </ul> <p><b>Environmental sustainability:</b></p> <ul style="list-style-type: none"> <li>- Sustainable transport (less traffic congestion)</li> <li>- Reduce air pollution</li> </ul>	<p><b>Economic sustainability:</b></p> <ul style="list-style-type: none"> <li>- Lower transaction costs and minimum investment thresholds</li> <li>- Improve the efficiency of the process</li> <li>- Provide accurate and unbiased recommendations (when identifying investment opportunities)</li> <li>- Enhance customization opportunities</li> </ul> <p><b>Social sustainability:</b></p> <ul style="list-style-type: none"> <li>- Inclusive financing (accessibility and affordable investment services)</li> <li>- Promote social empowerment (long-term financial planning to reduce income disparity)</li> <li>- Reduce inequality (knowledge gaps about investments, community engagement to increase income for vendors)</li> </ul> <p><b>Environmental sustainability:</b></p> <ul style="list-style-type: none"> <li>- Increase in investments in ESG-related stocks or EFT</li> <li>- Reduce consumption of natural resources (less paper waste, less food waste)</li> </ul>
------------------------	--	---	--

**Figure 1. Data analysis process for digital sustainability of BMI**



# Digital transformation in the public sector: Critical factors of Digital Justice Portal

## Abstract

*The use of ICT in e-government and particularly in justice sector offer new opportunities and more efficient services for citizens and users. Although this scientific area has attracted the interest of many academics and professionals in the justice field and despite the money spent on improving the performance of justice staff and the results of the courts, studies examining the acceptance and satisfaction of electronic justice system users are restricted. Therefore, the study integrates and assesses the IS success model and TAM in the context of e-justice services in Greece. Data was collected by 246 internal users and lawyers in Greece. As the use of the system by lawyers is different from that of employees, it seems that this group is more interested in an overall good picture of the system, ensuring the quality dimensions of the application, which will make it useful for its daily work activities.*

**Keywords:** E-justice system; Success; Satisfaction; IT strategy; Digital transformation

## 1.0 Introduction

Recent advances in Information Technology (IT) and Information Systems (IS) have influenced several industries and the public sector. IT and IS transform individuals, businesses and all public agencies providing with quick and secure access to all resources from a single point on. Implementing IS in government is part of a larger transformation cycle aimed at supporting government to provide safer, more reliable and more productive services to people, organizations and businesses (Angelopoulos et al., 2010; Charalabidis et al., 2019; Kitsios et al., 2009; Loukis and Charalabidis, 2011; Loukis and Tsouma, 2002; Oktal et al., 2016). Successful e-justice system, achieving a level of performance that primarily satisfies most internal users is important. The justice system has also, like different segments of the public sector, explored different avenues regarding the development of ICTs to increase the efficiency and effectiveness of services. In the justice system (known as "e-justice") various programming applications are being introduced to improve interaction and contact between different actors. In the e-justice, the collection of hardware, software, and networking tools built to help administrative court personnel and judges in soothing out their day-to-day work activities, such as case filing, scheduling, workflow, etc., is known as "e-court." The proliferation of trial technology has facilitated the introduction of new work practices and organizational procedures, thus lessening consistent expansion in the outstanding task at hand of courts and the

constant graduates of legal procedures (Agrifoglio et al., 2016; Oktal et al., 2016; Sachan et al., 2018; Stefanovic et al., 2016; Wang and Liao, 2008).

IT solutions have been built with the context of e-justice to enable lawyers and court staff to implement their day-to-day work activities. Such court-management technologies boost administrative staff's job performance as they have been promoting the implementation of innovative work practices. In information systems in courts the information is available to all users in the central electronic environment. Registered users are judges, solicitors, lawyers and other court staff. Electronic justice systems make access to this information fast and convenient. In addition, information systems in courts require users to execute all forms of information and records online. The most important aim of a justice information system is to build more accurate and faster information, documentation, and workflow of incompliant services in courts (Agrifoglio et al., 2016; Oktal et al., 2016).

In Greece, the goal of the November 2008-established Council for European Electronic Justice 2009-2013 Action Plan is to enhance both the availability of justice and the quality of justice provided. It recognizes that communication and information technology can play a significant role in improving the efficiency of justice systems by facilitating the day-to-day work of legal professionals and fostering collaboration amongst legal authorities. It also sets as its goal the establishment of a portal for European electronic justice. Through justice networks, attorneys, notaries, and judges can connect with one another, access legal databases, and obtain information about justice education. They also discover details about setting up international video chats. Previous researchers have used current IT/IS related models to help businesses adopt effective IS. Some of these models include the Technology Acceptance Model (TAM) (Davis, 1989), the Theory of Planned Behavior (TPB) (Ajzen, 1991), and the unified theory of acceptance and use of technology (Angelopoulos et al., 2010). The factors affecting IS and user behavior are significant for IT / IS implementation to be effective. Essentially, evaluation models were introduced to consider user needs and to analyze the dimensions and factors in system growth to increase their acceptance and satisfaction (Kamariotou and Kitsios, 2019; 2017; Kitsios and Kamariotou, 2017). The DeLone and McLean model of IS success (1992) (DeLone and McLean, 2003) is among the highest used to explain the effects of IS, and has been utilized as the reason for many researches in different countries (Wang and Liao, 2008). Although IS researchers have paid a lot of attention at the IS success model, limited papers have

been applied in order to assess the effectiveness of justice systems in courts. Therefore, recognizing the efficacy of e-justice systems, as well as the factors affecting the performance of court employees, clarifies a significant field of inquiry to bridge the gaps in literature and tackle future study.

Current research in the field of e-government has looked at public satisfaction as the end users. There are minimal research about the acceptance and satisfaction of internal users. Furthermore, current IS success models pay attention at system-centric assessment or organizational structure. Scholars have not yet addressed user-centric evaluations of IS in courts. To develop a successful e-justice system, it is necessary to accomplish a level of performance that mainly satisfies most internal users (Agrifoglio et al., 2016; Oktal et al., 2016; Stefanovic et al., 2016).

Several researchers and practitioners in court management have started to study the correlation between technology and individual and organizational success in the justice sector (e.g., Agrifoglio et al., 2016; Hamin et al., 2012; Velicogna, 2007). Researchers studied the trend implementing an exploratory approach, often described by a detailed study of the local justice system, utilized court systems, and performance of the court. While these trends have a significant contribution to the literature of court management, there are still no empirical papers examining the relation between technology and performance in court and, especially, the impact of justice systems on court staff's actions in terms of the use of IS.

Consequently, this paper aims to explore the factors influencing the acceptance and satisfaction of e-justice system users in Greece. Data was collected by 246 internal users and lawyers in Greece. Factor Analysis on detailed items of user acceptance and satisfaction constructs has been applied.

The layout of this paper is as follows sections: The next section, after a brief introduction to this area, is the theoretical background in respect of the satisfaction in e-government and justice. Section 3 explains the methodology, while Section 4 shows survey findings. Finally, conclusions are presented in Section 5 and the paper ends.

## **2.0 Theoretical background**

### **2.1 Information Systems acceptance and satisfaction**

The introduction of IT and computer technology into public administration brought new administrative practices and led to what is now called e-government. E-

government strengthens transparency, efficiency and public accessibility and is increasingly acknowledged as a central pillar to facilitating the transformation of public governance (Sachan et al., 2018). IT, moreover, has transformed government; it provides new opportunities for delivering better, more reliable and competitive services to people and businesses and its acceptance by employees and citizens is a top priority for governors. Therefore, the creation of a conceptual model for the acceptance of digital technology in the public sector, such as that proposed by Sang et al (2009), is particularly useful for developing future political and strategic decisions to enhance the usage of such services.

Much of the literature focuses on users' satisfaction with the development of services in e-government, as the success of such initiatives depends largely on the percentage of their use (Sachan et al., 2018). User acceptance is expressed mainly through the TAM. It is applied to understand individuals' attitudes towards the use of technology, which can lead to further acceptance and adoption. That is to say, the attitude formed by TAM represents the attitude formed towards the use of technology. It is considered as one of the earliest and most widely accepted research approaches; it is a dominant model in the field of technology and in the use of IS, along with the theory of IS success suggested by DeLone and McLean. According to the TAM model, the significant factors that influence the adoption and usage of digital technologies are perceived ease of use and perceived usefulness, with Davis (1989) being its main exponent. According to Davis, the model can be used to investigate the frequency at which users use a specific technology, the characteristics of the system, and the reasons users ultimately accept or reject it. In conducting a research on users of two information systems in a Canadian company and evaluating the variables used in the initial research, Davis said that both perceived usefulness and ease of use are strongly associated with self-reported system indicators; and, therefore, the final degree of acceptance and frequency of use of a system by its end users depends directly on what motivates each user.

Weerakkody et al. (2016) attempt to fill a research gap by exploring the significance of users' trust in the efficiency of a system and its information in the UK, and to what extent cost affects satisfaction. The five dimensions highlighted in their paper have significantly affect users' satisfaction with services in the public sector. According to Anwer et al. (2016), a thorough evaluation of these services will help highlight their strengths and weaknesses, identify their new guidelines and compare their



organization locally, nationally and internationally. For this reason, they are proceeding with an analysis and assessment of the current state of Afghanistan's e-government services, through a combination of evaluation approaches. Sachan et al. (2018) investigate users' satisfaction of e-government services and therefore suggest a model, incorporating the TAM into the process. This research can help app developers gain an idea of the needs of users in order to enhance the design and implementation of these systems. According to Wirtz et al. (2016), the key difficulty for local e-government portals is to define the most important dimensions affecting user satisfaction. For this reason, they develop a model to satisfy the users of such gates, using mixed methods. Also, the research of Danila et al. (2014) explores user intentions and the use of e-government services; it presents a framework that combines the TAM, the designed behavior theory and the DeLone and McLean success model, in order to explore the factors influencing the purpose and the use of such services. Skordoulis et al. (2017) study the TAXIS information system and examine the satisfaction of users with its use, using a multi-criteria methodology. Wang et al. (2008) develop and validate a success model of e-government systems, based on the revised DeLone and McLean success model, that records the multidimensional and interdependent nature of these systems. The main aim of Horan et al (2006)'s work is to create a means for the success of e-government, as shown by the users of such e-services. Regardless of whether their model will be used in the future, they point out that as these services are more widespread, it is necessary to understand the manner in which they are perceived by the taxpayer. The research of AL Athmay et al. (2016) was conducted to investigate the dimensions affecting the adoption of e-government services in the United Arab Emirates, considering the end-user. They are interested in knowing the significance of satisfied users and the effect they have on user intention for these services.

However, system developers are also considered employees, since they are primarily called upon to use the new applications either voluntarily or out of compulsion. Dukic et al. (2017) examine the level of computer skills of staff in the public sector and the degree to which they uphold e-government. Using a questionnaire from Croatian central government officials, they concluded that the official felt they were very specialized and did not resist the change. It is considered that some improvements in e-services need to be made. Stefanovic et al. (2016) also explore the success of such systems from the angle of employees. The findings verify the validity of the DeLone

and McLean model in e-government. Floropoulos et al. (2010) investigate the TAXIS system using employees in Public Financial Services. This is interesting since this system is applied in a country with a strong taxation system that is mandatory. Terpsiadou et al. (2009), in their study, concluded that most users are generally satisfied with the features of the system. Agrifoglio et al. (2016) use data collected by the administrative staff of two Italian courts to evaluate a management case; while Oktal et al. (2016) surveyed internal justice services in Turkey. Al-Busaidy et al. (2009) carried out a survey of civil servants from three e-government-related ministries. It is revealed in the survey that there is a strong link between the following factors: efficiency, accessibility, availability and trust.

Wang and Liao (2008) using the DeLone and McLean (2003) IS success model investigated the effect of information quality, quality of service, quality of system and use on user satisfaction for e-government technologies. Their results conclude that authorities in the public sector should develop IS which will execute accurate and useful information and a user-friendly system for users to accept. Additionally, the findings of their study highlighted that quality of information has a greater impact on user satisfaction and perceived net benefit than quality of service and system. Therefore, managers in the public sector will concentrate on executing up-to-date and accurate information. Many scholars explored the effect of information quality, service quality, system quality and use on the satisfaction of employees who used municipal e-government systems. The findings of these studies concluded that the quality of service and the technical quality are increasing the satisfaction of staff. Employees have therefore the intention to use systems with a high degree of usability, user-friendliness, and ease of use. User satisfaction is a significant factor for the benefits of local government workers, such as increased efficiency, work performance and effectiveness (Sachan et al., 2018; Stefanovic et al., 2016).

In e-government in particular, scholars have measured user satisfaction which adapt three factors: quality of the information, quality of the service and quality of the system. The first factor tests the content of IS containing variables such as precision, currency, timeliness of performance, reliability, completeness, mindfulness, ease of use and adequate amount of information. Level of service quality allows workers in the public sector to carry out their day-to-day work activities. Therefore, factors such as information production, the user-friendly interface, system compatibility and technical staff skills are essential to help users. The third aspect pertains to IS

production efficiency. Quality of service involves variables such as information completeness, precision, format, currency, importance, timeliness, accuracy, validity, usability, and conciseness to calculate the user satisfaction impact on this aspect. Internal justice system users indicated that the quality of system and service has a direct but not high and positive impact on user satisfaction. Their expectations are focused on the quality of information, perceived ease of use and the interface of the system because the main goal is the improvement of their work. Users require timely information by accessing data in real-time; correct information, fewer incorrect data entries and more consistent data entry across users over time. If the procedure related to legal assistance is complicated, the speed of the system is poor and technical staff cannot provide the help needed, therefore users are not equipped to use the court system. Court administrative staff suggested that the greater the effect it would have on job efficiency, the more system is used and the court employees are pleased with it. Findings from previous surveys thus indicate courts that the availability of information influences user satisfaction rather than the efficiency and usage of the system (Agrifoglio et al., 2016; Oktal et al., 2016).

## **2.2 Information Systems in courts**

Looking at the literature of IS, one will not come across a plethora of examples from applications in the field of justice, as is the case in other areas. The introduction of IT in e-government provides uncomplicated and systematic access to all public services for citizens, companies and other public organizations, with the use of such technologies being considered one of the main factors that changed the administration of justice (Velicogna, 2007). As part of public-sector reforms in e-government, e-justice information systems are contributing to a fairer, more efficient and transparent justice system. And although the challenges of introducing such systems in the courts are real, they are becoming more and more common, both in Europe and all over the world (Urbach et al., 2010).

An example is the case study of (Agrifoglio et al., 2016), on the effectiveness of such applications in Italy. The researchers say that although the use of these applications in Italian courts is mandatory for employees, the use of documents is a widespread activity, while the usage of applications is not compulsory for judges and lawyers. Using a questionnaire and applying the DeLone and McLean model, their results conclude that there is an important link between the success of the Electronic Justice

model and the quality of the system, which is a crucial dimension in the use of such systems. De Vuyst and Fairchild (2006) present the Phoenix project, which was launched in Belgium in 2001; it is the key step in a strategy for implementing e-justice systems based on electronic files and open source standards. Comparing this plan with similar systems in other countries, some challenges may hinder the success of the project: the security of the system, the required legislative changes and the change in the mentality of justice staff. Lienhard and Kettiger (2017), in a case study of Switzerland, expressed their view that justice management should mainly or solely secure the effective defense of legal rights, as well as the proper distribution of public resources; while the continuous communication within the bodies involved both domestically and internationally is confirmed by research as a particularly critical element. It is worth mentioning here that there is a development of an evaluation model of Turkey's national justice information system, as presented in Oktal et al (2016)'s study. After developing the basic theories of information systems, the authors use variables from various models to conduct research on 8,840 justice officers in Turkey; they identify the main problem of the system, which is the quantity and complexity of the processes involved in its implementation. Therefore, users are expected to simplify the functions and interfaces of the system. It should be noted that this study is the first assessment model of an effective electronic justice system, to consider the internal user, in Turkey.

Some examples of e-justice applications are also found in Greece. Sarantis (2017)'s research explores the effectiveness of Greek courts and the problems that arise and presents an upgraded system that aims to improve their performance. This is the first report in the Integrated Justice Case Management System (IACS). At the same time, Sarantis (2017) study the case of electronic criminal record in Greece. The general purpose of a criminal record information system is, according to the authors, to develop an electronic criminal record entry that will produce an automated version of the applicant's certified criminal record and concurrently provide the citizen with the possibility of submitting an electronic application. Also, the study of Deligiannis and Anagnostopoulos (2017) presents the ICT use standards in Greek courts, by judges and justice officers, in addition to the degree of acceptance, perceived usefulness, perceived ease of use and intention of the user to work with the new system of the IACS. Using the TAM model and a small sample of participants in the research, they

conclude that, although those involved are largely familiar with ICT technologies in the courts, they appear cautious in using this new integrated justice system.

Going beyond European borders, it is worth mentioning Malaysia; it is a country with a highly advanced electronic justice system, whose current legal system was shaped by the combination of Islamic and British legal systems, as well as local customary law. Saman et al. (2013) in their two studies, examine this peculiar relationship and the way it coincides with the developments of technology. In their first study, they look at the case study of a comprehensive electronic case management system. In particular, with a qualitative interpretive approach, they focus on the four main types of applications contained in the E-court's general plan (E-court), noting that the key issues in implementing electronic files are access, security and interoperability. In their second investigation, they observe the electronic application of Sharia law (E-Shariah) in the country's courts. Their research also adopts a qualitative approach and the data they collected helps understand how technology is used both in the court process and in the overall file management cycle. In addition, Hamin et al. (2012) explore the benefits and achievements that have been achieved after the adoption of ICT in the civil justice system of the supreme courts of Malaysia. Through personal semi-structured interviews, they conclude that users benefit from the use of technology and that each of the ICT applications contributes to positive changes in the country's justice system.

However, the implementation of innovative electronic applications in the field of justice often involves risks that need to be addressed promptly for the applications to be successful. Rosa et al. (2013) refer to the risks of developing an information system in the field of justice while examining the case of a similar system in Cape Verde, a developing country in Africa. Different experiences are initially analyzed worldwide, with the important example being Singapore, which was the first country, according to the authors, to develop and implement a justice system. They identify various factors with high risk for the development, implementation and evaluation of such systems, concluding that although the examples they cite are global, coming from different countries and with different content, they share the same risk factors. They conclude that good cooperations between all involved groups and proper training are key steps to eliminate any risks. Kitoogo and Bitwayiki (2010) suggest how to take advantage of current opportunities and evaluation methods to develop IS for courts in Uganda. They refer to the weaknesses of the existing system and propose

changes, hoping that their work will end up in a general framework that can be implemented in other countries as well. The study of Kuhimbisa et al. (2017) focuses on tackling real design gaps in utility dimensions in integrated e-justice projects in the same African country, Uganda. The literature and research were used to identify gaps in information, technology, processes, skills, and management structures in such systems. The new justice information system in Rwanda is finally presented by Watson et al. (2017) and specifically the basic functions of the system and how it is implemented. It is an award-winning system in Africa, which was launched in 2016 and serves as the only entry point for all departments of justice. They point out the key points for its success, with the training of public officials before, during and after its installation as a key first step in tackling any restrictions.

### **3.0 Methodology**

#### **3.1 Sampling**

In the present work, the samples that will be examined are two and concern the two groups of users of the information system: the employees in the court in Thessaloniki and lawyers. 125 lawyers who use e-justice system in the court in Thessaloniki and 121 employees in the court in Thessaloniki completed the questionnaire. We used two user groups in order to compare the factors that are more significant for each group. The results of this comparison will help the developers of e-justice systems in order to improve the aspects that are significant for each group and increase the efficiency and effectiveness of the system.

#### **3.2 Questionnaire**

The questionnaire used comes from previous research and incorporates the two main research trends derived from the literature about user satisfaction and technology acceptance (Davis, 1989; DeLone and McLean, 2003; Oktal et al., 2016; Rai et al., 2002). In other words, it is based on both the DeLone and McLean success model for IS and Davis' TAM. Such a combination model helps identify the degree to which a specific system fulfills its demands and proves its value, through the visual gaze of its immediate recipients, its users. Moreover, the use of variables in both models allows for a more comprehensive view of the application of such information systems, as it incorporates both objective and subjective elements of their definition (Agrifoglio et al., 2016; Oktal et al., 2016).

The stability, accuracy and suitability of the hardware and software that provide the required information can be described as system quality. That is, it is recognized by technical characteristics related to the network and computer equipment, and the determining criteria for its evaluation are the performance characteristics of the systems under investigation (Oktal et al., 2016). According to the prevailing success model, the quality of a system belongs to the technical level, while the quality of information belongs to the semantic level (Rai et al., 2002).

Quality of information is a multidimensional concept of understanding, relevance, completeness and effectiveness of information generated by an IS (Delone and McLean, 2003). Information system literature agrees that system quality and performance affect users' behavior (Delone and McLean, 2003; Rai et al., 2002) and that system quality and information quality are twofold factors of user satisfaction. Therefore, users who consider a system to be accurate, precise and timely would find it more pleasant and rewarding to use it (Agrifoglio et al., 2016).

Service quality benefits from a comparison of the preferences of users with the actual output of the services (Parasuraman et al., 1985). It includes two perspectives, adequacy and access and is essentially expressed through the help provided to users to perform their work and the support provided where required. Therefore, it is considered that better service quality will increase the efficiency and satisfaction of the internal user (Oktal et al., 2016). Furthermore, the findings of the current literature show that the success of the implementation of IS is largely identified by the quality factors (Stefanovic et al., 2011).

In defining the perceived ease of use, Davis (1989) refers to "the extent to which a user of a specific system thinks that it would be used without effort". Perceived ease of use positively influences perceived usefulness, because the simpler the system is to use, the more helpful it becomes (Oktal et al., 2016). Previous research confirms the importance of system usability in user satisfaction (Hudson et al., 2018). Indeed, users' attitude towards a system is identified by the perception it generates about its usefulness and ease of use, ie the perceived usefulness and perceived ease of use. This mindset shapes the willingness to use a system, which in turn describes the system's actual use.

According to Davis (1989) perceived usefulness can be identified as "the extent to which a user thinks that the use of a specific system would increase his or her work performance". When measuring this variable, it is analyzed whether the work is

completed quicker and more effectively and whether the use of the system is beneficial to the organization or company that uses it. The perceived usefulness has a strong causal correlation with user satisfaction and perceptions of utility derived from personal assessments of information systems (Rai et al., 2002).

To operationalize the above-mentioned constructs above the five-point Likert scale was used. Analysis of the data was implemented using Factor Analysis.

#### 4.0 Results

The internal consistency and reliability, calculated via Cronbach's alpha, ranged from 0.904 to 0.968, exceeding the minimally required 0.70 level (Newkirk et al., 2003).

Table 1 shows the Cronbach's alpha coefficient for all variables.

Variables	No. of items	Cronbach a for employees	Cronbach a for lawyers
Satisfaction	5	0.917	0.899
System quality	5	0.904	0.899
Information quality	7	0.968	0.890
Service quality	7	0.930	0.921
Perceived ease of use	5	0.907	0.893
Perceived usefulness	3	0.961	0.899

**Table 1. Reliability analysis of the questionnaire items for employees and lawyers.**

Table 2 presents the principal component analysis using the Maximum Likelihood Estimate and the extraction of factors with Promax with Kaiser Normalization method. The factor loadings and cross loadings provide support for convergent and discriminant validity.

Factors	Items	Loadings
System quality	Accessibility to the system	,929
	User-friendly	,890
	Easy to use	,878
	Usability	,728
	Integration with other systems	,835
Information quality	Precise information	,661
	Up-to-date information	,595
	Sufficient information	,639



	Reliable information	,635
	Useful information	,572
	Accessibility	,504
	Output of information	,599
Service quality	Readiness for service	,925
	Safe transactions	,846
	Availability	,910
	Individual attention	,900
	Specific needs of users	,979
	Skills for technical employees	,795
	Willingness of technical employees	,773
Perceived ease of use	Ease of use	,556
	Ease of learning	,500
	Interaction	,574
	Skills	,551
	Flexibility	,523
Perceived usefulness	Improvement of job performance	,895
	Accomplishment of tasks more quickly	,711
	Easier job	,814

**Table 2. Factor loadings.**

## 5.0 Discussion

Examining the results from the group of lawyers, it appears that indeed, the quality of the system, the quality of the information, the quality of the service and the perceived usefulness of the system are the variables that positively affect the overall satisfaction of the specific users, while the perceived ease of use of the system does not seem to affect it much. In this group, the quality of the system affects the overall satisfaction more than the other variables. As the use of the system by lawyers is different from that of employees, it seems that this group is more interested in an overall good picture of the system, ensuring the quality dimensions of the application, which will make it useful for its daily work activities.

Comparing the results of the present work with those of the authors of the articles in the literature review, it is worth noting that their findings are mainly coincidental. Indeed, according to previous research, all three dimensions of quality have

immediate and positive effects on the satisfaction of internal users of respective systems, with each of these dimensions playing a more important role in overall satisfaction, depending on the research under consideration (Agrifoglio et al., 2016; Oktal et al., 2016; Stefanovic et al., 2011). Nevertheless, the ease of use and perceived usefulness, the two key dimensions in accepting information systems have a positive sign in most surveys, that refer to the extent to which users think the system can help them perform their work better (Hardyanto et al., 2018; Oktal et al., 2016; Rai et al., 2002; Sachan et al., 2018).

## **6. Conclusion**

### **6.1 Limitations and suggestions for future research**

During the survey, there was a restriction on the size of the sample. This paper investigated the satisfaction of users with the new application, in all the courts that use it on a pilot basis. However, the choice to conduct the survey only in the courts of Thessaloniki made the sample relatively small compared to most empirical surveys that have used the questionnaire method. In addition, without the creation of a broader geographical analysis, it has not been possible to evaluate justice performance at a broader organizational level. Although evaluating the satisfaction of employees in the courts under study is a first step in understanding the performance of all users, future could use a representative national sample of users, so any generalization of results should be done with special care.

Behavioral IS usage models could be used by future researchers in order to clarify IS usage in various settings (such as operational, tactical and strategic level) where IS usage can be measured through time spent on the system. The results of the study reflect the attention provided to enhancing the efficiency and performance of e-justice systems by court users, authorities and suppliers of applications to consider these factors in the design and use of court systems. Besides, this study is helpful to justice authorities and professionals in order to develop those systems more effectively and to carefully consider these factors in the development and use of court systems. Due to the growing use of IT for the supply of public services, a greater understanding of such constructs necessary for increased acceptance. For agencies which provide e-justice services, it may also be crucial.

Extending the research nationwide to the justice services that currently pilot the application will give more and more reliable research results, highlight any

differences in the use of the system from region to region and present an overall picture of satisfaction or not of its users. Finally, a new study at a later date and after a long period of application of the information system would allow us to compare the results of the two surveys, to highlight any failures and errors of the application and to highlight the role of technology in the performance of Greek courts. In this way, the findings would provide useful information to the justice authorities and IT professionals to improve the application, its extension and use by other parties involved in the court, such as justice officers. This will undoubtedly contribute to a faster and more just administration of justice.

## References

- Agrifoglio, R., Metallo, C. and Lepore, L. (2016) *Success factors for using case management system in Italian courts*, Information Systems Management, 33 42-54.
- Ajzen, I. (1991) *The theory of planned behaviour*, Organizational Behavior and Human Decision Processes, 50 179-211.
- Al-Busaidy, M., Weerakkody, V. and Dwivedi, Y.K. (2009) *Factors Influencing eGovernment Progress in Oman: An Employee's Perspective*, In Proceedings of the 15th Americas Conference on Information Systems (AMCIS), San Francisco, California, USA.
- Angelopoulos, S., Kitsios, F., Kofakis, P. and Papadopoulos, T. (2010). *Emerging Barriers in E-Government Implementation*, In Electronic Government. EGOV 2010. (Eds. Wimmer, M.A., Chappelet, J.L., Janssen, M., Scholl, H.J.) Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, vol. 6228, pp. 216-225.
- Angelopoulos, S., Kitsios, F. and Babulac, E. (2010) *From e to u: Towards an innovative digital era*, In Ubiquitous and Pervasive Computing: Concepts, Methodologies, Tools, and Applications. (Ed. Symonds, J.) IGI Global Publishing, Hershey, PA, USA, pp. 1669–1687.
- Anwer, A.M., Esichaikul, V., Rehman, M. and Anjum, M. (2016) *E-government services evaluation from citizen satisfaction perspective: A case of Afghanistan*, Transforming Government: People, Process and Policy, 10 139-167.
- Athmay, A.A.A.A., Fantasy, K. and Kumar, V. (2016) *E-government adoption and user's satisfaction: an empirical investigation*, EuroMed Journal of Business, 11 57-83.
- Charalabidis, Y., Loukis, E., Lachana, Z. and Alexopoulos, C. (2019) *Future Research Directions on the Science Base and the Evolution of the Digital Governance Domain*, In Proceedings of the International Conference on Information Systems (ICIS 2019), Munich, Germany.
- Danila, R. and Abdullah, A. (2014) *User's Satisfaction on E-government Services: An Integrated Model*, Procedia - Social and Behavioral Sciences, 164 575-582.
- Davis, F.D. (1989) *Perceived usefulness, perceived ease of use, and user acceptance of information technology*, MIS Quarterly, 13 319-340.
- Deligiannis, A.P. and Anagnostopoulos, D. (2017) *Towards Open Justice: ICT Acceptance in the Greek Justice System the Case of the Integrated Court*

- Management System for Penal and Civil Procedures (OSDDY/PP)*, In Conference for E-Democracy and Open Government (CeDEM), Krems, Austria.
- De Vuyst, B. and Fairchild, A. (2006) *The Phenix project: a case study of e-justice in Belgium*, In 8th International Conference on Electronic Commerce: The new e-commerce - Innovations for Conquering Current Barriers, Obstacles and Limitations to Conducting Successful Business on the Internet, New York.
- Delone, W.H. and McLean, E.R. (2003) *The DeLone and McLean model of information systems success: a ten-year update*, Journal of management information systems, 19 9-30.
- Dukić, D., Dukić, G., Bertović, N. (2017) *Public administration employees' readiness and acceptance of e-government: Findings from a Croatian survey*, Information Development, 33 525-539.
- Faro, B., Abedin, B. and Cetindamar, D. (2021) *Hybrid organizational forms in public sector's digital transformation: A technology enactment approach*, Journal of Enterprise Information Management (in press).
- Floropoulos, J., Spathis, Ch., Halvatzis, D. and Tsiouridou, M. (2010) *Measuring the success of the Greek Taxation Information System*, International Journal of Information Management, 30 47-56.
- Hamin, Z., Bahrin, O. M. and Mohamad, A.M. (2012) *Benefits and achievements of ICT adoption by the High Courts of Malaysia*, In 2012 IEEE Symposium on Humanities, Science and Engineering Research, Kuala Lumpur.
- Hardyanto, W., Purwinarko, A. and Adhi, M.A. (2018) *End-user satisfaction analysis on library management system unnes using technology acceptance model towards national standard of integrated library*, Journal of Physics: Conference Series, 983 012006.
- Horan, T.A., Abhichandani, T. and Rayalu, R. (2006) *Assessing User Satisfaction of E-Government Services: Development and Testing of Quality-in-Use Satisfaction with Advanced Traveler Information Systems (ATIS)*, In Proceedings of the 39th Hawaii International Conference on System Sciences (HICSS'06), Kauia, HI, USA.
- Hudson, D., Kushniruk, A., Borycki, E. and Zuege, D. (2018) *Physician Satisfaction with a Critical Care Clinical Information System Using a Multimethod Evaluation of Usability*, International Journal of Medical Informatics, 112 131-136.
- Kitoogo, F.E. and Bitwayiki, C. (2010) *E-Justice Implementation at a National Scale: The Ugandan Case*, In E-Infrastructures and E-Services on Developing Countries. AFRICOMM 2009. E-Infrastructures and E-Services on Developing Countries. (Eds. Villafiorita A., Saint-Paul R. and Zorer A.) Springer, Berlin, Heidelberg, vol. 38, pp. 40-49.
- Kamariotou, M. and Kitsios, F. (2019) *Critical Factors of Strategic Information Systems Planning Phases in SMEs*, In Information Systems, EMCIS 2018, Springer LNBIP 341. (Eds. Themistocleous, M. and Rupino da Cunha, P.) Springer Nature, Cham, Switzerland, pp. 503–517.
- Kamariotou, M. and Kitsios, F. (2017) *An empirical evaluation of strategic information systems planning phases in SMEs: determinants of effectiveness*, In Proceedings of the 6th International Symposium and 28th National Conference on Operational Research, Thessaloniki, Greece.

- Kitsios, F., Angelopoulos, S. and Papadogonas, T. (2009) *Innovation and Strategic Management in e-Government Services*, International Journal of Trade in Services, 1 35-42.
- Kitsios, F. and Kamariotou, M. (2017) *Strategic IT alignment: Business performance during financial crisis*, In Advances in Applied Economic Research, Springer Proceedings in Business and Economics. (Eds. Tsounis, N. and Vlachvei, A.) Springer, pp. 503-525.
- Kuhimbisa, E., Baguma, R. and Nakakawa, A. (2017) *A Model for Developing Usable Integrated Case Management Information Systems*, In International Conference on Social Implications of Computers in Developing Countries, Yogyakarta, Indonesia.
- Lienhard, A. and Kettiger, D. (2017) *Between management and the rule of law. Results of the research project "basic research into court management in switzerland"*, International Journal for Court Administration, 8 7-19.
- Loukis, E. and Charalabidis, Y. (2011) *Why do eGovernment projects fail? Risk factors of large information systems projects in the Greek public sector: An international comparison*, International Journal of Electronic Government Research (IJEGR), 7 59-77.
- Loukis, E.N. and Tsouma, N. (2002) *Critical issues of information systems management in the Greek Public Sector*, Information Polity, 7 65-83.
- Newkirk, H.E., Lederer, A.L. and Srinivasan, C. (2002) *Strategic information systems planning: too little or too much?*, Journal of Strategic Information Systems, 12 201-228.
- Oktal, O., Alpu, O. and Yazici, B. (2016) *Measurement of internal user satisfaction and acceptance of the e-justice system in Turkey*, Aslib Journal of Information Management, 68 716-735.
- Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1985) *A conceptual model of service quality and its implications for future research*, Journal of Marketing, 49 41-50.
- Rai, A., Lang, S.S. and Welker, R.B. (2002) *Assessing the validity of IS success models: An empirical test and theoretical analysis*, Information systems research, 13 50-69.
- Rosa, J., Teixeira, C. and Pinto, J.S. (2013) *Risk factors in e-justice information systems*, Government Information Quarterly, 30 241-256.
- Sachan, A., Kumar, R. and Kumar, R. (2018) *Examining the impact of e-government service process on user satisfaction*, Journal of Global Operations and Strategic Sourcing, 11 321-336.
- Saman, W.M.W.S. and Haider, A. (2013) *E-Shariah in Malaysia: technology adoption within justice system*, Transforming Government: People, Process and Policy, 7 256-276.
- Sang, S. and Lee, J.-D. (2009) *A Conceptual Model of e-Government Acceptance in Public Sector*, In Proceedings of the 3rd International Conference on Digital Society (ICDS 2009), Cancun, Mexico.
- Sarantis, D. (2017) *The Challenge of Accelerating Greek Judicial Procedure*, In International Conference on Electronic Government and the Information Systems Perspective, Lyon, France.
- Skordoulis, M., Alasonas, P. and Pekka-Economou, V. (2017) *E-government services quality and citizens' satisfaction: a multi-criteria satisfaction analysis of TAXISnet information system in Greece*, International Journal of Productivity and Quality Management, 22 82-100.

- Stefanovic, D., Marjanovic, U., Delić, M., Culibrk, D. and Lalic, B. (2016) *Assessing the success of e-government systems: An employee perspective*, Information & Management, 53 717-726.
- Stefanovic, D., Mirkovic, M., Anderla, A., Drapsin, M., Drid, P. and Radjo, I. (2011) *Investigating ERP systems success from the end user perspective*, Technics Technologies Education Management, 6 1089-1099.
- Terpsiadou, M.H. and Economides, A.A. (2009) *The use of information systems in the Greek public financial services: The case of TAXIS*, Government Information Quarterly, 26 468-476.
- Urbach, N., Smolnik, S. and Riempp, G. (2010) *An empirical investigation of employee portal success*, The Journal of Strategic Information Systems, 19 184-206.
- Velicogna, M. (2007) *Justice systems and ICT - What can be learned from Europe?*, Utrecht Law Review, 3 129-147.
- Wang, Y.S. and Liao, Y.W. (2008) *Assessing eGovernment systems success: A validation of the DeLone and McLean model of information systems success*, Government Information Quarterly, 25 717-733.
- Watson, A.C., Rukundakuvaga, R. and Matevosyan, K. (2017) *Integrated Justice: An Information Systems Approach to Justice Sector Case Management and Information Sharing*, International Journal for Court Administration, 8 1-9.
- Weerakkody, V., Irani, Z., Lee, H., Hindi, N. and Osman, I. (2016) *Are U.K. Citizens Satisfied With E-Government Services?. Identifying and Testing Antecedents of Satisfaction*, Information Systems Management, 33 331-343.
- Wirtz, B.W. and Kurtz, O.T. (2016) *Local e-government and user satisfaction with city portals – the citizens' service preference perspective*, International Review on Public and Nonprofit Marketing, 13 265-287.

# On the Temporal Experiences of Remote Workers

Efpraxia D. Zamani, Durham University Business School

## Abstract

*This paper advocates for an existentialist viewpoint that recognizes time as both objective and subjective, allowing for an exploration of the emotional dimension of time. Utilizing a study involving remote workers, the research investigates affective temporal experiences, the role of technology in mediating these encounters and the emerging observed behaviours. Inspired by Existentialism and the Negative Theology of Time, the central research question addressed is: "How does technology influence chronopathic experience?" This approach extends beyond conventional productivity concerns to understand how technology commodifies personal time and shapes distinct temporal experiences.*

**Keywords:** remote work, time, temporality, affective experience, knowledge work

## 1.0 Introduction

The concept of time is a complex and multifaceted notion, which has been explored across various academic disciplines, yielding diverse and sometimes contradictory perspectives on what time is, how temporality is perceived, and the true ontology of time (Bergmann, 1992; Sider, 2006). However, while IS scholars have delved into time and temporality, they have largely approached time in a binary manner. Time is either viewed as linear and objective, tied to the notion of clock time (Ancona et al., 2001) or as subjective, seen through a practice-based lens, where time is considered a social construct (Orlikowski & Yates, 2002).

However, time is a fundamental aspect of the human experience, influencing it profoundly. The unstoppable nature of time means that time is finite and therefore people are limited in how they can use it to meet their goals. Yet, viewing time as either subjective or objective oversimplifies temporal experiences and human responses. Within the IS discipline, going beyond these concepts is critical as insights in terms of affective temporal experiences can inform the design of IT artefacts, that up to date, have been rather concentrated around concepts on efficiencies and productivity.

This paper argues for an existentialist perspective that recognizes time as both objective and subjective, which allows exploring the emotional aspect of time. The paper draws from a study that involves remote workers and delves into the intricate

relationship between remote workers, time and technology, examining temporality, the emotional impact of the passage of time on remote workers, and the role of technology in mediating these experiences.

Drawing from Existentialism and the Negative Theology of Time (Theunissen, 1986), time in this study is seen as *aion*, meaning an eternal and continuous flow with no beginning or end. Within this context, two central concepts are leveraged: chronopathic experience (or what can be understood as the affective aspect of temporal experiences) and chronotelic behaviour, i.e., the way in which individuals attempt to overcome the passage of time.

The research question addressed is: "How does technology influence chronopathic experience?" This approach goes beyond conventional productivity and performance considerations to understand how personal time becomes commodified and how technology shapes distinct temporal experiences.

The paper starts with an overview of existentialism and the existentialist perspective of time through the Negative Theology of Time. It elaborates on the concepts of chronopathic experience and chronotelic behaviour. The study's methods and framing are discussed, followed by the presentation of findings. The paper concludes by considering theoretical and practical contributions and acknowledging limitations.

## **2.0 Time in Existentialism**

Existentialism places great importance on the notion that "death [is] a pivotal aspect of life's meaning", and it thus underscores that choice and possibilities are boundless, which imposes an inevitable responsibility on each individual (Vandekerckhove, 2020, p. 130). In simpler terms, existentialism suggests that life gains significance because one's time is limited.

This philosophy revolves around the individual and their role in life, aiming to make sense of the human condition, life's circumstances, and the contradictions that shape existence (Gardiner, 2002). Within existentialism, the individual is viewed as free to make choices and construct their own meaning in the pursuit of an authentic existence. Importantly, however, this freedom does not detach the individual from their contextual conditions (MacMillan et al., 2012).

Existentialists consider time a fundamental component of the human experience, intricately woven into one's sense of freedom and responsibility. They believe that



individuals conquer time through a continuous exercise of agency, that allows them to exert control over time and life's events. As such, time is seen as a resource that can be harnessed to create meaning and achieve one's goals within the framework of personal responsibility and values (Hellstrom & Hellstrom, 2002). At the same time, however, existentialism emphasises that time is finite and irreversible, which lends urgency to the human condition, compelling individuals to make their existence and choices meaningful. Such a perspective predominantly directs attention to the present moment, where individuals are encouraged to make the most of it. Yet, humans, seen as reflective entities who seek continuity and certainty amid their freedom, often experience anxiety, uncertainty, despair, and even fear in relation to time (Berenskötter, 2020; Theunissen, 2005).

## **2.1 Negative Theology of Time**

As previously discussed, existentialists assign great significance to the present moment and how individuals navigate it. However, this focus on the present doesn't imply that individuals exist solely in a static state; rather, according to Theunissen, the present can only be understood as the "hermeneutic of the possible" (Thornhill, 1998, p. 8). In this view, the present is influenced by past experiences and shaped by expectations of the future (Theunissen, 1986). This perspective aligns with Heidegger's concept of temporality, which consists of the past (thrownness), present (fallenness), and future (projection), all interconnected in a temporal unity that contributes to our existence (Heidegger, 1953; Wheeler, 2020).

In essence, the past, present, and future, although distinct, are more than a linear sequence; they are interwoven within a temporal unity (Heidegger, 1953; Wheeler, 2020). The objective for people is to exert influence over the present in ways that pave the path for a better future (Söderbäck, 2013). Anticipatory resoluteness regarding future "ways of being" becomes a crucial aspect of this experience (Wheeler, 2020). In focusing on the present, humans do so with the expectation that the future holds the promise of being different and improved (Thornhill, 1998).

The Negative Theology of Time, previously introduced as a framework for exploring the emotional experience of time, is instrumental in both organizational and non-organizational contexts. Holt and Johnsen (Holt & Johnsen, 2019) employed this approach to delve into the essence of time, how it is felt, and how it is experienced. They underscore that despite time offering infinite possibilities, its inherent finitude,

marked by the certainty of death, defines its boundaries. Consequently, due to the finite nature of time and the organizational realm's tendency to dominate over all disposable time (Marx, 1939), time is often associated with suffering, despair, boredom, restlessness, malaise, and isolation (Fisherl, 1993; Lange, 2016; Thornhill, 1998) (*pathos*).

## **2.2 Chronopathic Experience/Chronotelic Behaviour**

The terms *chronos* and *kairos* originate from Aristotelian writings in Greek philosophy. In contemporary usage, they are often used interchangeably to refer to time and time-related concepts, yet each emphasizes different aspects of time. *Kairos* underscores the quality rather than the quantity of time, focusing on the right or opportune moment, while *chronos* pertains to the duration or passage of time.

A key aspect of the concept of the chronopathic experience relates to the individual's engagement with the passing of time and how it is experienced. It involves considering the past (e.g., past actions and choices) during the present moment to shape future actions.

The subjective perception of time becomes problematic when individuals perceive the present as meaningless, a barrier separating them from their desired future. Several examples illustrate this, including boredom resulting from the repetitive nature of life (Egenberger, 2012), confusion stemming from intersecting temporalities (Goodbrey, 2015), and the perception of time as futile when individuals constantly await an unattainable future (Johnsen et al., 2019).

Chronotelic behaviors are responses to the affective temporal pressures individuals experience in their attempt to assert control over time when it dominates them (Johnsen et al., 2019). These behaviours aim to challenge the linear organization of time (Rothbauer & Cedeira Serantes, 2021), which demands maximum productivity. Strategies to overcome modern temporality can be categorized into three classes (Blattner, 2011; Hammer, 2011): attempting to recover a better temporality, seeking to overcome the meaningless present, and resisting the logics of efficiency and commodification with the hope of an emancipatory future. These behaviours can take various forms, such as inmates reading books for pleasure to manage excess time (Garner, 2020), individuals focusing on small tasks to imbue the present with meaning (Johnsen et al., 2019), or entrepreneurs abandoning established courses of action due to a constant struggle against time (Branzei & Fathallah, 2023). Despite the variety of

forms that chronotelic behaviours might take, the common element across them is the objective of enduring and overcoming the passage of time.

### 3.0 Methodology

The study builds on semi-structured interviews with knowledge workers between late 2021 and mid 2022 who had transitioned to remote work during the COVID-19 pandemic and, who continued to do so post the pandemic. Participants were primarily professionals working in R&D, and Higher Education. Most were based in the UK, but they all had diverse ethnic background (e.g., African-Caribbean, White/Black EU, White/Black British, Asian<sup>1</sup>). Age-wise, they were all between the ages of early 30s to late 50s. The interviews were conducted online aimed to understand their experiences with remote work, technology usage, and how they manage their workdays. The questions were broad, serving as conversation starters, followed by probing questions to delve into their temporal experiences, including their use of technology to manage time.

The data analysis began with a preliminary examination of the empirical material, accompanied by the development of analytical summaries in the form of memos for each interview. The focus of the analysis was on identifying emerging themes, particularly those closely connected to chronopathic experiences. Two major types of chronopathic experiences were identified: ennui and anxiety. Further analysis around these two experiences using Grounded Theory Method (GTM) techniques (Charmazian approach for constructionist GTM) and the Negative Theology of Time theory as a sensitizing device, revealed three primary triggers related to the use of technology (monitoring and surveillance, dependencies, false urgencies), which subsequently led to three distinct types of chronotelic behaviours: prioritizing, ring-fencing, and multitasking (Table 1).

Table 1. Categories, Focused Codes and Initial Codes

Category	Focused Codes	Initial Codes
<i>Chronopathic experiences</i>	Ennui	What's the point of being there Wasting time

<sup>1</sup> The ethnic background markers used are those typically used by the UK government – it is acknowledged that this is limiting and based on certain assumptions and generalisations.

		Feeling trapped in a meeting
	Anxiety	Navigating temporal differences Reminded of productivity Working longer Feeling chased
<i>Triggers of chronopathic experience</i>	Monitoring and Surveillance	Expectations for productivity (monitored) Controlling others/Being controlled Presenteeism
	Dependencies	Beholden to technology Babysitting the technology
	Creating False Urgencies	Technology decides on importance Notifications are not always urgent Providing recommendations Source of stress
<i>Chronotelic Behaviours</i>	Prioritising	Task prioritisation Flagging emails Categorising
	Delimiting/Ring fencing	Time boxing Blocking time Status changing
	Multitasking	Doing many things at once Being pseudo efficient Constantly flipping between tasks

#### 4.0 Findings

The circumstances and the typical workday of participants varied greatly, but largely what they had in common was they all felt, at some point, the paradox of time: time flying by/time standing still. To the extent that this relates to technology, the resulting experiences discussed resembled either ennui or anxiety.

Ennui can be understood as a lack of interest and dissatisfaction, and it is often described as a passive feeling, whereby the individual simply “surrender[s] to emptiness” (Barbalet, 1999, p. 634), and it is often associated to boredom; yet, contrary to boredom, ennui is as a deeper feeling that involves ‘feeling empty’ or ‘emptiness’ as a result of lack of purpose in one’s life or meaning in the activities in which they engage (including life).

In this study, **ennui** was expressed particularly in terms of how technology often acts as an instrument for managerial control (**monitoring and surveillance trigger**), rather than as a true productivity tool, whereby remote workers see no meaning in using an IT tool for the purpose of enabling perceived organisational control:

*“I always felt that half of our meetings are useless, redundant, and we only have them so that we show our face, declare our presence. It could be, the reason we meet,*

*it could had been an email. And that would be the end of it, and we wouldn't need to waste an hour of our lives. (...) It's crazy they are asking us to have our cameras on. I haven't had a proper home office for so long. I had one, but now [partner] also works from home, and I am working from the kitchen, the dining room, at the living room. It is weird. I don't see the point.” (i15)*

What the extract above indicates is a high level of irritation because of 'useless' meetings and mandates for switched-on cameras. While this can lead to presenteeism, in this case it is also understood as a way for the organisation to extend its control outside the periphery of its physical boundaries and into the private space of the remote worker.

**Anxiety** was expressed through feelings of worry and fear in terms of what might or might not happen, especially in the future (**dependency and false urgencies triggers**). In our findings, such IT-induced anxiety relative to the temporal dimension emerged primarily because of one's dependency on technology and how it helps remote workers navigate time and temporal differences:

*“Managing the different time differences and time zones, you know, so if I'm working with somebody in, say, [country], for instance, and they're like a day ahead of me, you know. (...) I use calendars quite a lot, several calendars. And some of the calendars I can set, I can set them for the country that I'm working with, do you know what I mean? (...) I can set my calendar to [country] time and it will also tell me what the time, what the local time is here. So I use several different calendars. (...) So just imagine, just using, setting up different calendars according to the country (...) it's important for me to not get into a situation where I am working across multiple time zones because that's when I start dropping things, you know. Even though I have reminders, it's too much and I need time, I need time to move between, from one to the other thing, from one time zone to the next. And, you know, get my mind to understand that” (i9)*

In other cases, automated notifications make participants wonder about the accuracy of the information communicated, and, to the extent that such notifications are meant to improve productivity, they can be an additional source of stress, further leading to feelings of anxiety:

*“Previously I had used MyAnalytics by Microsoft that helps you understand how you spend your time. But after a point I stopped, because I saw no value and because it was annoying me and because it kept on sending me too many notifications and these annoyed me even more, because it was stressing me out [prompting why it was stressful] So basically, these tools are meant to be supportive, to help you be more productive but in the end it was frustrating me because you don’t know, so these stuff, they don’t really know how you actually spend your time, and what they count is the hours on your calendar, how much time I spend in meetings and so on. And if I forget to remove a meeting [that was cancelled]. And it sends you reports, you worked X days on this. Ok, I get irritated even knowing this, I don’t want to know. ‘Well done, you messed up’, I know this, now I have you, too, to tell me that.’ (i8)*

### **Triggers of Chronopathic Experiences**

Technology can trigger a chronopathic experience under the following circumstances: when remote knowledge workers perceive the organizational use of technology as an effort to monitor their performance and engagement; due to technology's tendency to create artificial "priorities" and "urgencies"; and because it can both blur and erect barriers between remote workers and their colleagues, as well as between work and their personal lives.

Technology is viewed as a tool that can be and often is employed for monitoring productivity and performance, as well as for surveilling the engagement levels of remote workers in work-related tasks. Many participants questioned whether their organizations trust them enough to meet deadlines and fulfil tasks while working outside the organization's physical premises. They also commented on the use of mandatory camera usage during online meetings and the frequency and utility of frequent online meetings for updates and other purposes. In several instances, shared online calendars are utilized to indicate availability (e.g., when an individual is available for an online meeting) and "busyness" (e.g., how engaged someone is). Participants noted that such calendars can potentially be exploited by management to scrutinize the activities of workers, including when and possibly from where they work.

*“I'm at home in the UK, you don't need to judge me that I may be abroad, because people know that all my family is abroad. And I felt that if it wasn't in an online setting, I wouldn't have worried about what people would say about me or what I'm doing” (I1)*

Based on our analysis, these uses of technology often lead to presenteeism and feelings of frustration, rather than fostering genuine engagement. This is because technology is perceived as a tool that management employs to exert control (e.g., through mandated camera usage and public shared calendars) and monitor perceived engagement and productivity based on time spent. However, in knowledge work, engagement and productivity are typically measured by output and quality rather than by the amount of time spent.

Technology serves as an essential tool for remote workers, enabling them to engage in work-related activities, collaborate with colleagues, coordinate with their team, and participate in various work tasks. Consequently, technology gives rise to dependencies, as remote workers rely on these technological tools. However, the extent of these dependencies can sometimes prove counterproductive, leading to frustration, especially when the technology malfunctions or becomes unavailable. Simultaneously, technology introduces distractions from primary tasks, often through features designed to enhance productivity or task management. For instance, automated reminders for scheduled meetings and notifications for pending or upcoming tasks are valuable for tracking purposes. Yet, these features also foster dependency and can become distracting, particularly when they extend to personal devices:

*“I used to have reminders and notifications. And that really, that really didn't work for me. I was babysitting my phone. I was babysitting my calendar, I was babysitting my email. And it really got to me to the point of I felt like I was too dependent and distracted. So notifications... I couldn't see something, an alert on my phone, for example, and not check it. I stopped doing that” (I15)*

These reminders, notifications, and "smart" technological features can also serve as triggers for chronopathic experiences, as remote workers feel compelled to respond to them within a limited timeframe. Notifications can arrive at any time during the day or night, exerting pressure on the worker, regardless of the true urgency or importance of the message's origin. In such scenarios, these technological features become sources of stress, continually reminding workers of pending tasks and upcoming events, regardless of whether they have sufficient time to address them.

Moreover, modern technology incorporates background analytics tools that can automatically schedule activities like focused work time and help with time management. However, these tools can also result in situations where technology, by learning from an individual's work patterns, determines the significance of certain tasks, people, and when these activities should occur. In the study, examples included the automated scheduling of breaks and lunches, as well as the ability for "important contacts" to override blocked time slots and focus periods, allowing them to schedule meetings in a person's already busy schedule. In essence, such technological tools have the potential to create a sense of artificial urgency.

*“So I might have a colleague, I think my boss can message me on Teams to say are you free [...] and Teams will tell you, only the most important people will get through the focus time [...] maybe because they see how much I interact on my – like on one corner of my screen it kind of says things like, the people you interact with more, your team, and it’s not about who I line manage.” (15)*

### **Chronotelic Behaviours**

Remote workers often utilize technology to overcome the chronopathic experience. In this study, such behaviours take the form of prioritizing important tasks or responsibilities, setting boundaries to protect time, and multitasking to maximize efficiency.

Prioritizing tasks is a common chronotelic behaviour among remote workers due to the limited nature of time. This behaviour helps remote workers navigate the challenges associated with time constraints. They make choices about what tasks to focus on or who to engage with, considering the constraints they face. These



constraints may revolve around deadlines, task significance, or personal/professional obligations, such as requests from valued colleagues or supervisors.

For example, one participant expressed their approach to prioritizing tasks, giving importance to personal considerations, like helping a friend:

*"I try to prioritize things, based on deadlines or who's asking. But also based on what is important for me. I did not use to do that. I used to prioritize based on what input they needed from me to continue working. But now I don't do that. If my boss asks, it's different, or if [friend] needs help..." (I15)*

One common approach to prioritizing tasks is flagging emails for follow-up and setting reminders. This method helps individuals manage their workdays more efficiently and reminds them to address tasks based on their importance and urgency:

*"I usually check my emails while I'm doing the dog walking first, actually, so it's like a pre-email check. Just to see if there are any, what's waiting for me, you know. I guess that's why the very first thing really is to see, you know, has anything happened that needs my immediate attention? You know, basically, can I get on with the jobs that I had to do today as normal or do I have to deviate and do something different in order to respond to whatever's occurred overnight in my inbox?" (I2).*

Multitasking is another prevalent chronotelic behaviour, driven by the perception that "time flies" and there's never enough of it. Participants explained their continuous efforts to maximize productivity within their finite workday. This often results in multitasking behaviour, which can take various forms. One form involves utilizing pockets of time (time fragments) for smaller tasks or household chores that do not demand intense focus:

*"There is no clear boundary really, and okay, I have always been working like this. I have to do school pick up/drop off, maybe pop by the supermarket, I will probably do a wash load. So I basically find pockets of time here and there to do home chores, like maybe between meetings" (I13).*

Alternatively, multitasking may involve managing tasks in parallel, combining personal and work-related activities rather than switching between them constantly. For example, I5 mentioned how the accessibility of Teams on their phone allowed them to seamlessly transition from a work-related debriefing to parenting responsibilities, all within the same timeframe:

*"Because Teams works on the phone as well, so I had a colleague message me saying, 'Are you free? We need to do a debrief about [redacted],' and I said, 'Sorry, I'm bringing my [child] to [classes], but I can call you later.' So while I was waiting for [child's name], Teams rang. I rang on Teams, and it was fine."*

Despite the sense of accomplishment derived from multitasking and accomplishing numerous tasks in a short span, it can also lead to tasks being done less effectively due to frequent distractions. This may cause multitasking to be regarded as a pseudo-efficient behavior:

*"I am very easily distracted by other technology, so it might be my phone. But it's not always personal stuff, it's often other work. It's about trying to be - at least how I see it, it's trying to be even more efficient. So I'm thinking, 'Okay, I'm sitting here in this meeting, but this portion of the meeting isn't relevant to me, so can I do a bit of other work in that time as well?' Yeah, so it's not boredom, it's a kind of pseudo-efficiency. I don't know if it really is efficient or not" (I2).*

Chronotelic behaviours may also entail meticulous time management, with time management involving the careful allocation of tasks across the workday. This ensures that less critical tasks do not take precedence over others and that work does not encroach on personal life. This may involve dedicating specific time blocks to particular tasks, commonly known as time boxing:

*"And I try to also time box, my tasks, the things that I have to do for work, like meetings, filling in reports, preparing reports, and whatnot. I'm trying not to say that I'm going to do this; I say that today, I'm going to spend half an hour doing this. And if that half an hour is not enough, I just put the report aside and work on the next thing. I'll just pick it up next time I have time" (I13).*

Another aspect of this behaviour is using online calendars to indicate availability or unavailability. As these calendars are often shared with colleagues and management, blocking out time slots typically allows for focused work on mentally demanding tasks. It also preserves personal time, such as lunch breaks, as those slots become unavailable for others to schedule meetings:

*"I do block out my lunch hours so that anybody who tries to book a meeting at that time can see that I'm unavailable" (I10).*

## **5.0 Discussion**

Remote workers today have at their disposal a wide array of technologies, most of which are user-friendly and offer the flexibility of access anytime and anywhere. While this has entailed that remote workers can switch easily between personal and professional tasks to accommodate e.g., family demands during the workday, the ubiquity and availability of technology have blurred the boundaries between personal and professional technology use (Choroszewicz & Kay, 2019). As a result, distinguishing between work and non-work activities has become increasingly challenging. Along these lines, it is interesting to explore how IT artefacts, such as shared calendars and collaboration suits influence the affective aspect of remote workers' experience and their relationship to time and temporality.

"How does technology influence chronopathic experience?"

Like earlier studies (e.g., Fonner & Stache, 2012; Park et al., 2020; Prasopoulou et al., 2006), this study confirms that the constant connectivity and accessibility provided by these technologies have made workers, whether consciously or not, vulnerable to the demands of their organizations. We enrich the remote work literature by detailing how online tools, like shared calendars, email applications, and collaboration tools have dissolved the boundaries between work and personal time, and explicitly account for the implications of these for the affective side of remote workers' temporal experience.

We consider that the study brings to the fore important implications for both theory and practice. These technologies have profound implications for time spent outside of work and emotional experiences related to time. The primary rationale behind

introducing technology into the workplace has traditionally been to boost productivity and create more free time for engaging in high-value tasks (Davenport, 1993). The findings show however that this is very often not the case, as the ways in which technology is used and enforced can have negative implications for remote workers, who experience anxiety and stress.

In terms of theoretical implications, we contribute to the extant literature by offering an alternative interpretation for some of the most frequently observed behaviours, such as multitasking. Multitasking has been framed as the result of negative perceptions and feelings regarding progress towards meeting certain goals, whereby multitasking is understood as self-interruption (Adler & Benbunan-Fich, 2013), and a recent large scale study has found that multitasking behaviour is particularly prevalent during online meetings, with detrimental effects for wellbeing (Cao et al., 2021). What this study proposes is that such behaviours, rather than mere attempts to catch up with one's progress, can be understood as attempts to overcome the passing of time, i.e., they have a deep affective origin. As such, there is a clear and important affective component in terms of relating to time that goes beyond claims and concerns of performance and productivity.

Another contribution is that the study offers an explicit account in terms of the role of technology towards these experiences and behaviours. There is an abundance of studies that focus on remote work and the role of technology (e.g., Hafermalz, 2021; Hardill & Green, 2003; Parra et al., 2022; Waizenegger et al., 2020), but most typically technology is black boxed and treated as something that enables collaboration and communication across spatiotemporal boundaries (some exceptions do exist, as for example (e.g., O'Leary & Cummings, 2007)). Contrary to this, in this study we provide a nuanced account where technology, in the form of diverse tools used for collaboration, coordination and monitoring, contributes to the construction of negative affective experiences that can be at times debilitating for remote workers. In other words, technology, rather than being a 'facilitator' and emancipatory by creating efficiencies, establishing and maintaining connections and offering work opportunities, can trigger negative temporal experiences.

## **6.0 Conclusions**

As ICTs become increasingly intertwined with our daily routines and work practices, they become integral to human temporality (Venters et al., 2014), and thus shifting our attention to the affective aspect of temporal experiences can be a useful tool towards understanding more deeply how remote work is experienced.

In this study, chronopathic experiences involve ennui and anxiety, and remote workers follow different behaviours towards overcoming and challenging these. It is important to note however that three identified behaviours are not mutually exclusive; instead, it was noticed that remote workers might change behaviour and move between. There are however certain limitations to this study. How and why certain behaviours may co-exist and relate to each other is subject to each participant's contextual circumstances, and possibly other factors, that require further research. While this study does not offer evidence as to what such factors might be, it could be that whether and how remote workers move between these types of behaviours might relate to how they think about the value of their time and of the task (time worthiness) (Feldman & Greenway, 2021), i.e., whether a task is considered as sufficiently important and worthy of one's time. Similarly, such 'movements' might be attributed to whether and to what extent one's own personal aspirations and long-term objectives align with organisational objectives and demands (goal congruence) (Zhang et al., 2023), and/or with those of their team members and peers more broadly (Choi & Cho, 2019). In terms of limitations, it is also highlighted that the investigation focused on knowledge remote workers, i.e., those who can reasonably work remotely. Other workers (e.g., manual workers) do not have this option. As such, future generalisations are inherently limited.

Concluding, this study calls for a comprehensive understanding of the multifaceted impacts of technology on remote workers. This involves exploring the emotional and affective dimensions of temporal experiences, monitoring the effects of technology on power dynamics, autonomy, and control, and examining the evolving relationship between work modalities, time, and the persistent influence of technology on both the work and non-work domains.

## References

- Adler, R. F., & Benbunan-Fich, R. (2013). Self-interruptions in discretionary multitasking. *Computers in Human Behavior*, 29(4), 1441–1449. <https://doi.org/10.1016/j.chb.2013.01.040>

- Ancona, D. G., Goodman, P. S., Lawrence, B. S., & Tushman, M. L. (2001). Time: A New Research Lens. *Academy of Management Review*, 26(4), Article 4. <https://doi.org/10.5465/amr.2001.5393903>
- Barbalet, J. M. (1999). Boredom and social meaning. *The British Journal of Sociology*, 50(4), 631–646. <https://doi.org/10.1111/j.1468-4446.1999.00631.x>
- Berenskötter, F. (2020). Anxiety, time, and agency. *International Theory*, 12(2), 273–290. <https://doi.org/10.1017/S1752971920000111>
- Bergmann, W. (1992). The Problem of Time in Sociology: An Overview of the Literature on the State of Theory and Research on the 'Sociology of Time', 1900-82. *Time & Society*, 1(1), Article 1.
- Blattner, W. (2011). Review: *Philosophy and Temporality from Kant to Critical Theory*. Notre Dame. Philosophical Reviews. <https://ndpr.nd.edu/reviews/philosophy-and-temporality-from-kant-to-critical-theory/>
- Branzei, O., & Fathallah, R. (2023). The End of Resilience? Managing Vulnerability Through Temporal Resourcing and Resisting. *Entrepreneurship Theory and Practice*, 47(3), Article 3. <https://doi.org/10.1177/10422587211053809>
- Cao, H., Lee, C.-J., Iqbal, S., Czerwinski, M., Wong, P. N. Y., Rintel, S., Hecht, B., Teevan, J., & Yang, L. (2021). Large Scale Analysis of Multitasking Behavior During Remote Meetings. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3411764.3445243>
- Choi, O.-K., & Cho, E. (2019). The mechanism of trust affecting collaboration in virtual teams and the moderating roles of the culture of autonomy and task complexity. *Computers in Human Behavior*, 91, 305–315. <https://doi.org/10.1016/j.chb.2018.09.032>
- Choroszewicz, M., & Kay, F. (2019). The use of mobile technologies for work-to-family boundary permeability: The case of Finnish and Canadian male lawyers. *Human Relations*, 001872671986576. <https://doi.org/10.1177/0018726719865762>
- Davenport, T. H. (1993). *Process Innovation: Reengineering Work through Information Technology*. Harvard Business School Press.
- Egenberger, S. (2012). Michael Theunissen: Fortune and Misfortune of Temporality. In J. Bartley Stewart (Ed.), *Kierkegaard's Influence on Philosophy. German and Scandinavian Philosophy* (pp. 187–209). Ashgate.
- Feldman, E., & Greenway, D. (2021). It's a Matter of Time: The Role of Temporal Perceptions in Emotional Experiences of Work Interruptions. *Group & Organization Management*, 46(1), Article 1. <https://doi.org/10.1177/1059601120959288>
- Fisherl, C. D. (1993). Boredom at Work: A Neglected Concept. *Human Relations*, 46(3), Article 3. <https://doi.org/10.1177/001872679304600305>
- Fonner, K. L., & Stache, L. C. (2012). All in a day's work, at home: Teleworkers' management of micro role transitions and the work-home boundary. *New Technology, Work and Employment*, 27(3), Article 3. <https://doi.org/10.1111/j.1468-005X.2012.00290.x>
- Gardiner, P. (2002). *Kierkegaard: A Very Short Introduction*. OUP Oxford.
- Garner, J. (2020). Experiencing time in prison: The influence of books, libraries and reading. *Journal of Documentation*, 76(5), Article 5. <https://doi.org/10.1108/JD-07-2019-0128>

- Goodbrey, D. M. (2015). Distortions in Spacetime: Emergent Narrative Practices in Comics' Transition from Print to Screen. In R. Pearson & A. N. Smith (Eds.), *Storytelling in the Media Convergence Age* (pp. 54–73). Palgrave Macmillan UK. [https://doi.org/10.1057/9781137388155\\_4](https://doi.org/10.1057/9781137388155_4)
- Hafermalz, E. (2021). Out of the Panopticon and into Exile: Visibility and control in distributed new culture organizations. *Organization Studies*, 42(5), Article 5. <https://doi.org/10.1177/0170840620909962>
- Hammer, E. (2011). *Philosophy and Temporality from Kant to Critical Theory*. Cambridge University Press.
- Hardill, I., & Green, A. (2003). Remote working—Altering the spatial contours of work and home in the new economy. *New Technology, Work and Employment*, 18(3), Article 3. <https://doi.org/10.1111/1468-005X.00122>
- Heidegger, M. (1953). *Being and Time* (Translated by Joan Stambaugh). State University of New York Press.
- Hellstrom, T., & Hellstrom, C. (2002). Time and innovation in independent technological ventures. *Human Relations*, 55(4), 407–426.
- Holt, R., & Johnsen, R. (2019). Time and Organization Studies. *Organization Studies*, 40(10), Article 10. <https://doi.org/10.1177/0170840619844292>
- Johnsen, R., Berg Johansen, C., & Toyoki, S. (2019). Serving time: Organization and the affective dimension of time. *Organization*, 26(1), Article 1. <https://doi.org/10.1177/1350508418763997>
- Lange, E. L. (2016). The Critique of Political Economy and The New Dialectic. Hegel, Marx, and Christopher J. Arthur's 'Homology Thesis'. *Crisis and Critique*, 3(3), Article 3. <https://doi.org/10.5167/UZH-127750>
- MacMillan, S., Yue, A. R., & Mills, A. J. (2012). Both How and Why: Considering Existentialism as a Philosophy of Work and Management. *Philosophy of Management*, 11(3), 27–46. <https://doi.org/10.5840/pom201211318>
- Marx, K. (1939). *Grundrisse. Foundations of the Critique of Political Economy* (M. Nicolaus, Trans.). Penguin Books. <https://www.marxists.org/archive/marx/works/1857/grundrisse/index.htm>
- O'Leary, M., & Cummings, J. N. (2007). The Spatial, Temporal, and Configurational Characteristics of Geographic Dispersion in Teams. *MIS Quarterly*, 31(3), Article 3.
- Orlikowski, W. J., & Yates, J. (2002). It's About Time: Temporal Structuring in Organizations. *Organization Science*, 13(6), Article 6. <https://doi.org/10.1287/orsc.13.6.684.501>
- Park, Y., Liu, Y., & Headrick, L. (2020). When work is wanted after hours: Testing weekly stress of information communication technology demands using boundary theory. *Journal of Organizational Behavior*, 41(6), Article 6. <https://doi.org/10.1002/job.2461>
- Parra, C. M., Gupta, M., & Cadden, T. (2022). Towards an understanding of remote work exhaustion: A study on the effects of individuals' big five personality traits. *Journal of Business Research*, 150, 653–662. <https://doi.org/10.1016/j.jbusres.2022.06.009>
- Prasopoulou, E., Pouloudi, N., & Panteli, N. (2006). Enacting new temporal boundaries: The role of mobile phones. *European Journal of Information Systems*, 15, 277–284.
- Rothbauer, P. M., & Cedeira Serantes, L. (2021). Reading time: Exploring the temporal experiences of reading. *Journal of Documentation*, ahead-of-print(ahead-of-print). <https://doi.org/10.1108/JD-11-2020-0200>

- Sider, T. (2006). Quantifiers and Temporal Ontology. *Mind*, 115(457), Article 457. <https://doi.org/10.1093/mind/fzl075>
- Söderbäck, F. (2013). Being in the Present: Derrida and Irigaray on the Metaphysics of Presence. *The Journal of Speculative Philosophy*, 27(3), 253–264. <https://doi.org/10.5325/jspecphil.27.3.0253>
- Theunissen, M. (1986). *The Other: Studies in the social ontology of Husserl, Heidegger, Sartre, and Buber* (C. Macann, Trans.). The MIT Press. <https://mitpress.mit.edu/books/other>
- Theunissen, M. (2005). Kierkegaard's Concept of Despair. In *Kierkegaard's Concept of Despair*. Princeton University Press. <https://doi.org/10.1515/9780691216195>
- Thornhill, C. (1998). Intersubjectivity and openness to change. Michael Theunissen's negative theology of time. *Radical Philosophy*, 88(March/April), Article March/April.
- Vandekerckhove, W. (2020). COVID, Existentialism and Crisis Philosophy. *Philosophy of Management*, 19(2), 127–132. <https://doi.org/10.1007/s40926-020-00140-2>
- Venters, W., Oborn, E., & Barrett, M. (2014). A trichordal temporal approach to digital coordination: The sociomaterial mangling of the CERN grid. *MIS Quarterly*, 38(3), Article 3.
- Waizenegger, L., McKenna, B., Cai, W., & Bendz, T. (2020). An affordance perspective of team collaboration and enforced working from home during COVID-19. *European Journal of Information Systems*, 29(4), Article 4. <https://doi.org/10.1080/0960085X.2020.1800417>
- Wheeler, M. (2020). Martin Heidegger. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Fall 2020 Edition)*. <https://plato.stanford.edu/archives/fall2020/entries/heidegger>
- Zhang, G., Wang, H., & Li, M. (2023). Leader narcissism, perceived leader narcissism, and employee outcomes: The moderating effect of goal congruence. *Journal of Business Research*, 166, 114115. <https://doi.org/10.1016/j.jbusres.2023.114115>



# From Connections to Classrooms: Utilizing Social Media platforms for Enhanced Engagement: The case of LinkedIn.

*Research In progress*

## **Abstract**

*This study addresses the impact of technology in higher education by examining the potential of social media platforms, particularly LinkedIn Learning, in enhancing student engagement. Despite the prevalence of social networks in education, limited literature explores LinkedIn's classroom applications. The research delves into the relationship between students' self-regulation, self-efficacy, collaboration, and their engagement using LinkedIn Learning online courses. A comparative design assesses changes in student engagement before and after participating in LinkedIn courses. This study uses the data from a survey of postgraduate students in a 2-year MBA Global program at a UK and employs confirmatory factor analysis and structural equation modelling to evaluate the measurement model and test research hypotheses.*

**Keywords:** student engagement, self-regulation, self-efficacy, collaboration, online studies, LinkedIn.

## **1.0 Introduction**

In recent years, higher education has witnessed a surge in the integration of innovative technologies to enhance the learning experience. Learning management systems, virtual reality, and social networking sites have become integral components of modern educational strategies (Aldadouh et al., 2020; Tan & Hsu, 2017). A significant milestone in educational literature can be found in the systematic review conducted by Al-Qaysi et al. in 2023. This comprehensive review addressed the current dearth of knowledge and called for an exploration of how social media can effectively be harnessed to bolster student engagement and participation. Among the plethora of social networking sites, LinkedIn has emerged as a prominent contender for pedagogical applications, especially concerning students' employability and professional networking (Healy et al., 2022). Despite its potential, literature on the utilization of LinkedIn within the classroom context remains relatively scarce, as highlighted by López-Carril et al. in 2022.

Furthermore, according to Tight (2020), student engagement has been among the most popular topics in higher education over the last four decades. Existing literature has

proved the relationship between engagement, self-regulation, collaboration, and self-efficacy independently. Pintrich (2000) and Zimmerman (2000a) viewed academic self-regulation as students' proactive-ness in learning by managing themselves (in terms of cognitive process, attention, behavior, and emotion) and the environment with the purpose of achieving their goals. This view by Pintrich and Zimmerman is based on social cognitive theory (Puustinen & Pulkkinen, 2001). In social cognitive theory, human cognition is a function of behavioral and environmental factors, which means people can exert self-control instead of merely reacting to external environment or inner impulses. This implies that the impact of learning environment on students' learning behaviours is mediated by their self-possessed attributes such as goals, perception, knowledge, skills.

Recognizing this enduring importance of student engagement in higher education, the primary objective of this paper is to delve into the intricate process by which LinkedIn can effectively promote students' engagement and active participation within the academic sphere. In other words, this study aims to address the following research question: *How do students' self-regulated behaviours, self-efficacy and academic motivation promote student engagement with the application of LinkedIn online courses in higher education?*

## **2. Literature review and hypothesis development**

Studies by Miao and Ma (2022) and Commissiong (2020) confirmed the positive impact of self-regulation on learning engagement in online setting. Pintrich (2000) and Zimmerman (2000a) viewed academic self-regulation as students' proactive-ness in learning by managing themselves (in terms of cognitive process, attention, behavior, and emotion) and the environment with the purpose of achieving their goals. Self-regulation is even more important in online courses which require a high level of self-management. When students exhibit self-regulated behaviors, they tend to perform actions conducive to their studies such as engaging and interacting activities. Thus, the first hypothesis in this paper proposes that:

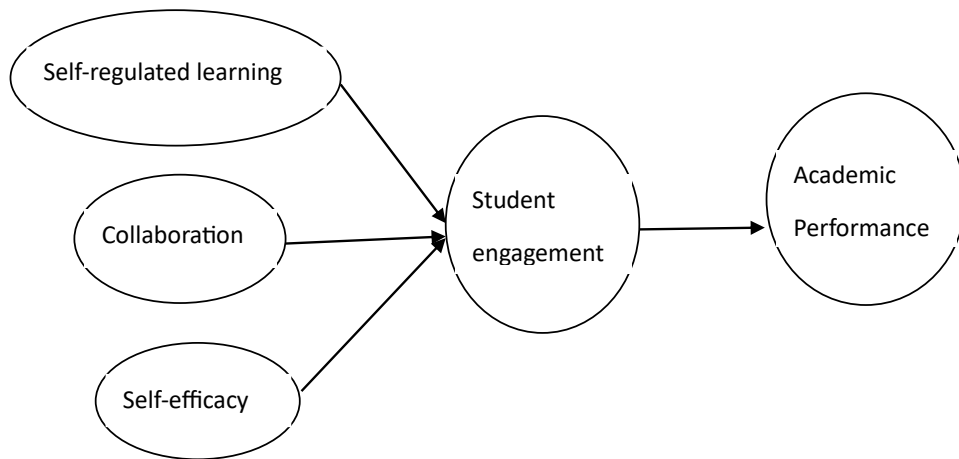
*H1. Self-regulation is positively associated with learning engagement.*

In addition to self-regulation, online engagement is said to be predicted by individual self-efficacy (Stan et al., 2022; Wolter et al. 2023). The term “self-efficacy” can be understood as one’s abilities to organize and perform actions necessary for goal attainment (Bandura, 1997). Especially in ambiguous learning environment such as online, those with high level of self-efficacy leads to higher learning engagement (Koob et al., 2021). Furthermore, considering the important role of self-efficacy in self-regulated learning, it is likely that self-efficacy also exerts an impact on engagement. The self-regulated learning (SRL) model by Zimmerman (2000b) demonstrated the important role of self-efficacy in self-regulated learning. In this model, besides goal setting and strategic planning, students’ self-efficacy is in the first phase of SRL process. Self-efficacy is accordingly a key component shaping SRL process under the social cognitive perspective. The second hypothesis is proposed as followed:

*H2. Online self-efficacy is positively associated with learning engagement.*

Moreover, collaboration in classrooms also plays a key role in learning engagement. Online collaborative learning strategy is found to enhance students’ engagement and performance (Gaad, 2022). The collaborative approach creates a congenial learning atmosphere where students appreciate opportunities to work with their classmates, increases students’ satisfaction, improves critical thinking, and helps students to gain better understanding of knowledge (Bharucha, 2017). The study on teaching and learning methodologies by Griffin & Howard (2017) showed that collaborative teaching strategies promote interaction and social support within the class. Students stated that they were both cognitively and behaviorally engaged in collaborative learning setting (Xu et al., 2023). Therefore, the next hypothesis is proposed as follow:

*H3. Collaboration is positively associated with learning engagement.*



**Figure 1. Research Model.**

A student's performance is said to be associated with factors of cognition, behaviors, and affections (Deci & Ryan, 2000; Vallerand et al., 1992; Vallerand et al., 2008). The feelings that students experience and behaviors that they exhibit have a direct impact on their performance. A successful student is a motivated one who can achieve both social and academic goals (Pintrich & De Groot, 1990). Motivated students are more attentive, show more progress, increase efforts, and show better persistence than un-motivated ones (Zimmerman & Schunk, 2007). Similarly, engaged students are more likely to commit time and efforts to learning activities and more involved in the learning process. Existing literature shows that student engagement positively influences learning performance and outcomes (Ievgenia et al., 2021; Lu & Cutumisu, 2022; Tran & Aspiras, 2022; Zahriban Hesari et al., 2022). Therefore, the next hypotheses are:

*H4. Student engagement is positively associated with academic outcomes.*

Finally, meta-analysis of forty years of evidence shows that it is not absolute that digital technology fosters student engagement (Tamim et al., 2011) but rather, using appropriate tool is vital (Popenici, 2013) because "technology can amplify great teaching, but great technology cannot replace poor teaching" (OECD, 2015, p. 4). Nevertheless, existing literature found that application of digital technology in higher education can improve student self-regulation, self-efficacy and increase engagement (Bond et al., 2020; Koehler & Mishra, 2005; Salaber, 2014). Due to such contrasting viewpoints in literature, it is worth exploring whether LinkedIn courses improve learners' self-efficacy, self-regulated learning, collaboration with others, and have an

indirect impact on student engagement and performance. Therefore, it is expected that the relationship between learners' collaboration, self-efficacy, self-regulated learning and student engagement and performance will be improved after the students participate in LinkedIn online classes as proposed in the following hypothesis.

*H5. Participation in LinkedIn online classes will result in a significant increase in student engagement compared to their engagement levels before participating.*

### **3. Research Design and Methodology**

#### **3.1 Research Design**

To answer the research question, we design research setting in which we can compare the student engagement level before and after participating in LinkedIn online classes. This aims to find out whether social media platforms such as LinkedIn can contribute to promoting student engagement. In addition, with this setting we investigate the psychological process of students with the purpose of understanding what factors (academic motivation, self-regulated learning and self-efficient) affect student engagement, then from student engagement to their academic performance.

#### **3.2. Development of measurement instruments**

A questionnaire instrument has been adapted from prior research. Particularly, self-regulated learning construct is based on the study of Bandura (1997) and Bandura et al. (2009). For self-efficacy, we apply the scale of Schwarzer & Jerusalem (1995). To measure students' collaboration, the Self-assessed Collaboration Skills (SACS) is used, which consists of three dimensions: information sharing, learning, and team support (Hinyard et al., 2019). Finally, we use the measurement of Dixson M. D. (2015) for student engagement.

All items were measured with a seven-point Likert scale (from 1 = totally disagree to 7 = totally agree). The questionnaire will be first evaluated through a group discussion with five experts from the education department. Then, the questionnaire will also be evaluated with students both offline (13 users) and online (21 users) using an interactive format. This will be followed by an offline pre-test (5 respondents) and an online pre-test (3 respondents) to administer the questionnaire in a non-interactive

format. All three evaluations and pre-tests will be aimed at improving the questionnaire regarding the measurement scales and the question wordings.

### **3.3. Data collection**

The respondents are all postgraduate students studying a 2-year MBA Global course at a UK university. In the second year of the course all students complete a personal and professional skills module which requires them to complete 80 hours of self-directed study using LinkedIn Learning based on their chosen course specialism (marketing, HRM, supply chain, finance, or entrepreneurship).

The first survey will be conducted before the students start participating in LinkedIn Learning. The same group of students will participate in the second survey after they have completed 80 hours of self-directed study using LinkedIn Learning, using the same set of questionnaires.

In addition, we provide a clear set of instructions and how to respond to the questionnaire. To ensure the reliability of the data, respondents were further screened using a set of questions in the beginning of the second survey, ensuring that the respondent completed LinkedIn online classes. Mandatory items in the questionnaire were indicated as such and were asked in a set sequence.

### **3.4. Data analysis**

We will start with confirmatory factor analysis to assess the measurement model based on the two-step approach proposed by Anderson and Gerbing (1988). In the second step, we utilize structural equation modelling to test the research hypotheses and research model (below) using statistical software programs. Finally, we will use other statistical analysis techniques (e.g., paired t-tests or other appropriate statistical tests) to compare the levels of student engagement before and after their participation in online classes.

## **References**

- Aldadouh, T.Z., Nokelainen, P., Korhonen, V. (2020). Technology and social media usage in higher education: the influence of individual innovativeness, *Sage Open*, 10 (1): 1–20.
- Al-Qaysi, N., Granić, A., Al-Emran, M., Ramayah, T., Garces, E., Daim, T.U. (2023). Social media adoption in education: A systematic review of disciplines, applications, and influential factors. *Technology in Society*, 73, 102-249.

- Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S.-L. (2009). Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education*, 12(1), 1–6.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman
- Bharucha, J. P. (2017). Building student engagement through collaborative practice in business management education. *International Journal of Virtual and Personal Learning Environments*, 7(2).
- Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *International Journal of Educational Technology in Higher Education*, 17 (1), p. 2
- Commission, M.A. (2020). *Student Engagement, Self-Regulation, Satisfaction, and Success in Online Learning Environments*. Dissertations. Walden University. Available online: <https://scholarworks.waldenu.edu/cgi/viewcontent.cgi?article=9794&context=dissertations>
- Deci, E.L., & Ryan, R.M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Deci, E.L., & Ryan, R.M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11, pp. 227–268
- Dixson M. D. (2015). Measuring student engagement in the online course: the Online Student Engagement Scale (OSE). *Online Learning*, 19(4).
- Gaad, A.L.V. (2022). The Effects of Online Collaborative Learning (OCL) on Student Achievement and Engagement. *IAFOR Journal of Education: Studies in Education*, 10(3).
- Griffin, C. P., & Howard, S. (2017). Restructuring the College Classroom: A Critical Reflection on the Use of Collaborative Strategies to Target Student Engagement in Higher Education. *Psychology Learning & Teaching*, 16(3), 375-392.
- Healy, M., Cochrane, S., Grant, P., & Basson, M. (2023). LinkedIn as a pedagogical tool for careers and employability learning: a scoping review of the literature. *Education + Training*, 65(1), pp. 106-125.
- Hinyard, L., Toomey, E., Eliot, K., & Breitbach, A. (2019). Student perceptions of collaboration skills in an interprofessional context: Development and initial validation of the self-assessed collaboration skills instrument. *Evaluation & the Health Professions*, 42(4), 450-472.
- Ievgenia, K., Bogdan, G., & Hannan, X. (2021). The Relationship Between Student Engagement and Academic Performance in Online Education. In 2021 5th International Conference on E-Society, E-Education and E-Technology (ICSET 2021). Association for Computing Machinery, New York, NY, USA, 97–101.
- Koehler, M., & Mishra, P. (2005). What happens when teachers design educational technology? The development of technological pedagogical content knowledge. *Journal of Educational Computing Research*, 32(2), 131–152.
- Koob, C., Schröpfer, K., Coenen, M., Kus, S., and Schmidt, N. (2021). Factors influencing study engagement during the COVID-19 pandemic: a cross-sectional study among health and social professions students. *PLoS One*, 16: e0255191.

- López-Carril, S., Alguacil, M., Anagnostopoulos, C. (2022). LinkedIn in sport management education: Developing the students' professional profile boosting the teaching-learning process. *The International Journal of Management Education*. 20(1): 100611.
- Lu, C., & Cutumisu, M. (2022). Online engagement and performance on formative assessments mediate the relationship between attendance and course performance. *International Journal of Educational Technology in Higher Education*, 19(2).
- Miao, J., & Ma, L. (2022). Students' online interaction, self-regulation, and learning engagement in higher education: The importance of social presence to online learning. *Frontiers in Psychology*. 13.
- OECD (2015). *Students, computers and learning*. PISA: OECD Publishing Retrieved from [http://www.oecd-ilibrary.org/education/students-computers-and-learning\\_9789264239555-en](http://www.oecd-ilibrary.org/education/students-computers-and-learning_9789264239555-en)
- Pintrich, P., Brown, Donald R, Weinstein, Claire Ellen. (2012). *Student Motivation Cognition and Learning*. Hoboken: Routledge.
- Pintrich, P. R. (2000). "The role of goal orientation in self-regulated learning". In *Handbook of self-regulation*, Edited by: Boekaerts, M., Pintrich, P. R. and Zeidner, M. 451–502. San Diego, CA: Academic.
- Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of educational psychology*, 82(1), 33.
- Popenici, S. (2013). *Towards a new vision for university governance, pedagogies and student engagement*. In E. Dunne, & D. Owen (Eds.), *The student engagement handbook: Practice in higher education*, (1st ed., pp. 23–42). Bingley: Emerald.
- Puustinen, M., & Pulkkinen, L. (2001), Models of self-regulated learning: A review, *Scandinavian Journal of Educational Research*, 45 (3), pp. 269-286
- Sabah, N.M. (2022). The impact of social media-based collaborative learning environments on students' use outcomes in higher education. *International Journal of Human-Computer Interaction*, 39 (3), 667-689.
- Salaber, J. (2014). Facilitating student engagement and collaboration in a large postgraduate course using wiki-based activities. *International Journal of Management in Education*, 12 (2) (2014), pp. 115-126
- Stan, M.M., Topală, I.R., Necşoi, D.V., Cazan, Ana-Maria. (2022). Predictors of Learning Engagement in the Context of Online Learning During the COVID-19 Pandemic. *Frontiers in Psychology*. 13.
- Schwarzer, R., & Jerusalem, M. (1995). *Generalized Self-Efficacy scale*. In J. Weinman, S. Wright, & M. Johnston, *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: NFER-NELSON
- Tamim, R.M., Bernard, R.M., Borokhovski, E., Abrami, P.C., Schmid, R.F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research*, 81 (1) (2011), pp. 4-28
- Tan, P.J.B., & Hsu, M. H. (2017). Developing a system for English evaluation and teaching devices. *International Conference on Applied System Innovation (ICASI)*, pp. 938–941.
- Tran, T. V., & Aspiras, O. (2022). College students' engagement and academic outcomes in online



- learning during the COVID-19 pandemic. *Modern Psychological Studies*, 28(1), Article 3. Available at: <https://scholar.utc.edu/mps/vol28/iss1/3>
- Zimmerman, B. J. (2000a). "Attaining self-regulation: A social-cognitive perspective". In *Handbook of self-regulation*, Edited by: Boekaerts, M., Pintrich, P. R. and Zeidner, M. 13–39. San Diego: Academic Press.
- Zimmerman, B.J. (2000b). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25(1), 82–91.
- Wolters, C.A., Iaconelli, R., Peri, J., Hensley, L.C., Kim, M. (2023). Improving self-regulated learning and academic engagement: Evaluating a college learning to learn course, Learning and *Individual Differences*. 103: 102282.
- Vallerand, R.J., L.G. Pelletier, M.R. Blais, N.M. Briere, C. Senecal and E.F. Vallieres, 1992. The academic motivation scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement*, 52(4): 1003-1017.
- Vallerand, R.J., L.G. Pelletier and R. Koestner, 2008. Reflections on self-determination theory. *Canadian Psychology*, 49(3): 257-262.
- Xu, B., Stephens, J.M. & Lee, K. (2023). Assessing Student Engagement in Collaborative Learning: Development and Validation of New Measure in China. *Asia-Pacific Education Researcher*.
- Zahriban Hesari, A., Ekrami, M., & Sarmadi, M. (2022). Analyzing the relationship between student engagement in Online Learning and Achievement with mediator variables of Readiness and Learning Climate during the Covid-19 epidemic. *Future of Medical Education Journal*, 12(4), 26-32.

# Impact of WhatsApp Groups on Teaching and Learning in Higher Education: A Review

## Abstract

*The impact of WhatsApp groups on teaching and learning in higher education institutions is growing, as both instructors and students see its potential to improve the learning experience. The study adopted a systematic literature review following the PRISMA guidelines and 26 articles were analyzed. The results showed a rise in the usage of WhatsApp during COVID-19 and an expansion of the usage in higher education in teaching and learning. Furthermore, WhatsApp groups have benefited students and lecturers in teaching and learning in different academic fields and a positive impact was seen. However, some challenges were realized due to the introduction of WhatsApp groups in teaching and learning. This study has brought a new voice to the body of literature on digital technologies.*

**Keywords:** WhatsApp Group, Teaching and Learning, Higher Education, Social Media, ICT, Online Technology

## 1 Introduction and Background

Social network applications have played an important role in teaching and learning for a long time transforming the higher education landscape and changing established approaches to teaching and learning. In this digital age, one noticeable phenomenon is the growing usage of messaging networks such as WhatsApp as educational tools within higher education institutions. The platform allows people with shared interests to collaborate irrespective of geographical distances and times when they form groups (Koole, 2009; Rambe & Bere, 2013). WhatsApp, which began as a personal messaging service, has rapidly found its way into the academic world, providing a versatile platform for communication, collaboration, and participation. The impact of WhatsApp groups on teaching and learning in higher education institutions is becoming increasingly important, as both educators and students see its potential to improve the learning experience. WhatsApp was chosen as the social media platform for this study because it is a popular platform in most African universities. Furthermore, the user-friendliness (Sari & Putri, 2019) and ease of use (Widodo,

2019) of WhatsApp are also a positive factor for learners who lack technical knowledge.

Higher education institutions are increasingly resorting to WhatsApp groups to facilitate communication, collaboration, and information exchange among students and educators in an era of quickly changing digital technology (Udenze & Oshionebo, 2020; Nyamupangedengu et al., 2023). While the impact of WhatsApp groups keeps growing in educational spaces, their recognition remains on the periphery of educational policy documents, which do not list their integral part in teaching and learning within higher education. This paper reviews the literature to assess the impact of WhatsApp groups on teaching and learning in higher education, given challenges such as high data costs and institution restrictions (Wang & Chen, 2009; Snoussi, 2019) associated with the Learner Management System (LMS). In some countries, such as Bulgaria and India, institutions integrate applications within their LMSs (Mahapatra et al., 2016; Chikurteva et al., 2020). While technological integration has the potential to improve teaching and learning, it also raises serious concerns about the overall influence on pedagogical practices, student engagement, privacy and security, exclusion and academic achievement in higher education. To get well-informed about integration benefits, it is important to assess the impact that these applications, such as WhatsApp groups, bring to teaching and learning in higher education, considering both the benefits and drawbacks. The key research question for the study is:

*What is the impact of WhatsApp groups on teaching and learning in higher education?*

The remainder of the paper is structured as follows: Section 2 discusses the literature review on WhatsApp groups in teaching and learning, Section 3 describes the research methodology, Section 4 presents the findings, Section 5 discusses the findings, Section 6 concludes the paper and gives the contributions of the study, and Section 7 provides a recommendation.

## **2 Literature Review**

Research arguments are contextualised within other related studies by identifying gaps and acknowledging opposing and supporting viewpoints. We carried out a systematic literature review to achieve comprehensiveness (Okoli & Schabram, 2010) and position our understanding of the phenomenon of interest within prior studies. We discuss the literature review under the following sub-themes: understanding WhatsApp groups, digital tools in education, WhatsApp groups in education, and supporting the argument with a theoretical perspective. We used structural coherence (Locke & Golden-Biddle, 1997) to identify how the literature on WhatsApp groups is linked.

### **2.1 Digital Tools in Education**

According to (Bond et al., 2018), the prevalence of digital tools in education has witnessed significant growth in recent years, a trend projected to persist into the future, even impacting the workplace, as suggested by (Colbert et al., 2016). These digital tools have brought about a transformative shift in education by furnishing educators and learners with innovative avenues for content engagement, peer collaboration, and personalised learning experiences. Researchers (Jere et al., 2019; Rahaded et al., 2020) highlight that digital tools streamline communication and cooperation between students and teachers. For instance, platforms such as WhatsApp enable students to collaborate on projects and gain knowledge from one another, irrespective of geographical boundaries (Koole, 2009; Duncheon & Tierney, 2013; Mao et al., 2017). Nonetheless, the integration of digital tools into the educational landscape raises pertinent challenges on equity, privacy, and the imperative for comprehensive teacher training. In light of the evolving educational milieu, it becomes imperative for educators and institutions to judiciously harness digital tools to optimise their advantages while concurrently addressing the attendant concerns, thereby ensuring that technology acts as a catalyst, rather than an impediment, in enriching educational experiences. The substantial engagement facilitated by digital tools on platforms like WhatsApp enhances personalised learning prospects, given the extensive participation of peers in collaborative endeavours.

## **2.2 WhatsApp Groups in Education**

A WhatsApp group is a chat space within WhatsApp that enables multiple users to communicate, share information, and collaborate. A WhatsApp group is created by a person who takes on the role of group administrator by default. Depending on the group's purpose and the admin's preferences for joining participants, WhatsApp groups can have a few to hundreds of members. New members can join a group without an invitation from the admin by clicking invite links generated by group admins, but, of course, this has the potential to attract people with no shared interests, although it is useful for large groups or public communities.

One of the most common uses of WhatsApp groups is the sharing of classroom communication, such as assignment submission reminders, important class updates, and broadcasting class achievements and important dates. According to (Rahmadi, 2020) study, students acknowledge their potential passively through WhatsApp groups. Further, WhatsApp groups have been found to increase student motivation (Susilawati & Supriyatno, 2020) and contribute to active learning and student engagement (Dahdal, 2020; Nugroho, 2022). In addition, WhatsApp groups are one of the fastest ways to share resources among students in class and provide immediate feedback (Spencer & Hiltz, 2003). Using WhatsApp groups, teachers and students can share educational materials such as PDFs, links to online articles, videos, and more. However, it should be noted that WhatsApp groups have some challenges which educators need to deal with to get the best out of these platforms.

Given the chaotic nature of WhatsApp groups, there is a need for proper control to ensure the privacy of participants. It is crucial to consider alternatives for students without WhatsApp access or who prefer not to use it; otherwise, the platform becomes a stumbling block to learning. Dualising communication, though it is time consuming potentially helps to deal with the aforementioned challenge. WhatsApp's capabilities in voice and text assist immensely in catering for students with differential physical challenges, which in essence helps students understand the information quickly. For WhatsApp groups to be used effectively in education, educators and students should be aware of both their advantages and limitations. We conclude that the discussed literature suffers from progressive incoherence. We then

use connectivism as the underpinning theory to position our argument within existing frames.

### **2.3 Theoretical Framework - Connectivism**

This study examines how information is disseminated within WhatsApp groups using the connectivism learning theory by George Siemens. As part of the theory adopted for this study, the following concepts are highlighted: diversity of opinions, networks, chaos, openness, and self-organizing (Siemens, 2005; Goldie, 2016). The concept of connectivism has been proposed as a learning theory for digital natives who rely on networks, with active nodes considered to be the most reliable connections within WhatsApp groups (Dzvapatsva, 2020). A WhatsApp application offers a fast, easy, and cheap way to communicate, especially if you're in a poor area (van den Berg & Mudau, 2022). The multiplicity of nodes (Downes, 2007) facilitates communication and information sharing among the group members. As soon as a student identifies an important node in a WhatsApp group, they start self-organizing. While connectivism has been criticized for its lackadaisical assumption that identification of key nodes is easy (Dunaway, 2011), we believe students themselves can quickly identify sources of reliable information within the WhatsApp group itself. In higher education, connectivism can inform teaching and learning methods, instructional design, and curriculum development.

Connectivism was considered appropriate in this study because it offers a theoretical framework that is consistent with the features of WhatsApp groups as technologically mediated and networked learning environments. Connectivism's influence on learning theories and instructional methods, along with its applicability in the current digital era, make it a suitable theoretical framework for this research. In light of the digital age, the theory's emphasis on the value of networks and connections in learning is extremely pertinent. Connectivism's application as a theoretical framework within the context of a methodical literature review allows for a thorough analysis of WhatsApp groups' effects on teaching and learning in higher education. In this inquiry, the researchers searched the examined papers for connectivist characteristics related to the subject matter. The results were categorized into themes.

### 3 Methodology

The research adopted a systematic literature review to assess the impact of WhatsApp groups on teaching and learning at higher education institutions. A systematic literature review was adopted because it allows the researchers to comprehensively survey and synthesize existing research on the subject. The purpose of the systematic literature review approach was to offer evidence-based insights into the usefulness of WhatsApp groups as a teaching and learning aid. Systematic literature review is critical in the context of higher education, where educational strategies must be informed by evidence-based approaches. By using this strategy, the researchers are assured of including a wide range of viewpoints, conclusions, and study techniques about the effects of WhatsApp groups in educational settings. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standard was adopted as a guide for locating previously published works (Moher et al., 2009). The researchers followed the PRISMA guidelines to conduct the research due to its rigour, comprehensiveness, and reproducibility (Okoli & Schabram, 2010). The following inclusion criteria were used:

- WhatsApp groups
- Teaching and learning
- Higher education

On the other hand, articles were excluded due to the following reasons:

- Not focusing on WhatsApp groups
- Not focusing on teaching and learning
- Not focusing on higher education
- Not written in English
- Repeating articles
- Full articles are not available

The researchers conducted a literature search from two databases namely Scopus and IEEE Xplore, as shown in Figure 1. The researchers used the following search strings in the databases:

*Scopus: WhatsApp groups in teaching and learning in higher education*

*IEEE Xplore: WhatsApp groups in teaching and learning*

The researchers used different search strings because the search string used in the Scopus database retrieved only 4 articles in the IEEE Xplore. In trying to expand the research, the researchers decided to open up the search string, hoping to retrieve more articles. The study retrieved a total of 51 articles from both databases. The articles

were screened using the inclusion and exclusion criteria for the study, and 25 articles were excluded. The researchers analyzed 26 articles that met the inclusion and exclusion criteria. The whole process of searching is illustrated in Figure 1.

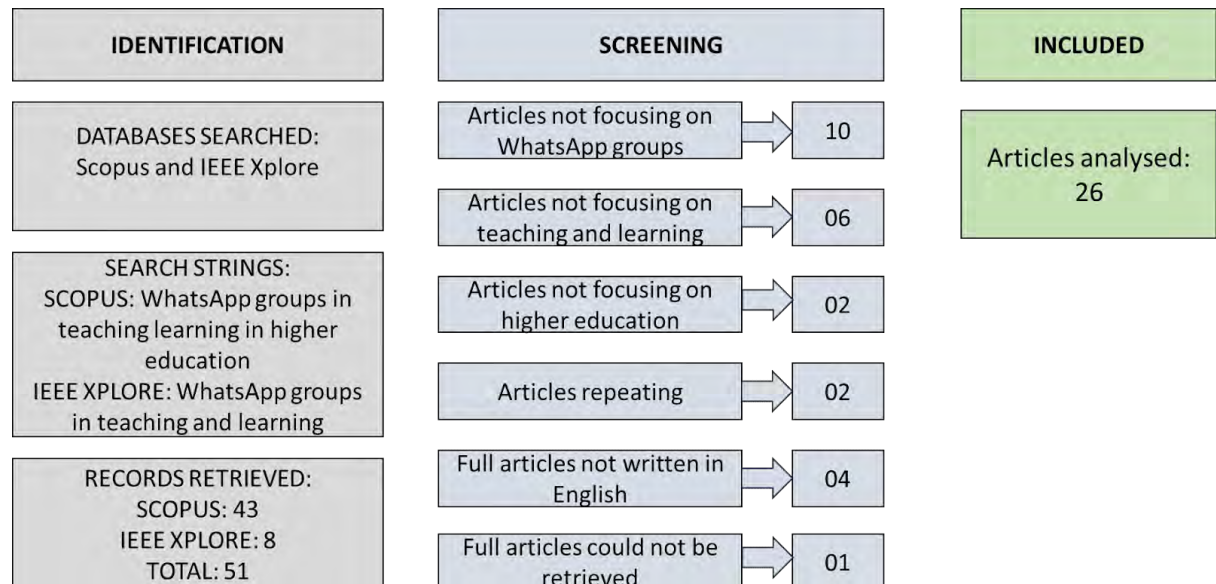


Figure 1. Diagrammatic representation of the methodology

## 4 Results

The results are presented in the form of themes. Table 1 shows the documents that were analyzed by year of publication.

Year	Frequency	Frequency (%)
2023	2	7,69%
2022	4	15,38%
2021	5	19,23%
2020	8	30,77%
2019	1	3,85%
2018	3	11,54%
2017	1	3,85%
2016	2	7,69%
Total	26	100%

Table 1. Documents by Year

Table 1 shows that the majority of the papers (65.38%) dealing with WhatsApp groups on teaching and learning with an emphasis on higher education were published



between 2020 and 2022. Only 26.93% of the articles were published before 2020. It can be seen from the table that the discussion about WhatsApp groups in teaching and learning within higher education was sparked more in 2020 when there was a sharp increase. This period coincided with the COVID-19 pandemic.

The results from the analyzed articles were grouped into themes, namely, adaptability and lifelong learning, networked learning and social learning, technology-mediated learning, benefits of WhatsApp groups in teaching and learning and challenges of WhatsApp groups in teaching and learning. The results are discussed under these themes.

#### **4.1 Adaptability and Lifelong Learning**

Literature shows that WhatsApp groups were used before the introduction of COVID-19 for teaching and learning. However, articles analyzed show that the use of WhatsApp groups has been expanded since the introduction of COVID-19. Before COVID-19, WhatsApp groups were largely utilized for administrative or communication purposes, and they were mostly used by students without the assistance of lecturers (Yu & Motlhabane, 2022). During COVID-19, WhatsApp groups were used to support students when physical classes were suspended in a way to curb the spread of the virus. With the coming of COVID-19, WhatsApp groups have been used to support and supplement academic and professional activities at universities since the outbreak of COVID-19 (van den Berg & Mudau, 2022; le Roux & Parry, 2022). During the era of COVID-19, lecturers shared lecture recordings in various formats (such as audio and videos) in WhatsApp groups for students to access content at any time and from any location, lowering their anxiety caused by feeling alone and left behind (Tunjera & Chigona, 2022).

WhatsApp groups have provided students with a platform to acquire knowledge, actively talk and learn, score higher, and retain more than didactic lectures (Klein et al., 2018). Outside of school hours, the educational intervention of WhatsApp for the course has been seen to improve students' knowledge of the subject (So, 2016). Students saw the potential of the WhatsApp groups for ubiquitous learning and had a positive attitude about this app as a teaching and learning platform (Rahmadi, 2020). Using WhatsApp groups as a handy teaching tool has been shown in studies to boost

academic performance (Alsharif et al., 2020). This made WhatsApp groups the most popular online learning medium throughout the COVID-19 pandemic because they were the easiest to use and access (Pramana et al., 2021).

#### **4.2 Networked Learning and Social Learning**

According to the articles reviewed, WhatsApp groups were used to support students from different disciplines, such as health sciences, pathology, pre-service teachers, law, physics, and ICT as they were learning in their respective groups (Jannat et al., 2022; Tunjera & Chigona, 2022; Yadav et al., 2021; Alsharif et al., 2020; Grover et al., 2020; Klein et al., 2018; Basitere & Ivala, 2017). WhatsApp groups were used to facilitate teaching and learning at the undergraduate and postgraduate levels (Lee et al., 2023; van den Berg & Mudau, 2022; Al-Omary et al., 2016).

Through WhatsApp groups, students were able to engage with each other at their leisure and seek advice from their mentors without hesitation or opposition (Klein et al., 2018). WhatsApp groups assisted lecturers in creating a welcoming environment that assists students outside of the classroom in completing assignments and studying course content (Al-Omary et al., 2016). As a result, WhatsApp groups boost communication in teaching and learning, improve student-to-lecturer contact, student-to-student interaction, student-to-lecturer closeness, and student-to-student intimacy (Robles et al., 2019; Ujakpa et al., 2018). Students exposed to WhatsApp groups were quite satisfied with the utilization of the WhatsApp chat group (Robles et al., 2019). Learning through WhatsApp groups improved student interest in courses as the students received resources in various formats that could cater to different learning styles (Ujakpa et al., 2018). Studies reported positive outcomes in the use of WhatsApp groups to support teaching and learning in courses such as law and ICT (So, 2016). Furthermore, there was a favourable attitude toward the use of WhatsApp groups for project learning in higher education and positive effects such as boosting competency and promoting collaborative learning (Berewot & Fibra, 2020).

#### **4.3 Technology-Mediated Learning**

WhatsApp groups were the most used online learning platforms by lecturers at universities during COVID-19 compared to online platforms such as Microsoft Teams, Zoom, and Google Class (Mursyidin et al., 2021; Figueras-Maz et al., 2021).

Lecturers and students preferred WhatsApp groups because of immediate, speedy, simple, and low-cost communication and interaction, resulting in a sense of belonging for most of the students (van den Berg & Mudau, 2022). Interestingly, students preferred the use of WhatsApp groups in education even before COVID-19 for generating better interaction among students and contributing to the collective building of knowledge (Martins et al., 2018). While these groups are predominantly associated with chaos, they are self-organising with students' ability to identify those who are always providing credible information (Goldie, 2016). Furthermore, WhatsApp groups are an effective communication medium in higher education for students and instructors to maintain communities of practice (Nuuyoma et al., 2020).

#### **4.4 Benefits of WhatsApp Group in Teaching and Learning**

Lecturers and students appreciate the benefits that come with WhatsApp groups in teaching and learning. WhatsApp groups enabled the exchange of a wider range of resources (e.g., texts, photographs, videos, and voice notes) to better fit the varied learning styles of students (Ramkissoon et al., 2020). Besides sharing resources, students performed academic-related activities on WhatsApp, such as group discussions, group studies, and informing educational agendas (Jannat et al., 2022). WhatsApp groups enable the facilitation of hybrid learning and provide a non-restrictive environment by facilitating meaningful interactions with instructors anywhere and at any time (Ramkissoon et al., 2020). The majority of students preferred WhatsApp because it allowed them to access content from anywhere and at any time (Mursyidin et al., 2021). As a result, WhatsApp groups enabled learning to continue even after class with or without lecturers, thereby extending the bounds of a classroom (Figueras-Maz et al., 2021). Moreover, WhatsApp is preferred because of its capabilities for interactivity, usability, respect for privacy, collaboration, and rapidity of feedback (Ramkissoon et al., 2020; Lee et al., 2023). With WhatsApp groups, students in teams benefited a lot, as they could collaborate easily. Teams could make use of group chats, and emojis, read receipts, and quote messages to increase collaboration (Lee et al., 2023). Students working in teams in WhatsApp groups were seen achieving their goals as they did their best to accomplish their tasks (Kurni & Saritha, 2021). Students were seen to solve problems faster when working in WhatsApp groups compared to when they worked in other platforms such as forums and learning management systems (Figueras-Maz et al., 2021). Learning through

WhatsApp groups has a favourable learning effect on students, and the durability of this effect was seen to be higher when the teaching material is delivered in short videos compared to virtual courses hosted through platforms like Skyroom (Jannat et al., 2022). WhatsApp groups also support students in developing countries characterized by challenges such as network connectivity issues, a lack of funds to buy data, and a lack of access to infrastructure (Tunjera & Chigona, 2022). WhatsApp groups brought inclusivity as most students were able to access WhatsApp since it is light on data (Tunjera & Chigona, 2022). In summary, WhatsApp groups were preferred over any other online technology in terms of their ubiquity of usage, efficiency in collaboration, accessibility, sense of presence, and effectiveness as communication tools (Lee et al., 2023; Klein et al., 2018).

#### **4.5 Challenges of WhatsApp Groups in Teaching and Learning**

Besides the benefits that come with a WhatsApp group, some challenges were encountered that affected teaching and learning. Since WhatsApp groups are mainly for socializing, students can be distracted when they do non-academic activities, hindering them from achieving their academic goals (Yu & Motlhabane, 2022). Furthermore, some messages passed in WhatsApp groups may have disrespectful tones; friends may send too many messages in the group that may distract students from academic activities; and also, irrelevant messages can be sent within the group that could divert the attention of students (Yu & Motlhabane, 2022; van den Berg & Mudau, 2022; Alsharif et al., 2020). It was discovered that WhatsApp groups require students who are ready to learn actively, collaboratively, and independently, as there might not be someone monitoring students' engagements (Rahmadi, 2020). Therefore, there is a need to educate students about the techniques that can be used to offset its negative impacts on human behaviour, such as disruptions, addiction, and a lack of responses (Nuuyoma et al., 2020).

## **5 Discussion**

Our study aimed to answer the following research question:

*What is the impact of WhatsApp groups on teaching and learning in higher education?*

Significant data from our findings progressively highlighted the positive impact of WhatsApp groups on learning in higher education, which in essence mirrored our literature. The ability of WhatsApp groups for teaching purposes is not explicitly reflected in our data. Although lecturers or facilitators use it for sharing lecture material, the teaching process involves students actively engaging with content, taking part in assessments, and providing feedback or remediation to the process. According to connectivism, learning happens in networks, and WhatsApp groups provide an excellent example of this since they allow users to share resources, participate in collaborative learning, and add to a group's body of knowledge. The immediacy of feedback (Spencer & Hiltz, 2003), also noted in literature is one of the greatest plus for WhatsApp groups, a thing that most platforms might not provide because they are not used as a student playground. Students are always on their mobile phones, and any message that comes through they are instinctively forced to look at it.

Connectivism recognizes how technology shapes learning experiences. WhatsApp's technology-mediated features support the idea by highlighting how people obtain and share information via digital technologies contributing to social presence. Since WhatsApp provides video and voice, the platform has a high social presence to the extent that students feel like they are in a face-to-face class. Voice and text are a great plus for the medium as they complement each other when used to explain concepts in WhatsApp groups.

Furthermore, connectivism emphasizes the need for learners to adapt to evolving information environments. WhatsApp groups facilitate adaptability by providing a responsive and adaptable learning environment that is consistent with the theory's emphasis on lifelong learning.

WhatsApp groups are of great importance when comes to content sharing and issuing important key dates for group members. Since the platform can be used for group tasks, completion is easier as communication is instantly received on both ends and participants can respond to a particular message by tagging it. The more they text and send voice notes, the more they learn to appreciate each other resulting in improved relations. The findings point to positive achievements for students participating in

WhatsApp groups owing to improved motivation. This has to be taken with caution as there could be many other factors contributing, such as the quality of the content.

Similar to our theoretical framework, WhatsApp groups can also be viewed as chaotic. These groups can be very useful for teaching and learning, but there is still a need to control the chaos which may negatively affect learning. This has been highlighted in our literature and appeared in our findings. Students on WhatsApp groups can identify important nodes primarily due to their face-to-face interaction experience, but this did not come out explicitly in our study. Perhaps this is because the reviewed articles did not use connectivism theory, which we believe would have resulted in similar results. It is crucial to note that WhatsApp groups are beneficial in teaching and learning if associated challenges are dealt with carefully. Our methodology was appropriate for getting conclusive data on the impact of WhatsApp groups on teaching and learning at a very broad level allowing us to get a conclusive voice on their impact.

However, the outcomes of learning, such as skill and competence are not well revealed in the study.

## **6 Conclusions and Contributions**

The objective of the study was to understand the impact of WhatsApp groups on teaching and learning in higher education. Despite the brevity of the progressive intertextual field developed in the reviewed literature for this manuscript, it exhibits rhetorical strategies that call for the construction of both cumulative progress and consensus regarding WhatsApp groups' effects on teaching and learning in higher education. The methodology applied allowed this study to get a broader perspective on the impact of WhatsApp groups on teaching and learning. The findings suggest that WhatsApp groups can contribute positively to teaching and learning in higher education.

According to (Locke & Golden-Biddle, 1997), reinterpreting existing work to show underlying consensus is the third synthesized coherence practice crucial for the contribution of a study which this research has done. A new voice has been added to the body of literature on digital technologies through this study. We believe that

integrating WhatsApp technology into higher education policy documents will be easier with more voices in WhatsApp groups.

Due to the study's cross-sectional nature, the results need to be taken with caution.

## 7 Recommendation

The study recommends increased use of WhatsApp groups for teaching and learning in higher education. It could be more beneficial if WhatsApp groups were integrated into institutions' Learner Management Systems.

## References

- Al-Omary, A., El-Medany, W. M., & Isa, K. J. E. (2016). The impact of SNS in higher education: A case study of using WhatsApp in the University of Bahrain. *2015 Fifth International Conference on E-Learning (Econf)*, 296–300. <https://doi.org/10.1109/ECONF.2015.72>
- Alsharif, A. T., Alsharif, B., Alsharif, L., Althagafi, N., Natto, Z. S., & Kassim, S. (2020). Effectiveness of WhatsApp as a Part of a Hybrid Learning Environment: An Opportunity for Post-COVID-19 Pandemic Pedagogy. *Journal of Contemporary Dental Practice*, 21(12), 1331–1336. <https://doi.org/10.5005/jp-journals-10024-2978>
- Basitere, M., & Ivala, E. N. (2017). An exploration of students' experiences of Blended Learning in a physics course at a university of technology. *Journal of Social Development in Africa*, 32(1), 23–43.
- Berewot, J., & Fibra, A. (2020). Whatsapp Messenger among Higher Education Students: A Study on Whatsapp Messenger Usage for Project Learning. *Journal of Physics: Conference Series*, 1477(4). <https://doi.org/10.1088/1742-6596/1477/4/042018>
- Bond, M., Marín, V. I., Dolch, C., Bedenlier, S., & Zawacki-Richter, O. (2018). Digital transformation in German higher education: student and teacher perceptions and usage of digital media. *International Journal of Educational Technology in Higher Education*, 15(1), 1–20. <https://doi.org/10.1186/s41239-018-0130-1>
- Chikurteva, A., Spasova, N., & Chikurtev, D. (2020). E-learning: Technologies, application and challenges. *2020 XXIX International Scientific Conference Electronics (ET)*, 1–4. <https://doi.org/10.1109/ET50336.2020.9238176>
- Colbert, A., Yee, N., & George, G. (2016). The digital workforce and the workplace of the future. *Academy of Management Journal*, 59(3), 731–739. <https://doi.org/10.5465/amj.2016.4003>
- Dahdal, S. (2020). Using the WhatsApp Social Media Application for Active Learning. *Journal of Educational Technology Systems*, 49(2), 239–249. <https://doi.org/10.1177/0047239520928307>
- Downes, S. (2007). Learning networks in practice. *Emerging Technologies for Learning*, 2(4), 20.
- Dunaway, M. K. (2011). Connectivism: Learning theory and pedagogical practice for networked information landscapes. *Reference Services Review*, 39(4), 675–685. <https://doi.org/10.1108/00907321111186686>

- Duncheon, J. C., & Tierney, W. G. (2013). Changing Conceptions of Time: Implications for Educational Research and Practice. In *Review of Educational Research* (Vol. 83, Issue 2). <https://doi.org/10.3102/0034654313478492>
- Dzvapatsva, G. P. (2020). *Contextualisation of Instructional Time Utilising Mobile Social Networks for Learning Efficiency: A Participatory Action Research Study for Technical Vocational Education and Training Learners in South Africa*.
- Figueras-Maz, M., Grandío-Pérez, M. D. M., & Mateus, J. C. (2021). Students' perceptions on social media teaching tools in higher education settings. *Communication and Society*, 34(1), 15–28. <https://doi.org/10.15581/003.34.1.15-28>
- Goldie, J. G. S. (2016). Connectivism: A knowledge learning theory for the digital age? *Medical Teacher*, 38(10), 1064–1069. <http://elearning.surf.nl/e-learning/english/3793>
- Grover, S., Garg, B., & Sood, N. (2020). Introduction of case-based learning aided by WhatsApp messenger in pathology teaching for medical students. *Journal of Postgraduate Medicine*, 66(1), 17–22. [https://doi.org/10.4103/jpgm.JPGM\\_2\\_19](https://doi.org/10.4103/jpgm.JPGM_2_19)
- Jannat, F., Alipour, S., Noori, F., Ansari, S., Ashtab, T., Eskandari, A., Sani, S. B., Orouji, M., Goodarzi, D., & Amshaki, F. D. (2022). Comparison of the Effectiveness of Breast Cancer Education through Two Virtual Methods for Increasing Knowledge in Nurses. *Nurse Media Journal of Nursing*, 12(1), 100–110. <https://doi.org/10.14710/nmjn.v12i1.41780>
- Jere, N. R., Jona, W., & Lukose, J. M. (2019). Effectiveness of Using WhatsApp for Grade 12 Learners in Teaching Mathematics in South Africa. *2019 IST-Africa Week Conference, IST-Africa 2019*, 1–12. <https://doi.org/10.23919/ISTAfrICA.2019.8764822>
- Klein, A. Z., Junior, J. C. D. S. F., Barbosa, J. L. V., & Baldasso, L. (2018). The educational affordances of Mobile Instant Messaging (MIM): Results of Whatsapp® used in higher education. *International Journal of Distance Education Technologies*, 16(2), 51–64. <https://doi.org/10.4018/IJDET.2018040104>
- Koole, M. L. (2009). Use of mobile learning apps in workplace learning. In *Mobile learning: Transforming the delivery of education and training* (pp. 25–47).
- Kurni, M., & Saritha, K. (2021). Applying collaborative learning for enhancing the teaching-learning process in online learning through social media. *International Journal of Emerging Technologies in Learning*, 16(16), 251–259. <https://doi.org/10.3991/ijet.v16i16.23829>
- le Roux, D. B., & Parry, D. A. (2022). An exploratory investigation of the use and effects of academic instant messaging groups among university students. *Education and Information Technologies*, 27(1), 1055–1080. <https://doi.org/10.1007/s10639-021-10631-y>
- Lee, C. E., Chern, H. H., & Azmir, D. A. (2023). WhatsApp Use in a Higher Education Learning Environment: Perspective of Students of a Malaysian Private University on Academic Performance and Team Effectiveness. *Education Sciences*, 13(3). <https://doi.org/10.3390/educsci13030244>
- Locke, K., & Golden-Biddle, K. (1997). Constructing Opportunities for Contribution: Structuring Intertextual Coherence and “Problematizing” in Organizational Studies. *Academy of Management Journal*, 40(5), 1023–1062. <https://doi.org/10.5465/256926>
- Mahapatra, J., Srivastava, S., Yadav, K., Shrivastava, K., & Deshmukh, O. (2016). LMS weds WhatsApp: Bridging digital divide using MIMs. In *Proceedings of*



*the 13th International Web for All Conference (pp. 1-4)*

- Mao, Z., Jiang, Y., Min, G., Leng, S., Jin, X., & Yang, K. (2017). Mobile social networks: Design requirements, architecture, and state-of-the-art technology. *Computer Communications*, *100*, 1–19. <https://doi.org/10.1016/j.comcom.2016.11.006>
- Martins, E. R., Geraldes, W. B., Afonseca, U. R., & Gouveia, L. M. B. (2018). The Use of WhatsApp in Learning: An Experience in Higher Education. *2018 Proceedings*. <https://aisel.aisnet.org/capsi2018>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264–269. <https://doi.org/10.7326/L23-0065>
- Mursyidin, M., Parlindungan, F., & Rahmatillah, R. (2021). Challenges in Online Learning during Covid-19 Pandemic: Lessons Learned from Universities in Indonesia. *TESOL International Journal*, *110–124*. <http://www.tesol-international-journal.com>
- Nyamupangedengu, E., Kazeni, M., Mandikonza, C., Stephen, M., Tshesane, H., Ajayi, E., Abdulhamid, L., Makonye, J., & Moloi, M. (2023). Mathematics and science teacher educators experiences of using the WhatsApp platform as a tool for supporting teaching during the Covid-19 pandemic. *Journal of Education (University of KwaZulu-Natal)*, *(92)*, 115-135
- Nugroho, N. (2022). Students' Engagement in Online Learning Using Whatsapp Group During Covid-19 Pandemic. *Linguistics and ELT Journal*, *10*(1), 23. <https://doi.org/10.31764/leltj.v10i1.8387>
- Nuuyoma, V., Mhlope, N. J., & Chihururu, L. (2020). The use of WhatsApp as an educational communication tool in higher education: Experiences of nursing students in Kavango East, Namibia. *International Journal of Higher Education*, *9*(5), 105–114. <https://doi.org/10.5430/ijhe.v9n5p105>
- Okoli, C., & Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *Sprouts: Msterdam, The Netherlands*. <https://doi.org/10.2139/ssrn.1954824>
- Pramana, C., Susanti, R., Violinda, Q., Yoteni, F., Rusdiana, E., Prihanto, Y. J. N., & Purwahida, R. (2021). Virtual learning during the COVID-19 pandemic, a disruptive technology in higher education in Indonesia. *Okma and Arkiang, Fajeri and Purwahida, Rahmah and Haimah, Virtual Learning During The COVID-19 Pandemic, A Disruptive Technology In Higher Education In Indonesia (February 2, 2021)*.
- Rahaded, U., Puspitasari, E., & Hidayati, D. (2020). The Impact of Whatsapp Toward Uad Undergraduate Students' Behavior in Learning Process. *International Journal of Educational Management and Innovation*, *1*(1), 55. <https://doi.org/10.12928/ijemi.v1i1.1515>
- Rahmadi, I. F. (2020). Whatsapp group for teaching and learning in Indonesian higher education what's up? *International Journal of Interactive Mobile Technologies*, *14*(13), 150–160. <https://doi.org/10.3991/ijim.v14i13.14121>
- Rambe, P., & Bere, A. (2013). Using mobile instant messaging to leverage learner participation and transform pedagogy at a South African University of Technology. *British Journal of Educational Technology*, *44*(4), 544–561. <https://doi.org/10.1111/bjet.12057>
- Ramkissoon, P., Belle, L. J., & Bhurosy, T. (2020). Perceptions and experiences of students on the use of interactive online learning technologies in Mauritius.

- International Journal of Evaluation and Research in Education*, 9(4), 833–839.  
<https://doi.org/10.11591/ijere.v9i4.20692>
- Robles, H., Guerrero, J., Llinás, H., & Montero, P. (2019). Online teacher-student interactions using WhatsApp in a law course. *Journal of Information Technology Education: Research*, 18, 231–252. <https://doi.org/10.28945/4321>
- Sari, F. M., & Putri, S. N. (2019). Academic Whatsapp group: Exploring students' experiences in writing class. *Teknosastik*, 17(2), 56-65
- Siemens, G. (2005). Connectivism: Learning as network-creation. *ASTD Learning News*, 10(1).
- Snoussi, T. (2019). Learning Management System in Education: Opportunities and Challenges. *International Journal of Innovative Technology and Exploring Engineering*, 8(12S), 664–667. <https://doi.org/10.35940/ijitee.l1161.10812s19>
- So, S. (2016). Mobile instant messaging support for teaching and learning in higher education. *Internet and Higher Education*, 31, 32–42. <https://doi.org/10.1016/j.iheduc.2016.06.001>
- Spencer, D. H., & Hiltz, S. R. (2003). A field study of the use of synchronous chat in online courses. *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, 10-pp. <https://doi.org/10.1109/HICSS.2003.1173742>
- Susilawati, S., & Supriyatno, T. (2020). Online Learning Through WhatsApp Group in Improving Learning Motivation in the Era and Post-Pandemic COVID -19. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 5(6), 852. <https://doi.org/10.17977/jptpp.v5i6.13670>
- Tunjera, N., & Chigona, A. (2022). Educators Synchronously Using Multiple Platforms and Devices for Teaching and Learning During COVID-19 Lockdown. *Journal of Educational Technology Development and Exchange*, 15(1), 81–101. <https://doi.org/10.18785/jetde.1501.06>
- Udenze, S., & Oshionebo, B. (2020). Investigating 'WhatsApp' for Collaborative Learning among Undergraduates. *Etkileşim*, 3(5), 24–50. <https://doi.org/10.32739/etkilesim.2020.5.92>
- Ujakpa, M. M., Heukelman, D., Lazarus, V. K., Neiss, P., & Rukanda, G. D. (2018). Using WhatsApp to support communication in teaching and learning. *2018 IST-Africa Week Conference, IST-Africa 2018, July*, 2–8.
- van den Berg, G., & Mudau, P. K. (2022). Postgraduate students' views on the use of WhatsApp groups as an online communication tool to support teaching and learning during COVID-19. *Perspectives in Education*, 40(1), 112–128. <https://doi.org/10.18820/2519593X/PIE.V40.I1.7>
- Wang, Y., & Chen, N. S. (2009). Criteria for evaluating synchronous learning management systems: Arguments from the distance language classroom. *Computer Assisted Language Learning*, 22(1), 1–18. <https://doi.org/10.1080/09588220802613773>
- Widodo, L. (2019). Users' perceptions of WhatsApp's usefulness in learning. *Open Journal for Information Technology*, 2(1), 1
- Yadav, S. K., Para, S., Singh, G., Gupta, R., Sarin, N., & Singh, S. (2021). Comparison of asynchronous and synchronous methods of online teaching for students of medical laboratory technology course: A cross-sectional analysis. *Journal of Education and Health Promotion*, 10, 1-. <https://doi.org/10.4103/jehp.jehp>
- Yu, K., & Motlhabane, M. G. (2022). WhatsApp's potential to broaden online teaching and learning: Perceptions of undergraduate students from one South African university. *Journal of Information Technology Education*, 21, 547–569

## RESPONSES FROM AUTHORS TO THE REVIEWERS

TOPIC: IMPACT OF WHATSAPP GROUPS ON TEACHING AND LEARNING  
IN HIGHER EDUCATION: A REVIEW

Reviewer ID	Reviewer's Comment	Response to the reviewers
1	This research can address a broader issue rather than focus on a single application. Why should the research community care about WhatsApp groups as the issue itself? Does the paper provide no implication for the readers who does not use the WhatsApp group? I do not think so, which is why I suggest the researchers address a broader issue that can grasp more of the reader's interest.	Provided justification as to why the authors focused on WhatsApp.
1	The current research question is more likely a "practical question" rather than a "research question."	The research question was reworked
1	The theoretical foundation has to be strengthened. Why is connectivism an appropriate theoretical framework for systematic literature review? How is it used in the systematic literature review? To justify the appropriateness of the theoretical framework is expected.	Justification was provided.
1	The research findings should be linked to the theoretical framework. The current findings have no link to the framework, and it is confusing.	The findings were linked to the theoretical framework
1	Moreover, further and deeper discussion on the research findings is also expected.	Discussion was improved
2	I suggest the first sentence to be removed. Start the introduction with the second sentence: Social network applications have ...	Done
2	- Why whatsapp was selected as the object remains unclear. Other media, platforms, or apps share similar features with whatsapp, and they might even have greater potential in the field of higher education. The authors need to provide further explanation regarding the reason for choosing WhatsApp in this study.	Justification was provided
2	- I suggest the authors could explain why adopting "connectivism" as the theoretical	Explanation provided

	framework is appropriate and sufficient to address the research question of this study.	
2	<p>Minor issues</p> <ul style="list-style-type: none"> <li>- Incorrect punctuation.</li> </ul> <p>Section 1 Introduction and Background: considering both the benefits and problems,.</p> <ul style="list-style-type: none"> <li>- Repetitive description of quantity in the sentence.</li> </ul> <p>Section 3 Methodology: The researchers conducted their search from two (2) databases ...</p>	Corrected
2	The authors should explain why adopted an approach of systematic literature review can address the research question	Justification provided
2	The research question seems to highlight how whatsapp influences communication and information sharing, but I did not observe this aspect being addressed in the results analysis.	The research question was reworked and the results were also reworked to link with the theoretical framework
2	The section 2.1 appears somewhat redundant. The authors could consider removing this section and adding literature related to the topic.	This has been moved and integrated into 2.2.

# The Role of Middle Managers in Steering Digital Transformation: A Dynamic Capabilities Perspective

## Abstract

*Middle managers play a pivotal role in refining an organisation's capabilities to achieve business goals. In the research on digital transformation, while top managers receive substantial attention for their strategic formulation, there is comparatively less focus on middle managers. This study, drawing upon dynamic capabilities theory, delves into how middle managers drive the execution of digital transformation by dynamically cultivating three crucial capabilities: sensing, seizing, and reconfiguration. Our research centres on a case study involving a large steel factory leveraging blockchain technology to establish automatous procurement and production processes, thereby evolving from a traditional factory to an ecosystem orchestrator. The study is currently in the data collection stage, progressing steadily. We are confident in presenting our research results during the upcoming conference and eagerly anticipate receiving additional insights from the scholarly audience.*

**Keywords:** Digital Transformation, Middle Manager, Blockchain, Dynamic Capabilities, Qualitative case study

## 1.0 Introduction

Digital transformation (DT) is a prominent topic in the business realm, with nearly 90% of the business executives anticipating positive impacts that information technology (IT) can bring to their companies (Bonnet et al., 2012). Despite this optimism, the complexity of DT leads to a high failure rate, approximately 90%, resulting in adverse impacts on organisations (Ramesh & Delen, 2021). DT represents a fundamental shift in numerous facets of companies, extending beyond mere technological changes (Bouncken et al., 2021; Klein, 2020; Tabrizi et al., 2019). Therefore, the efficacy of DT execution heavily relies on managers in both formulating and implementing DT strategies (Christodoulou et al., 2022).

Numerous scholars have extensively examined the functions of top managers in the formulation of DT strategies. Elbanna (2013) has underscored the critical role of top management in leading organisational transformation through the initiation of large information system (IS) projects. DT often involve dramatic changes in technologies,

value re-creation, structural change, financial aspects, which heavily depends on the strategic decision by top manager (Hess et al., 2016). Despite the prevalent acknowledgement of the significance, role, and actions of top managers (Matt et al., 2015; Westerman et al., 2014; Wrede et al., 2020), the contribution of middle managers is often overlooked. The neglect fails to recognise their integral role in the precise execution and operationalisation of DT initiatives (Nadkarni & Prüggl, 2021).

Middle managers play a distinct role within organisations, translating organisational strategies into operational tasks. This sets them apart from top managers, who predominantly focus on strategy formulation, and frontline employees, who primarily execute operational tasks (Wooldridge et al., 2008). Consequently, middle managers possess the essential knowledge, expertise, and contextual understanding to adeptly navigate the complexities of the organisation and drive impactful change. Their responsibilities extend beyond for execution; they are recognised as orchestrators and synthesisers within organisations (Christodoulou et al., 2022; Floyd & Lane, 2000). Given the distinctive role of middle managers, there is a significant need to explore their influences within DT. To do so, we propose the following research question (RQ): **How do middle managers exert their influence in steering DT projects?**

We argue that the theoretical lens of Dynamic Capabilities (Teece et al., 1997) provides a robust perspective for investigating strategic and operations shifts within organisations (Barreto, 2010; Peteraf et al., 2013; Schilke et al., 2018). Given that DT is the long journey with dynamic change and complexity, dynamic capabilities is particularly well-suited for this study exploring the continuously changing of strategy and implementation during the journey of DT. Our research setting involves in *Mega* company (a pseudonym), a prominent steel company in Taiwan. By conducting semi-structured interviews across various managerial levels and organisation functions, we are able to collect substantial data, unveiling novel insights pertaining to the execution of DT. While our study is ongoing, we anticipate the emergence of compelling findings from this case study. These findings will contribute to a deeper understanding of how middle managers operate in DT projects.

## **2.0 Literature Review**

### **2.1 The Role of Middle Managers in DT**

The concept of DT appeared in the 1990s, coinciding with the establishment of IT/IS solutions designed to address business challenges (Chatfield & Bjørn-Andersen, 1997). Evolving continually over subsequent decades, contemporary digital technology has become a foundational element interwoven into the fabric of modern organisations. Recognised for its transformative potential, digital technology has garnered increasing attentions from both researchers (Bharadwaj et al., 2013; Piccinini, 2015) and practitioners (Fitzgerald et al., 2014; Westerman et al., 2011), emerging as a powerful force capable of addressing strategic intents and guiding organisational transformations.

The introduction of DT can exert a profound impact on organisations, manifesting in both internal and external effects. Internally, it encompasses various aspects including strategy formulation (Bharadwaj et al., 2013; Matt et al., 2015), organisational structure (Ivančić et al., 2019; Selander & Jarvenpaa, 2016), business processes (Ivančić et al., 2019), capabilities (Li et al., 2018), culture (Vey et al., 2017) and leadership (El Sawy et al., 2020; Karimi & Walter, 2015), as well as the underlying business model (Berman, 2012). Externally, DT extends beyond organisational boundaries, engendering far-reaching interactions with a diverse array of external stakeholders, such as business partners, suppliers, clients, government entities, and entire ecosystems (Plekhanov et al., 2022). Furthermore, certain scholars have highlighted the strategic and dynamic changes within DT, suggesting its potential to facilitate comprehensive and holistic impacts on businesses (Chaniyas et al., 2019; Matt et al., 2015). The recognition of these multifaceted influences underscores the complexity and transformative potential associated with DT in organisational contexts.

In organizations, major change initiatives are typically organized by top-level management, and their endorsement plays an imperative role in determining the success of these endeavors (Mugge et al., 2020). Top-level executives possess the strategic authority and access to enabling them to champion transformative alterations within the organization. However, top-level managers are not the only role crucial to the

advancement and execution of these initiatives. Middle managers hold a distinctive and pivotal role in organisations, marked by their diverse set of multifaceted responsibilities and capabilities. Functioning as the precise executors of strategies devised by top-level managers, middle managers additionally serve as advocates and information synthesisers (Floyd & Wooldridge, 1992). This dual role enables them to provide support to both top-level executives and front-line employees in the decision-making process. Moreover, middle managers have the ability to exert influence on top-level managers through their innovative and emergent initiatives, thereby shaping the evolution of the strategic roadmap in response to organisational changes. Thus, middle managers are strategically positioned as intermediaries, connecting strategic dimensions to operational endeavours. This role contributes to fostering innovation, alignment, and harmonisation across both vertical and parallel groups within the organisational structure. This unique position therefore renders them significantly integral to the processes of DT (Chaniias & Hess, 2016; Floyd & Wooldridge, 1992).

With the growing recognition of the strategic significance of middle managers in organisations, the literature has identified four primary organisational dimensions for advancing research in this domain: their strategic involvement, their intermediary role, their contributions to innovation, and their facilitation for DT implementation (Chaniias & Hess, 2016; Floyd & Wooldridge, 1999). Despite the distinctive organisational responsibilities shouldered by middle managers in these dimensions, there is a notable scarcity of research investigating their specific roles in the context of DT (Nadkarni & Prügl, 2021). This underscores the imperative to delve into the unique role played by middle managers in steering the processes of DT.

## **2.2 Dynamic Capabilities**

Dynamic Capabilities Theory refers to a theoretical framework in strategic and organisational management theory (Teece et al., 1997). It emphasises an organisation's ability to adapt or change in response to shifting environments and competitive conditions. Dynamic capabilities involve the organisation's capacity to build and integrate its resources and competencies to address rapidly evolving challenges and/or opportunities. Due to its focused insights into how companies respond to continuous changes in technology and market dynamics throughout the entire journey of strategy



formulation and implementation, it is considered a robust theoretical perspective for investigating strategic and organisational shifts (Teece, 2007). Thus, the dynamic capabilities theory offers a compelling framework for examining DT due to its focus on the intricate and evolving nature of the transformation process. DT involves a continuous cycle of strategic shifts and implementation adjustments, reflecting the dynamic and fluid characteristics of DT. This theory provides a potent analytical lens to dissect the complexities of DT and unravel the evolutionary journey of how organizations adapt their strategies and implementation approaches throughout the DT process (Warner & Wäger, 2019).

Dynamics represent strategic capabilities that involve the construction, integration, and rearrangement of both internal and external resources, enable firms to effectively respond to the circumstances related to transformation (Teece, 2019). Prior literature has delineated three principal dimensions of organisational capabilities, *sensing*, *seizing*, and *reconfiguring*. Sensing involves the perception and assessment of opportunities and threats, seizing entails actions in response to identified opportunities and threats, and reconfiguring involves orchestrating and relocating resources to facilitate transformation activities while establishing an environment supportive of organisational members (Augier & Teece, 2009; Teece, 2007). Utilising the three dimensions of capabilities equips organisations to swiftly navigate and respond to dynamic changes in the market (Eisenhardt & Martin, 2000; Teece et al., 1997).

In Dynamic Capabilities, managers play a pivotal role determining success or failure of the developing such capabilities. Managers are required to utilise and establish the DC making organisation adapt to the dynamic environment and retain the competitive advantages (Heaton & Teece, 2013; Teece, 2016). In addition, Heaton and Teece (2013) highlighted that the role of middle managers involve in not only ordinary routine but also development of changes, including strategic planning and innovation initiating. Hence, adapting this theory enhances the value of exploring middle managers in DT. However, while Heaton and Teece's study (2013) offers an important theoretical concept, empirical study remains absent. These authors emphasised the pressing need for empirical research to underscore the practical significance of this aspect.

### 3.0 Research Methods

We contend that a qualitative approach is the appropriate foundation for exploring the influence to strategy formulation and implementation in DT by middle managers. This methodological choice is grounded in the belief that the nuanced interactions in DT projects can be best elucidated through an in-depth case study (Gerring, 2004). In this research, we embark on a case study involving *Mega* (a pseudonym), a traditional bank in Taiwan.

Mega, a prominent player in the Taiwanese banking sector, employs over 5,000 individuals and possesses a paid-in capital exceeding 82 billion. Its comprehensive range of services encompasses enterprise and personal finance, trust operations, investment, and deposit business. In recent years, Mega has recognized the impact of digital finance on the traditional financial landscape. To maintain its competitive edge, the bank initiated its DT project in 2019. Given its limited prior experience in this domain, Mega commenced its transformation journey by establishing fundamental applications and infrastructure. This foundational step has significantly impacted and comprehensively altered Mega's products, processes, business value, and operating environment.

Our research delves into the comprehensive digital transformation journey of Mega, a traditional bank that has undergone a remarkable metamorphosis through digitization, digitalization, and ultimately, DT. To achieve its DT, Mega meticulously constructed a digitization environment as the cornerstone for further transformation. This initial phase encompassed online and mobile banking, as well as e-services, which laid the groundwork for online banking and provided a solid foundation for DT. The subsequent stage of transformation, digitalization, focused on optimizing digital services and establishing a robust connection with customers. The last phase, DT, integrated cross-functional resources and utilized diverse technologies to create a digital service environment, thereby generating novel value for the bank's products and services. This case illuminates the complexities of the entire DT journey, highlighting the intricate strategies formulated and implemented across various functions and missions. It

presents a valuable opportunity to explore how middle managers influence strategy within DT projects.

For data collection and analysis, we conducted semi-structured interviews to explore and obtain rich data from interviewees. To outline the basic interview questions for understanding the role of middle managers in DT systematically and comprehensively, we integrated the dynamic capabilities theory framework into the different stages of Mega's DT. This framework provided guidance on setting interview questions and obtaining details of the processes in Mega's DT by sensing, seizing, and reconfiguring. Subsequently, the relevant data from interviews will be identified, labeled, and coded to summarize it. This approach will enable us to analyze and conclude the activities and connections between stakeholders in DT and explore the critical actions taken by middle managers.

Our study is currently in the negotiation stage with Mega to conduct interviews with essential stakeholders. We have conducted a preliminary interview with the president of the digital finance department, which leads the DT initiative in Mega, and have obtained the preliminary transcript of the viewpoint of it. We have secured consent to proceed with interviews involving six middle managers in six teams with different missions, including digital education, lean process, online platform, digital ecosystem, AI system, and big data. Additionally, we will invite multiple front-end employees across various DT teams to delve into the interaction between middle managers and front-end employees. These individuals play pivotal roles in the DT project. We anticipate that the gathered data will be comprehensive to analyze and elucidate how middle managers exert their influence in implementing DT projects, thereby steering the organization toward success in DT initiatives.

#### **4.0 Anticipated findings**

There have been limited studies exploring the role of middle managers in the DT process. Therefore, as discussed earlier, this lacuna underscores a pressing need for empirical investigations in this domain. In pursuit of this goal, our study employs an

in-depth case study approach to uncover and clarify the role of middle managers in connecting DT's strategic concepts and implementations.

Additionally, we adopt the Dynamic Capabilities Theory as a robust theoretical framework to dissect the dynamics of strategy formulation and execution across varied stages of DT. This approach enables us to trace the evolving nature of strategies and actions undertaken by middle managers in response to the unique challenges posed at each stage. Specifically, we anticipate that the three dimensions of Dynamic Capabilities Theory—sensing, seizing, and reconfiguring—will diversify into distinct characteristics across the stages of digitization, digitalization, and DT. This analysis contributes valuable insights into the practical realities of capabilities development, providing actionable guidelines for middle managers to enhance the likelihood of success in DT initiatives.

Through this research, we aim to elaborate the mechanisms that middle managers can orchestrate effective DT projects. The findings will not only enrich our understanding of the dynamic interplay between strategy and action in different stage of DT but also offer practical guidance to equip middle managers with the requisite skills and strategies to navigate the complexities of this transformative journey.

## **5.0 Conclusions**

This case study has undergone an intricate process of DT. Through a thorough exploration, we expect to derive research findings that are both comprehensive and valuable. The research is presently in the data collection phase, with an anticipation of presenting additional research results at the upcoming UKAIS conference.

Attending this event provides an opportunity to receive constructive feedback and suggestions from fellow participants, contributing to the refinement and enhancement of the overall research quality.

## References

- Augier, M., & Teece, D. J. (2009). Dynamic capabilities and the role of managers in business strategy and economic performance. *Organization science*, 20(2), 410-421.
- Barreto, I. (2010). Dynamic capabilities: A review of past research and an agenda for the future. *Journal of management*, 36(1), 256-280.
- Berman, S. J. (2012). Digital transformation: opportunities to create new business models. *Strategy & leadership*, 40(2), 16-24.
- Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. v. (2013). Digital business strategy: toward a next generation of insights. *MIS quarterly*, 471-482.
- Bonnet, D., Ferraris, P., Westerman, G., & McAfee, A. (2012). Talking 'bout a Revolution. *Digital Transformation Review*, 2(1), 17-33.
- Bouncken, R. B., Kraus, S., & Roig-Tierno, N. (2021). Knowledge-and innovation-based business models for future growth: Digitalized business models and portfolio considerations. *Review of Managerial Science*, 15(1), 1-14.
- Chanas, S., & Hess, T. (2016). Understanding digital transformation strategy formation: Insights from Europe's automotive industry.
- Chanas, S., Myers, M. D., & Hess, T. (2019). Digital transformation strategy making in pre-digital organizations: The case of a financial services provider. *The Journal of Strategic Information Systems*, 28(1), 17-33.
- Chatfield, A. T., & Bjørn-Andersen, N. (1997). The impact of IOS-enabled business process change on business outcomes: Transformation of the value chain of Japan Airlines. *Journal of Management Information Systems*, 14(1), 13-40.
- Christodoulou, I. P., Wasim, J., Reinhardt, R. J., & Ivanov, K. (2022). The strategic role of middle managers in the formulation and implementation of digital transformation projects. *Strategic Change*, 31(6), 613-622.
- Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they? *Strategic management journal*, 21(10-11), 1105-1121.
- El Sawy, O. A., Kræmmergaard, P., Amsinck, H., & Vinther, A. L. (2020). How LEGO built the foundations and enterprise capabilities for digital leadership. In *Strategic information management* (pp. 174-201). Routledge.
- Elbanna, A. (2013). Top management support in multiple-project environments: an in-practice view. *European Journal of Information Systems*, 22, 278-294.
- Fitzgerald, M., Kruschwitz, N., Bonnet, D., & Welch, M. (2014). Embracing digital technology: A new strategic imperative. *MIT sloan management review*, 55(2), 1.
- Floyd, S. W., & Lane, P. J. (2000). Strategizing throughout the organization: Managing role conflict in strategic renewal. *Academy of management review*, 25(1), 154-177.
- Floyd, S. W., & Wooldridge, B. (1992). Middle management involvement in strategy and its association with strategic type: A research note. *Strategic management journal*, 13(S1), 153-167.
- Floyd, S. W., & Wooldridge, B. (1999). Knowledge creation and social networks in corporate entrepreneurship: The renewal of organizational capability. *Entrepreneurship theory and practice*, 23(3), 123-144.
- Gerring, J. (2004). What is a case study and what is it good for? *American political science review*, 98(2), 341-354.

- Heaton, S., & Teece, D. (2013). The functions of middle and top management in the dynamic capabilities framework. *Kindai Management Review*, 1.
- Hess, T., Matt, C., Benlian, A., & Wiesböck, F. (2016). Options for formulating a digital transformation strategy. *MIS Quarterly Executive*, 15(2).
- Ivančić, L., Vukšić, V. B., & Spremić, M. (2019). Mastering the digital transformation process: Business practices and lessons learned. *Technology Innovation Management Review*, 9(2).
- Karimi, J., & Walter, Z. (2015). The role of dynamic capabilities in responding to digital disruption: A factor-based study of the newspaper industry. *Journal of Management Information Systems*, 32(1), 39-81.
- Klein, M. (2020). Leadership characteristics in the era of digital transformation.
- Li, L., Su, F., Zhang, W., & Mao, J. Y. (2018). Digital transformation by SME entrepreneurs: A capability perspective. *Information Systems Journal*, 28(6), 1129-1157.
- Matt, C., Hess, T., & Benlian, A. (2015). Digital transformation strategies. *Business & information systems engineering*, 57, 339-343.
- Mugge, P., Abbu, H., Michaelis, T. L., Kwiatkowski, A., & Gudergan, G. (2020). Patterns of digitization: A practical guide to digital transformation. *Research-Technology Management*, 63(2), 27-35.
- Nadkarni, S., & Prügl, R. (2021). Digital transformation: a review, synthesis and opportunities for future research. *Management Review Quarterly*, 71, 233-341.
- Peteraf, M., Di Stefano, G., & Verona, G. (2013). The elephant in the room of dynamic capabilities: Bringing two diverging conversations together. *Strategic management journal*, 34(12), 1389-1410.
- Piccinini, G. (2015). *Physical computation: A mechanistic account*. OUP Oxford.
- Plekhanov, D., Franke, H., & Netland, T. H. (2022). Digital transformation: A review and research agenda. *European Management Journal*.
- Ramesh, N., & Delen, D. (2021). Digital transformation: How to beat the 90% failure rate? *IEEE engineering management review*, 49(3), 22-25.
- Schilke, O., Hu, S., & Helfat, C. E. (2018). Quo vadis, dynamic capabilities? A content-analytic review of the current state of knowledge and recommendations for future research. *Academy of management annals*, 12(1), 390-439.
- Selander, L., & Jarvenpaa, S. L. (2016). Digital action repertoires and transforming a social movement organization. *MIS quarterly*, 40(2), 331-352.
- Tabrizi, B., Lam, E., Girard, K., & Irvin, V. (2019). Digital transformation is not about technology. *Harvard business review*, 13(March), 1-6.
- Teece, D. J. (2007). Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic management journal*, 28(13), 1319-1350.
- Teece, D. J. (2016). Dynamic capabilities and entrepreneurial management in large organizations: Toward a theory of the (entrepreneurial) firm. *European Economic Review*, 86, 202-216.
- Teece, D. J. (2019). A capability theory of the firm: an economics and (strategic) management perspective. *New Zealand Economic Papers*, 53(1), 1-43.
- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic management journal*, 18(7), 509-533.
- Vey, K., Fandel-Meyer, T., Zipp, J. S., & Schneider, C. (2017). Learning & Development in Times of Digital Transformation: Facilitating a Culture of

- Change and Innovation. *International Journal of Advanced Corporate Learning*, 10(1).
- Westerman, G., Bonnet, D., & McAfee, A. (2014). *Leading digital: Turning technology into business transformation*. Harvard Business Press.
- Westerman, G., Calm ejane, C., Bonnet, D., Ferraris, P., & McAfee, A. (2011). Digital Transformation: A roadmap for billion-dollar organizations. *MIT Center for digital business and capgemini consulting*, 1, 1-68.
- Wooldridge, B., Schmid, T., & Floyd, S. W. (2008). The middle management perspective on strategy process: Contributions, synthesis, and future research. *Journal of management*, 34(6), 1190-1221.
- Wrede, M., Velamuri, V. K., & Dauth, T. (2020). Top managers in the digital age: Exploring the role and practices of top managers in firms' digital transformation. *Managerial and Decision Economics*, 41(8), 1549-1567.

# Fostering inclusion in open strategy: three IT-enabled mechanisms

## Abstract

*Fostering inclusion has emerged as imperative but inherently complex challenge in the contemporary organisational landscape. This study delves into the role of technology in advancing the cause of inclusion within the context of Open Strategy (OS). Employing a qualitative research approach and drawing from a comprehensive dataset of forty-six expert interviews, we have discerned three underlying mechanisms of technology use in fostering inclusion: welcoming, supporting, and valuing participants. This study contributes to the extant body of OS literature by affording an understanding of technology's pivotal role in the realm of inclusion management. Furthermore, the findings of this study bear practical relevance for organizations striving to enhance inclusive practices, with an objective of mitigating adverse consequences associated with ineffectual diversity management practices, such as tokenism, bias, and discrimination.*

**Keywords:** Inclusion, Technology, Open Strategy

## Introduction

Recent studies in strategic management have embarked on the journey of Open Strategy (OS), which evolves from open innovation and other open forms of strategy making (Whittington et al., 2011). OS has two dimensions of inclusiveness (i.e., involving broader participants) and transparency (i.e., disseminating strategic content and information). In this study, our focus is on the dimension of inclusiveness.

Inclusion has been a key topic in industry and academic in the past decade. However, recent survey on inclusion at work (2022) shows that managers are struggling with effective inclusive practices. For instance, 21% of 1475 employers they surveyed believe that their senior leaders only pay lip service to inclusion management within the organisation. Therefore, it is crucial to find meaningful ways to foster inclusion in OS.

Indeed, one such way of increasing inclusion is through technology use. Many studies have also depicted OS as IT-enabled strategic decision making (Tavakoli et al., 2015, Morton et al. 2019). In these studies, few scholars have shed light on different platforms (e.g., social media) which create psychologically safe environment for increased participation (Baptista et al. 2017) with specific digital features (e.g., anonymity) (Amrollahi and Rolands, 2017) to be used enabling inclusion among extensive and diverse participants. However, our knowledge in how managers fostering inclusion in OS through technology is fragmented. This study, therefore, provides an integrative explanation of underlying mechanisms of how managers foster inclusion in OS through the use of technology. Our research question, therefore, is: *How does technology use by managers foster inclusion in open strategy?*

To answer this question, we adopted a qualitative approach. We interviewed forty-six managers who facilitated OS making. Our informants were purposefully selected (Patton, 2002) with diverse industrial backgrounds and countries to provide data generalizability and credibility. We explored the underlying mechanisms and corresponding components in fostering inclusion through technology use in OS. With this study, we contribute to OS literature by contributing three mechanisms of welcoming, supporting, and valuing participants through technology use. The richness of qualitative approach helped us to uncover the corresponding components of



each mechanism which provides an integrative picture of technology use in fostering inclusion in OS.

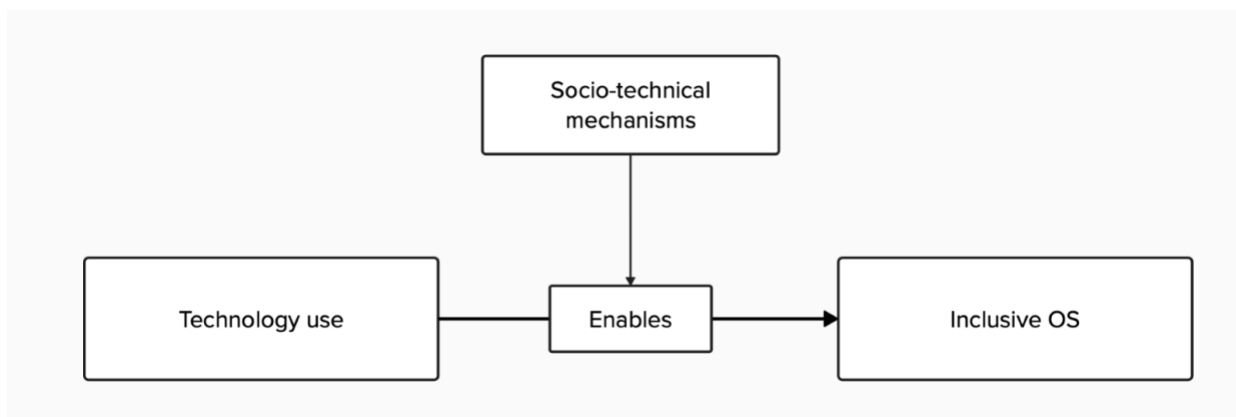
## **Research Background**

In management studies, inclusion is framed as a “strategic goal for diversity management” (Adamson et al., 2021). It serves as a key driver and basis for reaping the potential benefits of diversity (Ferdman and Deane, 2014). Combs et al., (2019, p. 279) further noted that inclusion involves “ways that organisations can maximise the benefits of diversity by fostering and promoting full rights, access, and privileges of employment and advancement to all organisational members”. To create an inclusive organisation, managers are often required to implement practices that mitigate negative effects of diversity management such as those grounded in discrimination, opposition, and prejudice (Ferdman and Deane, 2014). Additionally, they must address barriers that prevent individuals from full participating in organisational systems and from utilising their skills and potentials to the fullest extent (Adamson et al., 2021). Furthermore, it is worth noting that having inclusive practise in organisations, does not only mean including those marginalised groups, but also involves both majority and minority groups having equal opportunities to access and influence organisational systems and core issues (Dobusch, 2019).

Open Strategy is an example of attempting to be more inclusive in organisational core issues. OS is a form of strategic decision making that “widening the search for strategy ideas and improve commitment and understanding in strategy implementation” (Whittington et al., 2011, p.535). OS has two dimensions, namely inclusiveness and transparency. Inclusiveness refers to “the range of people involved in making strategy” (Whittington et al., 2011, p.531). and transparency refers to “the visibility of information about an organisation's strategy, potentially during the formulation process but particularly with regard to the strategy finally produced” (Whittington et al., 2011, p.536). OS is also IT-enabled (Tavakoli et al., 2015; Whittington et al., 2011). The use of technology enables collaborative creation, circulation of strategy-related content across the organisational boundaries (Kaplan, 2011).

Existing OS studies suggest that technology use can foster inclusion (Amrollahi and Rowlands, 2017; Mount et al., 2023; Schlagwein et al., 2017). Indeed, different studies explore different ways of how use of technology can lead to more inclusive decision making through different mechanisms. One of the mechanisms discussed in the literature is broadcasting strategic information to the participants, which indicates technology use is for informing and attracting potential participants and their contributions in OS. For instance, digital tools like PowerPoints presentations, web-based questionnaires, blogs, emails increase the participants engagement and encourage further participation (Morton et al. 2020). Another mechanism noted in the literature is enabling one to be their authentic self. For instance, with the feature of anonymity on the social media platforms, it provides psychological safe virtual environment for participants to be their authentic self while contribute meaningfully (Amrollahi and Rowlands, 2017; Mount et al., 2023). Lastly, technology use also empowers participants in OS making. For instance, features like voting and commenting on social media empower participants to influence strategy outcomes, allowing everyone’s input being heard (Dobusch and Kapeller, 2018).

Despite the contributions of these studies, our understanding in fostering inclusion in OS is still fragmented, as most of the literature only focusing only one mechanism of inclusion management through the use of technology. This study aims to uncover the underlying mechanisms associated with technology use in fostering inclusion in a more integrative way, Figure 1 highlighted the research focus of this study, in which we aim to explore how the use of technology can lead to a more inclusive strategy making context.



**Figure 1- Conceptual framework highlighting our area of research focus**

## Research Methods

We took a qualitative approach in this study. The richness of qualitative data enabled us to understand the socio-technical mechanisms that underpin management, making it ideal for to uncover the underlying mechanisms and their corresponding components of technology use in fostering inclusion in OS.

### Data collection

Our main data collection method was through expert interviews, like the study of Lorenz and Buchward (2023), expert interviews enabled us to gather first-hand and in-depth perspectives from our purposefully selected informants. During the period of April 2022 to March 2024, we conducted a total of forty-six expert interviews. In this period, we conducted three rounds of the expert interviews. In the first round, we conducted fifteen interviews, asked open and general questions centred around a) background of technology use in fostering inclusion in OS. In the second round, we conducted further twenty-five interviews and asked more explicit questions after the first-round interview analysis, deepened the questions to understand b) what the detailed practices that the informants conducted to use different technologies in their OS episodes and c) the outcomes of inclusion in OS supported by technology, both individual and organisational levels (Illustrated in [Table 2](#)). To assure that we reach theoretical saturation, which refers to “no additional data are being found whereby the sociologist can develop properties of the category” (Glaser and Strauss, 1967, p.61), we conducted further six interviews in the last round of expert interviews. To enrich our findings, we also engaged with publicly available documents, including user stories, white papers and online reviews of different technology use in fostering inclusion ([Table 1](#)).

Data source	Total and breakdown
Expert interviews	46 interviews (mean length 40 minutes) with 45 respondents generated.
Secondary data of publicity available information	Approx. 65 user stories; 56 platform white paper; approx. 850 Online reviews of the selected platform.

**Table 1-Data Collection**

Interview question categories	Interview questions
a) Background of technology use in fostering inclusion in OS.	Could you please share us a your most recent OS practice?
	Could you please tell us what inclusion means to you?
	Why do you think inclusion is important when your open your strategy making process?
	How did the platform benefit you to achieve inclusion?
	Do you think technology plays an important part in your inclusive OS practice?
b) Detailed practices of using technology to foster inclusion in an OS episode.	Could you please give us an example what did you do to foster inclusion?
	During your OS making process, how technology helped you in different stages of OS making to foster inclusion?
	How did technology help you to attract participants to join?
	How did technology help you to make participants engaged?
	How did technology help you to evaluate and incorporate participants' strategic inputs?
c) The outcomes of inclusion in OS supported by technology	How do you think the overall process you took help you to foster inclusion in your organisation?
	What are the impacts for individuals?
	What are the impacts for organisations?

**Table 2-Interview Questions**

### **Participants recruitment**

We used purposeful sampling to identify the key informants in this study (Patton, 2002). The informants qualified as key informants in this study were managers who had experiences and acted as a strategic decision facilitator (Ackermann, 1996) of using technologies to engage with diverse participants in OS ([Appendix A](#)). We reached out to our key informants mainly online and through two approaches: First, we reached out the users of certain digital platforms which enabled managers to include diverse and extensive participants in OS. We identified these platforms through a review website of these platforms (e.g., G2. com). The selection criteria for selected platforms in this study include a) having certain digital features enable managers to include extensive participants in strategy making; and b) having a large number of active users. As the result, we selected Mural, Miro, ThoughtExchange and Leapsome in this study, and we reached out their users through G2 platform, email, and LinkedIn.

Second, we found other informants by attending relevant Diversity and Inclusion (D&I) workshops, events and presentations. In this approach, we particularly looked informants who were specialised in D&I practices using technology in their career and recruited them accordingly through emails and LinkedIn. We assured all our informants understand the concept of OS before the interview, and guided them to share their story of using technologies in fostering inclusion from a particular OS episode during the interview process.

## Data analysis

We preconcept three types of mechanisms for fostering inclusion in OS through technology are welcoming, supporting and valuing participants from different OS literature. These preconceptions are important to the theory development in this study, yet not completely drive our data analysis. Like study conducted by Gregory and Kaganer et al. (2018), we deliberately managed these preconceptions to “avoid forcing-fitting existing theory onto our data (Birks et al. 2013)” but treating these preconceptions “more similar to additional slices of data to compare with our own data” (p.1230).

To do so, we adopted content analysis in our research (Patton, 2002, p.453), which “refer to qualitative data reduction and sense-making effort that takes a volume of qualitative material and attempts to identify core consistencies and meanings”. The core of content analysis is to find patterns and themes to answer the research question (Patton, 2002). The data analysis was taken three steps, grounded, inductive reasoning (Langley, 1999; Strauss and Corbin, 1998) as shown in [Table 3](#), to derive the three-underlying mechanisms and corresponding components for achieving inclusion through technology.

First, we categorise concepts that appear as inclusive practices through open coding (Strauss and Corbin, 1998). At this stage, we analysed our data from inductive approach, we coded our data line by line and focused on distinctive digital features and managerial practises that enabling inclusion from all strategy episodes we interviewed.

In the second stage, we conducted axial coding (Strauss and Corbin, 1998) to make connections between categories and subcategories and identified the link between the codes, and this is where we draw a definition of what would qualify as a corresponding component of inclusion in view of empirical evidence and relevant literature. We further categorised them into broader codes, where we labelled as mechanisms. In this stage, we derived three mechanisms and relevant 10 components, we labelled the former based on their empirical characteristics as welcoming, supporting and valuing (see [figure 2, 3 and 4](#)).

Finally, we focused on generating a model of fostering inclusion through technology ([Figure 5](#)). In this stage, we conduced selective coding (Strauss and Corbin, 1998) and illustrated the intertwined relationship among the three mechanisms for inclusion over time. The outcome of this stage is to produce a conceptual model of fostering inclusion through technology in OS.

All our data analysis followed intercoder reliability (O'Connor and Joffe, 2020) to provide the trustworthiness and credibility of the data coding and analysis in this study. We initially generated approximately 130 open quotes; 24 open codes, and 12 axial codes. We agreed approximately 70% of the quotes, 13 open codes and 6 axial codes during our first inter-coder agreement check between 6-7<sup>th</sup> September 2023. The agreement rate increased at our second inter-coder agreement check which was placed on 12<sup>th</sup> of September 2023, with 95% of the interview quotes agreed, 100% of agreement on 24 updated open codes, and 100% of agreement on 10 updated axial codes. We then conducted a final check on 21<sup>st</sup> of September 2023 of all coding analysis and achieved full agreement between the first two authors of the study.

Stages	Tasks	Outputs
1. Generating themes and patterns through open coding.	<ol style="list-style-type: none"> <li>1. Identify distinctive digital features and managerial practise for inclusivity practices.</li> <li>2. Ensure reliability, generality and credibility of open coding by establishing theoretical saturation.</li> <li>3. Achieve intercoder reliability between authors.</li> </ol>	Categorise emerging themes and patterns through open coding.
2. Identifying inclusion mechanism through axial coding.	<ol style="list-style-type: none"> <li>1. Conduct axial coding to identify mechanisms and relevant components in view of empirical evidence and relevant literature.</li> <li>2. Derive three mechanisms and each mechanism's relevant components.</li> <li>3. Consider how each mechanism relates to yet differs to other mechanisms.</li> <li>4. Enhance reliability, generality and credibility of axial coding by establishing theoretical saturation.</li> <li>5. Achieve intercoder reliability between authors.</li> </ol>	Identifying mechanisms (Figure 2,3 and 4)
3. Generate conceptual model of fostering inclusion through technology in OS.	<ol style="list-style-type: none"> <li>1. Build on the findings emerged from stages 2 and 3 to examine how the three mechanisms distinctive but relevant to each other both empirically and theoretically.</li> <li>2. Construct the conceptual model from the three mechanisms to indicate how managers can benefit from this model to fostering inclusivity through technology.</li> </ol>	Conceptual model (Figure 5)

**Table 3-Data Analysis Stages**

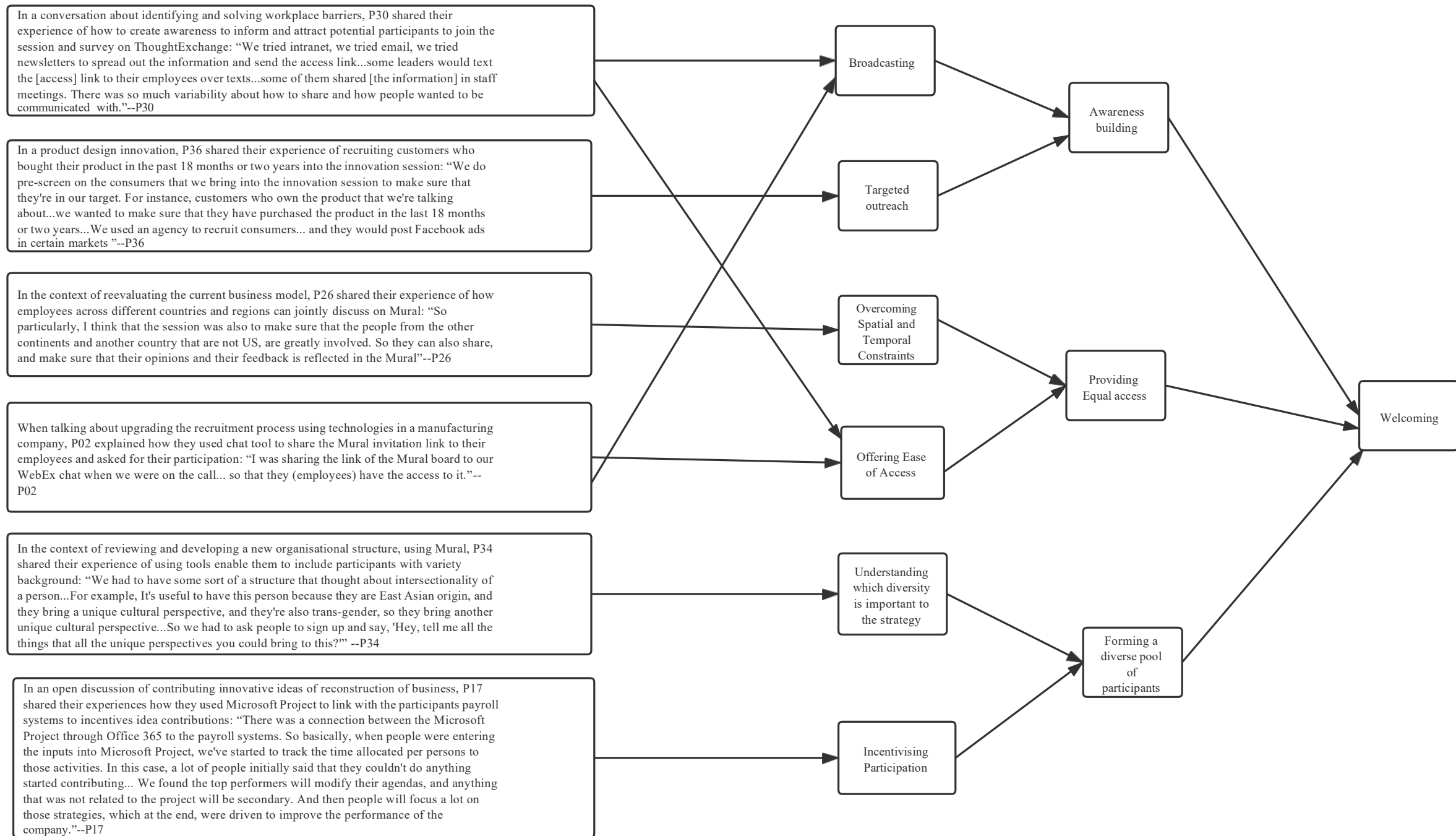


Figure 2- Exemplar coding of underlying mechanism 1: Welcoming

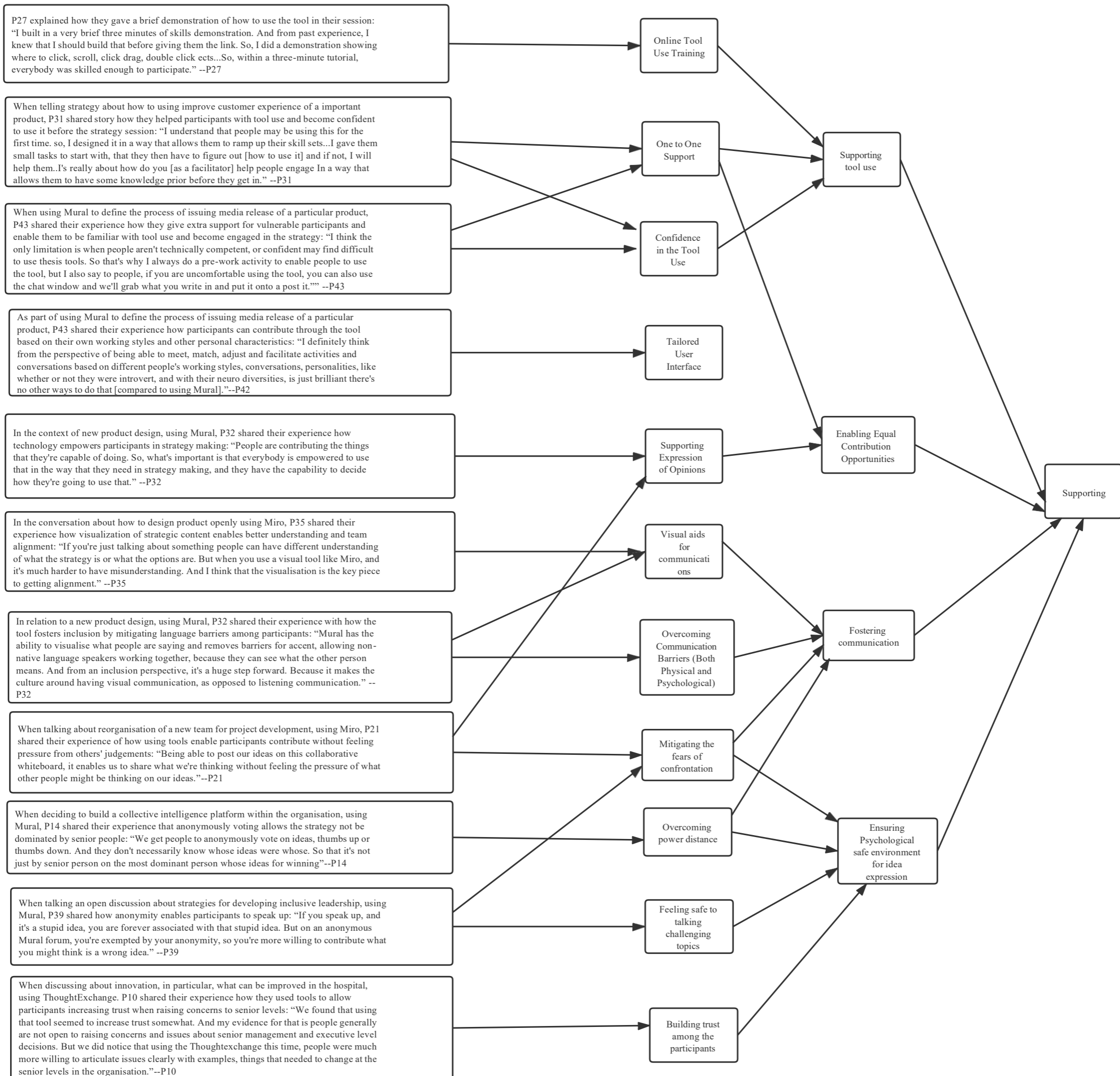


Figure 3- Exemplar coding of underlying mechanism 2: Supporting

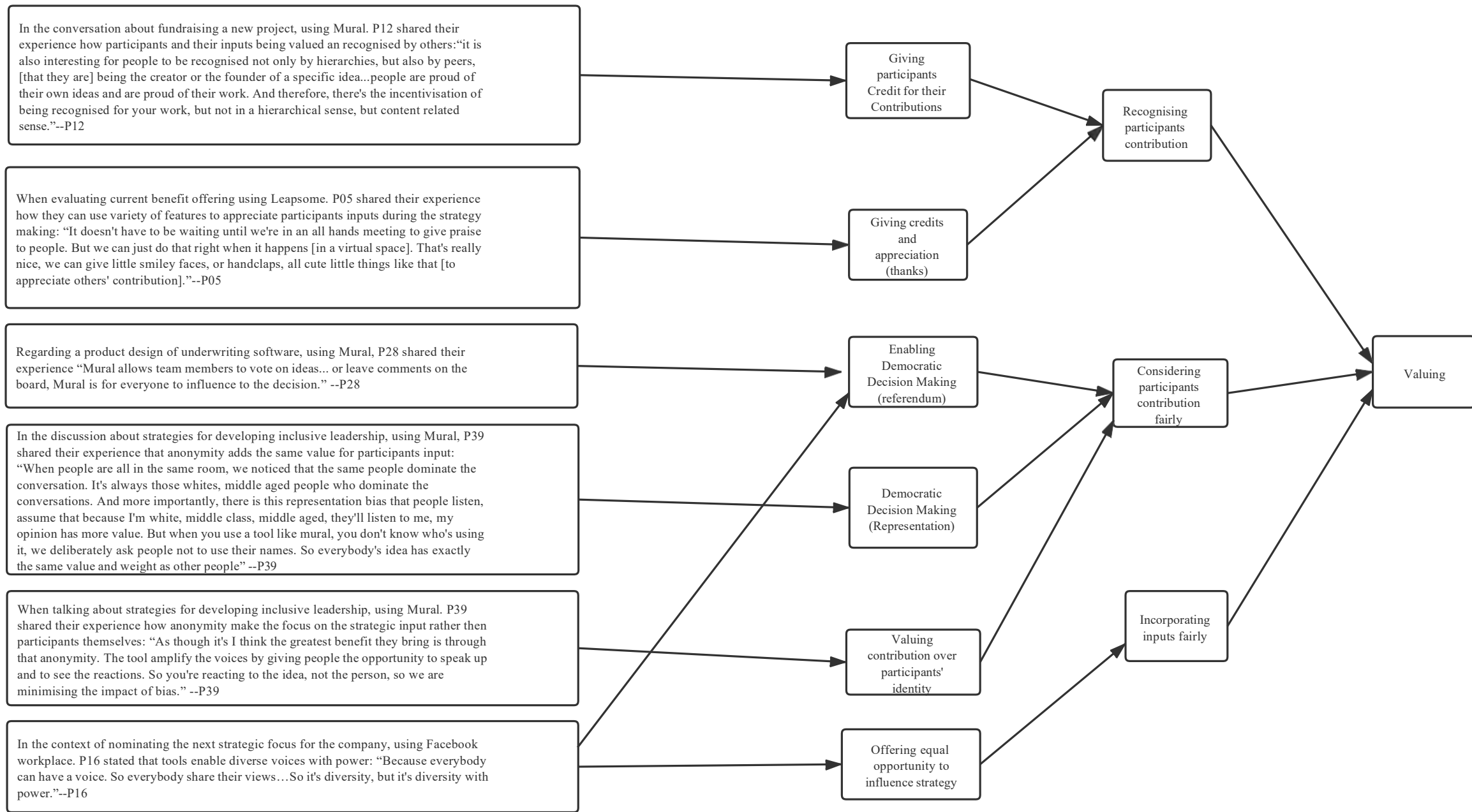


Figure 4-Exemplar coding of underlying mechanism 3: Valuing



## Findings/Theory Development

We approached the identification of underlying mechanisms relevant to managers' use of technology for inclusion by examining what our interview informants reported they were able to achieve inclusion through using technology. Our data analysis revealed that in the process of OS, managers do welcome, support and value participants to achieve inclusion (Figure 5). We further highlighted 10 corresponding components of the three mainchains (Table 4). For each mechanism, we outline the specific corresponding components that our informants mentioned in their process of OS and indicate their importance in fostering inclusion in OS. We feel that collectively, these three mechanisms with their corresponding components present an integrative and detailed reflection on how managers use technology to foster inclusion in OS.

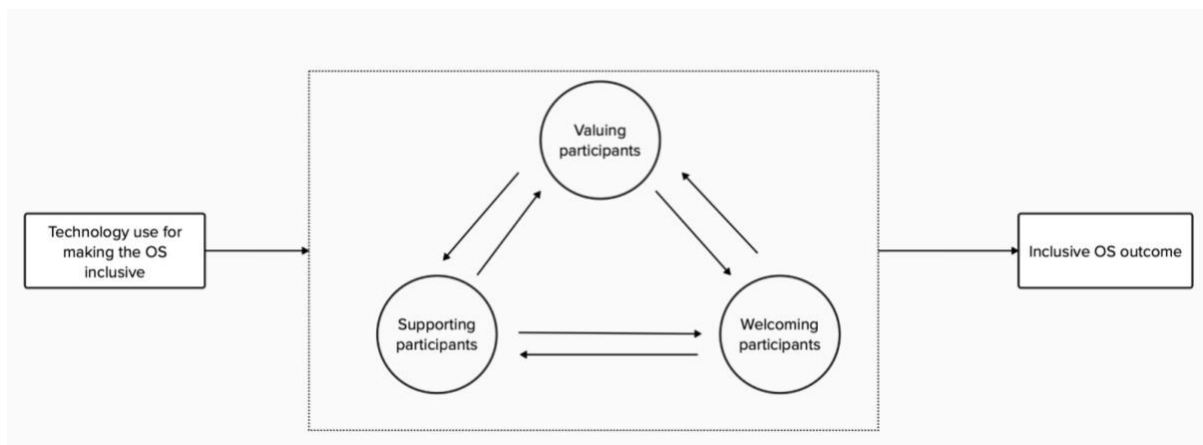


Figure 5-process of how managers foster inclusion through technology over time in OS.

Mechanisms	Definition	Components
Welcoming	The mechanism by which managers use digital tools to intentionally and actively to disseminate information and attract participants to join the OS.	<ul style="list-style-type: none"> <li>• <b>Building awareness:</b> The utilisation of digital tools to broadcast as well as targeted outreach awareness of OS to all and specific participants.</li> <li>• <b>Forming a diverse pool of participants:</b> The practices to form a pool of participants based on the context-specific diversity requirement of strategy decision.</li> <li>• <b>Giving opportunity of equal access:</b> The utilisation of digital tools to provide an equal and easy access to a virtually synchronised workplace without spatial constraints.</li> </ul>
Supporting	The mechanism by which managers use digital tools to intentionally and actively to increase participants engagement and commitment in the OS.	<ul style="list-style-type: none"> <li>• <b>Supporting tool use:</b> the practices and process of familiarizing and becoming knowledgeable about the proper and effective use of specific tool or software application.</li> <li>• <b>Enabling equal contribution opportunities:</b> The utilisation of digital tools to allow participants contribute their input equally regardless their diverse backgrounds and contexts.</li> <li>• <b>Fostering communication:</b> The utilisation of digital tools to produce better communication among the participants, including using visual aids, such as diagrams, images and pictures; and to overcome</li> </ul>

		<p>communication barriers, both physical and psychological.</p> <ul style="list-style-type: none"> <li>• <b>Creating psychologically safe environment:</b> The utilisation of digital tools to support participants to be true to their authentic self without worrying about negative consequences of speaking up.</li> </ul>
Valuing	The mechanism by which managers use digital tools to intentionally and actively recognise, appreciate, consider and incorporate participants' contributions in OS.	<ul style="list-style-type: none"> <li>• <b>Recognising participants contribution:</b> The utilisation of digital tools to give participants credits and appreciation for their inputs.</li> <li>• <b>Considering participants contribution fairly:</b> The utilisation of digital tools to produce democratic strategic making process.</li> <li>• <b>Incorporating inputs fairly:</b> The utilisation of digital tools to give participants' inputs further consideration and influence on the strategy outcomes.</li> </ul>

**Table 4-Mechansims and their corresponded components of technology use in fostering inclusion in OS.**

### **Mechanism 1: IT-enabled Welcoming**

For many of our informants, the first practise to make participants feeling included through technology is welcoming their participation in upcoming OS episodes. We identify welcoming as the first mechanism to managers in fostering inclusion through technology in OS. We define welcoming is a mechanism by which managers use digital tools to intentionally and actively to disseminate information and attract participants to join the OS. Our data revealed that to welcome participants, manager use tools to include all as well as target outreach certain groups of individuals who are under-represented in OS making.

#### **Building awareness**

Many managers use the digital tools with the aim to build awareness of an upcoming OS episode to all participants. For instance, P10 mentioned their experience how they used email, posters and QR codes to inform potential participants, especially to inform the ones who were not sitting in front of the computer all the time to join an upcoming innovation session on ThoughtExchange, where they jointly discussed things needed to be improved in the hospital services:

“So, every employee would receive a link to do the engagement survey and the link to access the ThoughtExchange through email...We also knew some people were not sitting at their computer all time, so we put posters up around the hospital that advertised it, and pointed people to it. We provided the QR code for people who forgotten the password, so they could just go to the poster scanning it and to join the survey right away.”

In many other OS episodes, managers do not only broadcast OS related information to inform all participants, but also targeted reach out specific participants and using different digital tools to attract their participation. For example, in a discussion concerning strategies for anti-racism in a university, P41 decided to include students with minority backgrounds by using newsletters and emails, particularly through targeted emails:

“It was an open call, and everybody could contribute to it. But our primary focus was on getting opinions and perspectives from staff and students of colour.

Thinking particularly students who are brown, and black. So, what we did was we sent out an initial invitation just through the university newsletter for people to the town hall. And then we also sent out emails to specific student groups. The groups were focused on race, ethnicity and cultural heritage.”

### **Forming a diverse pool of participants**

To assure that OS making includes meaningful inputs from diverse participants, many informants highlighted their experience to form a pool of participants based on the context-specific diversity requirement of the strategy decision. For instance, when talking about open strategies for developing inclusive leadership, P39 shared their experience of using Mural to include underrepresented groups and foster contribution:

“That is exactly the reason we use Mural, [It allows us] to be able to get the opinions and the views from underrepresented groups. Previously, we just had a Microsoft Teams channel and people could post an idea in there. And I'll be honest, the feedback was brutal...and for about nine months, no one has posted a single idea. About eight weeks ago, we relaunched the ideas forum in Mural. And just in those eight weeks, we have 30 ideas. Two of them which I believe are from underrepresented groups, are actively developing into our business strategy.”

On few occasions, managers also incentivise participation through technology, including linking participants strategic performance to their payroll systems through Office 365 (P17) and sending digital vouchers. In a product design innovation, P36 shared their experience of encouraging end customers participation by sending out digital vouchers on online survey:

“In those sessions, some of other tools that we used were digital survey, quick surveys like mobile surveys with consumers...So we've found some tools that fairly quickly and cheaply allow us to ask consumers questions. We've also found tools that allow us to recruit consumers locally to come in... So, I submitted a survey online, and consumers would answer it, who were in our target audience. And after a couple of days, and a few \$1,000 spent, we had talked to 200 consumers about their experience with the product.”

### **Providing opportunities of equal access**

Many of our informants believe that with the support of technologies, it provides equal opportunities for participants to join and access to a virtually synchronised workplace without spatial constraints. In the context of re-evaluating the current business model, P26 shared their experience of how employees across different countries and regions can jointly discuss on Mural:

“So particularly, I think that the session was also to make sure that the people from the other continents and another country that are not US, are greatly involved. So, they can also share, and make sure that their opinions and their feedback is reflected in the Mural.”

In fact, managers also mentioned the platform itself also provides equal and easy access to participants with a shared link. For instance, giving some examples of previous strategy projects that were opened to discussion, P24 mentioned all relevant stakeholders of projects can join the Mural as long as they have the access link:

“There was nothing in the tool that would have stopped anybody from joining. If you have the link, you can get into it.”

## **Mechanism 2: IT-enabled Supporting**

Our data also revealed managers use technologies to support their participants in fostering inclusion in OS by making the participants to be engaged in the OS making process. We highlighted supporting as the second mechanism in fostering inclusion in OS and identifying this mechanism by which managers use digital tools to intentionally and actively to increase participants engagement and commitment in the OS.

### **Supporting tool use**

Many of our informants indicated that during the process of OS, it is crucial to allow participants having practises and process of familiarizing and becoming knowledgeable about the proper and effective use of specific tool or software application. Our data revealed that managers can support participants being familiar with the tool use by providing online and large group training, one to one support, especially for digitally less competitive groups (i.e., elder people); tailoring user interfaces to meet specific participant’ needs and to boosting their confidence of tool use. For instance, P43 shared their experience of giving demonstration workshop of tool use before their OS session:

“The first thing that I wanted to do was to ensure that everyone could use Mural. So, what I did there was to create a place on the [Mural] board for people to go into before the actual meeting session. I gave them a short video of how to use mural and ask them to click on some post and notes and try to add in some information anonymously... I also asked them three questions before the meeting, that was mainly for me to understand how people were feeling about [using] it. So, I knew what to expect when I was facilitating the meeting. But also, to make sure that people were familiar with [using] Mural.”

Additionally, our data revealed that managers do give extra support for participants who are not digitally competitive. For example, when talking one Mural session strategy, which about improving customer experience of an important product, P31 shared a story how they helped participants who are not technically competitive by giving different contribution options on the tool and boosted their confidence in tool use:

“I think the only limitation is when people aren't technically competent, or confident may find difficult to use thesis tools. So that's why I always do a pre-work activity to enable people to use the tool, but I also say to people, if you are uncomfortable using the tool, you can also use the chat window and we'll grab what you write in and put it onto a post it.”

The platform itself also supports with the customised user preference, allowing participants to alter the user interface based on their own preference. For example, in the conversation about cost-effective strategy for future projects, using Mural. P24 shared their experience of how the tool enables participants to contribute with different features which based on their own selected preference:

“I've used the templates, and I've tailored that template for the audience...They can go into different sections of the tools...people can go in and they can create a virtual post and note, they can write a comment into that, they can put pictures into it, they can embed it in the video or any digital content, sounds, and links to websites, there's all sorts of ways they can add information into the tool that's relevant to the framework that's been set up.”

### **Enabling equal contribution opportunities**

Many of the informants responded that the platform they used enable participants to express their opinions equally regardless their diverse backgrounds and contexts. For example, P45 shared their experiences in a product design session, using Mural, all participants could contribute equally regardless their position in the organisation:

“It doesn't matter what kind of position you are in. You would still need to be able to go to this virtual area where everybody can contribute, also maybe [contribute] something that is beyond your job.”

Similarly, when developing a new website becomes necessary to attract more potential customers for an open talent recruitment service company, P07 believed tools like Leapsome gives everyone an opportunity in the organisation to contribute:

“I personally believe in that everybody comes with different experiences and different backgrounds. So just because you see something in a certain way does not mean somebody else can't add value or [having] a different perspective. So, I truly enjoy Leapsome give everybody an opportunity to have a blast.”

### **Fostering communication**

Our informants further described of using digital tools to foster better communication among the participants, particularly with the support of visual aids, allowing participants to increase their mutual understanding of strategic content and “reading the same words in the same way, with the same mental images” (P25). For instance, when talking about a product design of an underwriting software, using Mural. P28 shared their experience of how the tool enabled them to convey messages through visualisation for better understanding among the diverse participants and to collectively develop the product:

“What happens in most of the times, they (participants) do not communicate what there they have in their mind. But when comes to visually, they use these templates and sticky notes, they use this wire framing tools to convey their message. [Visualising] what they're actually thinking about the product, how it's going to look like and how the process flow would look like. That helps a lot.”

Our informants also highlighted the use of selected tools can overcome communication barriers among the participants, we noted these communication barriers can be both physical and psychological. For example, as part of the new product design development process, using Mural. P32 shared their experience how the tool fosters better communication and allows participants to feel included when using the tool to mitigate language barriers:

“Mural has the ability to visualise what people are saying and removes barriers for accent, allowing non-native language speakers working together, because they can

see what the other person means, and from an inclusion perspective, it's a huge step forward. Because it makes the culture around having visual communication, as opposed to listening communication.”

In other cases, communication barriers could be derived from social and power distance, using technology allows participants to mitigate these barriers and foster their communication. For example, in the discussion about strategies for developing inclusive leadership, using Mural, P39 shared their experience that anonymity overcome social barriers and foster better communication:

“When people are all in the same room, we noticed that the same people dominate the conversation. It's always those white, middle-aged people who dominate the conversations. And more importantly, there is this representation bias that people listen, assume that because I'm white, middle class, middle aged, they'll listen to me, my opinion has more value. But when you use a tool like mural, you don't know who's using it, we deliberately ask people not to use their names. So, everybody's idea has exactly the same value and weight as other people.”

### **Creating psychologically safe environment**

Our informants addressed that in order to support participants to feel being included, one of the important practices is to create a psychologically safe environment where participants can be true to their authentic self without worrying about negative consequences of speaking up. Our data revealed that using technology, especially with the feature of anonymity, managers can support participants to mitigate the fears of confrontation, overcome power distance, feel safe to talk challenging topics and build trust among the participants. For instance, P03 mentioned how Miro helps those participants who usually do not feel comfortable with public speaking:

“I think Miro is a nice tool, because it democratises the strategy process, and it allows everyone to share their ideas, even if they're not comfortable speaking up in front of a large group.”

Similarly, P23 shared how anonymity helps their participants to overcome power distance by expressing their thoughts without fear in an open discussion, which was about defining new structure of work using Miro:

“Some people don't want to raise their hand and speak up when there are leaders or there are different audience. Miro help us to be anonymously and that help them (participants) to express their opinions without fear. And it was very useful to have all these different opinions [from different participants] ...because they were heard without any bias.”

Our data further revealed that through the use of technologies, it allows participants to feel comfortable when talking sensitive and challenging topics. P41 shared their experiences of how Mural helped their strategy session in defining anti-racism strategies in a UK university:

“Mural is not necessarily as sort of confrontational as physical space might be. For example, when you're talking something about anti-racism, which is particularly an elusive topic. Having a digital space, is sort of creating some distance. Where when you're physically in a room together talking about this kind of stuff, you worry if

saying this is considered in appropriate and a little bit emotive. But having a sort of digital space makes it feel safer. It also brings more people to join.”

Our data indicated that technology also increase trust among the participants in the process of OS making. For instance, in relation to an innovation session of improving hospital services using ThoughtExchange. P10 shared their experience how they used tools to allow participants to increase trust when raising concerns to senior levels:

“We found that using that tool seemed to increase trust somewhat. And my evidence for that is people generally are not open to raising concerns and issues about senior management and executive level decisions. But we did notice that using the Thoughtexchange this time, people were much more willing to articulate issues clearly with examples, things that needed to change at the senior levels in the organisation.”

### **Mechanism 3: IT-enabled Valuing**

Valuing is the third underlying mechanism we found in our data in fostering inclusion through technology. We define this mechanism as one by which managers use digital tools to intentionally and actively recognise, consider and incorporate participants’ contributions in OS.

#### **Recognizing participants**

Our informants stated that they showed their valuation of participants’ contribution by giving credits and appreciation to their contributions first. For instance, when talking about the OS conversation of fundraising a new project, using Mural, P12 shared their experience how participants and their inputs being give credits and appreciated by others:

“It is also interesting for people to be recognised not only by hierarchies, but also by peers, [that they are] being the creator or the founder of a specific idea...people are proud of their own ideas and are proud of their work. And therefore, there's the incentivisation of being recognised for your work, but not in a hierarchical sense, but content related sense.”

Similarly, P05 shared their experience of how they can use variety of features to appreciate participants’ inputs during their strategy making of re-evaluating organisations’ benefit offering to end users:

“It doesn't have to be waiting until we're in an all hands meeting to give praise to people. But we can just do that right when it happens [in a virtual space] ...we can give little smiley faces, or handclaps, all cute little things like that [to appreciate others' contribution].

#### **Considering participants’ contributions fairly**

To value participants contributions, our informants further revealed their experiences of using technologies like online voting to produce democratic decision making. For example, when talking about an OS product design decision for an underwriting software development, using Mural, P28 shared their experience of using online voting through referendum which gives all participants a direct democratic decision power:

“Mural allows team members to vote on ideas... or leave comments on the board, Mural is for everyone to influence to the decision.”

We also noted in our data that managers value participants' contribution by considering their contributions fairly not only through referendum but also through representation of different diversity groups. The latter focus on providing equal weights on participants' contributions, such practice gives extra weight to minority voices and enable their voces to have the same weight as majority groups. For instance, when discussing about making strategies for developing inclusive leadership, P39 shared their experience of using Mural to incorporate more radical ideas into strategy outcome from underrepresented groups in the organisation:

“That is exactly the reason we use mural, to be able to get the opinions and the views from underrepresented groups. Within our organisation internally, we have an ideas forum, We're a small organisation with about 25 people. And it can be difficult in a in a small organisation to put forward radical ideas, particularly if you're new, particularly if you're not male, middle class, middle aged.”

Our informant also mentioned that technology use in OS also mitigate bias, allowing the focus centred around on the strategic content rather than the person who contributes. For instance, P39 shared their experience of how to use Mural, especially with the feature of anonymity to minimise the impact of bias:

“I think the greatest benefit the tool bring is through that anonymity. The tool amplifies the voices by giving people the opportunity to speak up and to see the reactions. So, you're reacting to the idea, not the person, so we are minimising the impact of bias”.

### **Incorporating inputs fairly**

Lastly, our informants revealed that using technologies giving everyone's contribution an opportunity to be incorporated into final strategy outcome. As informant 16 stated: “It is diversity, but it is diversity with power”. Other informants also illustrated how they use tools to value participants inputs and give generated inputs further considerations and opportunities to be incorporated into the strategy. For instance, in relation to the new organisational website development across different stakeholder and phases in a staffing company, using Leapsome, P07 shared their experience of how they valued participants' inputs with further consideration:

“When somebody has a valid point, what we did it's really about trying to understand where they're coming from, and what, and why they're feeling the way that they are. Even during the website development, we were so close to launch a site. And there were some great ideas that came from some of the latter conversations...We always take those into account and have managed to incorporate them into it... and we will always give further considerations of the ideas and could incorporate those ideas to further similar projects if they were not selected this time”.



## **Discussion**

In this paper, we deepened the role of technology in fostering inclusion in the context of OS. By conducting forty-six expert interviews with managers who facilitated OS episodes with diverse and extensive participants, this study provides an integrative view of technology use in fostering inclusion in OS. In particular, we uncovered the three underlying socio-technical mechanisms of IT-enabled welcoming, supporting and valuing participants, along with their ten corresponding components (Table 4) which intertwiningly foster inclusion in OS (Figure 5).

### **IT-enabled welcoming mechanism**

We suggest that technology use contributes to inclusive practises by welcoming all potential participants to join upcoming OS sessions. Our study contributes to studies (Morton et al., 2020) that tools like social media, webpages and blogs enables managers to broadcast strategic related information to inform potential participants. Beyond such informing, our study revealed digital tools (i.e., online survey and digital voucher) can not only build awareness to all potential participants, but using tools (i.e., emails, direct online chat) to reach out targeted participants to join the upcoming OS session. In this sense, inclusion is not only focusing on welcoming regular participants but also to welcome participants who are historically left out in the OS by sending out targeted invites. This study also consolidates that technology can mitigate geographic barriers to foster inclusion in group decisions (Wirtz et al., 2018). By demonstrating of using digital access, such as sharing a platform link, every participant can effortlessly join the OS session with a simple click on their devices, eliminating the need for in-person presence. Such digital access greatly increases participation and inclusion, bringing geographical gaps for diverse participants across different regions to work together. The study further noted that many of the informants aim to achieve inclusion by deliberately increasing diversity of the pool of the participants, such as including participant with expertise subject to strategy content, like the study Malhotra et al., (2017) conducted, and including minorities and underrepresented voices in the composition of the participants' pool.

### **IT-enabled supporting mechanism**

Our second mechanism suggests that technology use can help managers to foster inclusion by supporting participants to feel engaged and committed in OS. Doing so can eliminate organisational silence, which refers to when a participant withholds their ideas, concerns and opinions which could be valuable to others or thoughts that they wish they could express (Morrison and Milliken, 2003). Earlier studies noted that features like anonymity allows participants to be true for themselves in contributing OS (Amrollahi and Rowlands, 2017; Mount et al., 2023). This study further highlighted that participants with anonymity could express their opinions without fear of confrontation, bias and judgement from others, allowing more radical ideas to overcome the communication barriers and to be heard.

Participants can also feel supported with technology use when the tool meeting their specific needs in OS making. For instance, visualisation of strategic content by diagrams, pictures, tables allow participants to read and interpret the same strategic context in the same way, this beneficial to participants who may encounter language difficulties to understand the strategy content and foster better understanding in strategy making among the group of participants, similar to earlier studies addressed that visualisation through IT can increase sense-making among the groups in strategy

making (Platts and Tan, 2004). Further, earlier studies (Motschnig and Hagelkruys, 2020.) have suggested that user interfaces must be flexible to accommodate for customisation that are specific to diverse user needs. This study elaborated on this idea and highlighted that tailoring strategy process through a diverse range of digital features, such as modifications in strategy templates, dashboard icon, and text form, supports participants to engage with and contribute to OS visually in the way that suit their individual preferences. This customisation, in turn, fosters a greater sense of comfort, ultimately enhancing participants' willingness to share their strategic views.

While technology inherently affords participants the opportunity to experience inclusion by catering to their individual needs, it is imperative for managers to proactively provide additional support to facilitate a sense of inclusivity in the process of OS. This study posits that managers can aid participants in acquiring proficiency with the tool use prior to the commencement of OS sessions. This proactive approach proves participants, especially those who lack technical competence, notably the elderly individuals. This study suggests that managers can effectively facilitate online training workshops aimed at elucidating tool utilisation or offer personalised one-to-one support to boost participants confidence in handling these tools. Such practices are instrumental in bridging technological barriers, therefore, enabling all participants to contribute to the OS seamlessly.

#### **IT-enabled valuing mechanism**

The use of technology also affords each participant an opportunity to exert influence on the OS outcomes by equitably appraising their strategic contributions. Prior studies have demonstrated that empowering participants with strategy making authority through technological means, such as via the mechanisms of commenting and voting on each other's ideas (Dobusch and Kapeller, 2018). This study elucidates that, in order to value participants' strategic contributions, managers can leverage a diverse array of digital tools to recognise, consider and incorporate participants' strategic input individually. Thus, in turn, serves to fortify participants' sense of ownership over their strategic inputs and fostering inclusive OS making.

The study further highlighted that the use of technology in fostering inclusion should not only focus on providing equality but also to promote equity among participants in the process of OS. It is evident from our findings that merely providing each participant with an opportunity to contribute does not guarantee that their contribution will be adequately acknowledged and incorporated into the final strategic outcomes. We argue that even in democratic voting process, like referendum, everyone votes at the same time, contributions from minority groups may still be overshadowed and marginalised by those of the majority groups in the organisation. Therefore, it is imperative to accord extra weight to the perspectives of underrepresented groups to ensure equity in their recognition and valuation.

#### **Overall contribution to OS literature**

This study makes several contributions to the existing literature on OS. Firstly, it extends the current understanding of the inclusiveness dimension in OS. While a limited number of studies (Adobor, 2020; Yakis-Douglas et al., 2017) posit that inclusiveness entails the broadening of both internal and external participation in OS activities, this study extends the current idea, from an inclusion perspective, that the composition of the participant pool for OS episodes should adhere to diversity

requirements. In this context, inclusiveness is construed as not only encompassing all potential participants but also incorporating individuals whose diverse characteristics can contribute unique expertise to OS and are historically underrepresented in the organization. Despite few studies (e.g., Malhotra et al., 2017) in OS have alluded the role of diversity in participants' pool formation, this study emphasizes the reciprocal relationship between inclusion and diversity, and fostering inclusion requires effective diversity management. The contention here is that greater inclusion results in a more diverse composition of the participant pool, from which diverse inputs and ideas are synthesized.

Secondly, this study underscores the pivotal role of technology in fostering inclusion in OS through the introduction of three mechanisms and their corresponding components. While existing studies have hinted at the role of technology in welcoming participants by broadcasting strategic information to inform them to join the OS (Morton et al., 2020), supporting individuals in expressing their authentic selves through anonymity (Amrollahi and Rowlands, 2017; Mount et al., 2023), and valuing participants by decentralizing decision-making powers through voting and commenting (Dobusch and Kapeller, 2018), our research contributes to a more integrative understanding of technology's role in fostering inclusion. Specifically, this study expands upon the components associated with each mechanism, providing more detailed explanation of technology use in fostering inclusion in OS.

Thirdly, the model generated from these three mechanisms presents an intertwined relationship among them, indicating that technology use for inclusion practices should address all three, supplementing the existing OS studies that often only focus one of these mechanisms (i.e., Amrollahi and Rowlands, 2017). Such intertwined relationship confirms that inclusion management is an ongoing and dynamic process (Ferdman, 2017), which requires managers to consider variety of aspects (i.e. corresponding components) when fostering inclusion.

### **Overall contribution to inclusion literature**

This study also makes contribution to the broader literature on diversity and inclusion (D&I). As previously mentioned, it offers a distinctive perspective on how the utilization of technology can foster inclusion. While existing studies in D&I have addressed the role of technology in social inclusion, particularly with respect to marginalized groups such as refugees and individuals with disabilities (Andrade and Doolin, 2016; Manzoor and Vimarlund, 2018), our research contributes novel insights to the management of inclusion within organizational settings. Unlike prior works that often focus on identifiable disadvantaged groups, our study explores instances in which marginalized groups, such as workplace discrimination in gender, sex, and age (Dipboye and Colella, 2013) or even individuals who are unwilling to contribute (Adobor, 2019) may not be readily discernible, consequently being unintentionally excluded from accessing vital organizational considerations. In such contexts, the strategic use of technology emerges as a means to afford equal opportunities for all participants within the organization, facilitating their access and influence in decision-making processes. Technology use in fostering inclusion in organisational settings, therefore, framed as a good force that changes exclusionary practices that have dominated organisations (Adamson, 2021).

### **Implication for practicing managers**

Inclusion management is imperative, yet a challenging and complex practise for the organisations. We hope conducting this research can help managers prohibit doing inclusive practices effortlessly and prevent issues like tokensim (Guldiken et al., 2019). Thus, our model of three intertwined mechanisms, along with their corresponding components provides explicit guidance to managers on implementing inclusive practices with the support of technology.

## **Conclusion**

This study provides an integrative picture of technology use in fostering inclusion in OS. By adopting a qualitative approach with inductive reasoning, we uncovered the three mechanisms of welcoming, supporting and valuing participants in OS. The study further explained 10 corresponding components of the three mechanisms to present explicit views of what managerial practices can be conducted through relevant digital tools and features use to foster inclusion. We hope by conducting this study not only contributes to existing OS literature by deepen our understanding how technologies use can foster inclusion but also to provide practical implications for managers in technology use for broader inclusive practice within the organisation.

## Reference

- Adamson, M., Kelan, E., Lewis, P., Śliwa, M. and Rumens, N., 2021. Introduction: Critically interrogating inclusion in organisations. *Organization*, 28(2), pp.211-227
- Ackermann, F., 1996. Participants' perceptions on the role of facilitators using group decision support systems. *Group decision and negotiation*, 5, pp.93-112.
- Adobor, H., 2020. Open strategy: role of organizational democracy. *Journal of Strategy and Management*, 13(2), pp.310-331.
- Amrollahi, A. and Rowlands, B., 2017. Collaborative open strategic planning: a method and case study. *Information technology & people*, 30(4), pp.832-852.
- Andrade, A.D. and Doolin, B., 2016. Information and communication technology and the social inclusion of refugees. *Mis Quarterly*, 40(2), pp.405-416.
- Baptista, J., Wilson, A.D., Galliers, R.D. and Byngghall, S., 2017. Social media and the emergence of reflexivity as a new capability for open strategy. *Long Range Planning*, 50(3), pp.322-336.
- Brimhall, K.C. and Mor Barak, M.E., 2018. The critical role of workplace inclusion in fostering innovation, job satisfaction, and quality of care in a diverse human service organization. *Human Service Organizations: Management, Leadership & Governance*, 42(5), pp.474-492.
- Birks, D.F., Fernandez, W., Levina, N. and Nasirin, S., 2013. Grounded theory method in information systems research: its nature, diversity and opportunities. *European Journal of Information Systems*, 22(1), pp.1-8.
- Combs, G.M., Milosevic, I. and Bilimoria, D., 2019. Introduction to the special topic forum: Critical discourse: Envisioning the place and future of diversity and inclusion in organizations. *Journal of Leadership & Organizational Studies*, 26(3), pp.277-286.
- Dipboye, R.L. and Colella, A. eds., 2013. *Discrimination at work: The psychological and organizational bases*. Psychology Press.
- Dobusch, L., 2021. The inclusivity of inclusion approaches: A relational perspective on inclusion and exclusion in organizations. *Gender, Work & Organization*, 28(1), pp.379-396.
- Dobusch, L. and Kapeller, J., 2018. Open strategy-making with crowds and communities: Comparing Wikimedia and Creative Commons. *Long range planning*, 51(4), pp.561-579.
- Ferdman, B.M. and Deane, B., 2014. Diversity at work: The practice of inclusion.
- Ferdman, B.M., 2017. Paradoxes of inclusion: Understanding and managing the tensions of diversity and multiculturalism. *The Journal of Applied Behavioral Science*, 53(2), pp.235-263.
- Glaser, B., & Strauss, A. (1967). *The Discovery of Grounded Theory Strategies for Qualitative Research*. Mill Valley, CA Sociology Press.
- Gregory, R.W., Kaganer, E., Henfridsson, O. and Ruch, T.J., 2018. IT consumerization and the transformation of IT governance. *Mis Quarterly*, 42(4), pp.1225-1253.
- Goldman, G.A. and Kruger, J., 2021. Viewing 'open strategy' as a viable approach to the management of strategy. *Journal of Economic and Financial Sciences*, 14(1), p.579.
- Guldiken, O., Mallon, M.R., Fainshmidt, S., Judge, W.Q. and Clark, C.E., 2019. Beyond tokenism: How strategic leaders influence more meaningful gender diversity on boards of directors. *Strategic Management Journal*, 40(12), pp.2024-2046.
- Inclusion at Work (2022). Findings from the inclusion and diversity survey. CIPD.
- Kaplan, S., 2011. Strategy and PowerPoint: An inquiry into the epistemic culture and machinery of strategy making. *Organization Science*, 22(2), pp.320-346.
- Konrad, A.M., Kramer, V. and Erkut, S., 2008. The impact of three or more women on corporate boards. *Organizational dynamics*, 37(2), pp.145-164.
- Langley, A., 1999. Strategies for theorizing from process data. *Academy of Management review*, 24(4), pp.691-710.

- Lirio, P., Lee, M.D., Williams, M.L., Haugen, L.K. and Kossek, E.E., 2008. The inclusion challenge with reduced-load professionals: The role of the manager. *Human Resource Management: Published in Cooperation with the School of Business Administration, The University of Michigan and in alliance with the Society of Human Resources Management*, 47(3), pp.443-461.
- Lorenz, F. and Buchwald, A., 2023. A perfect match or an arranged marriage? How chief digital officers and chief information officers perceive their relationship: a dyadic research design. *European Journal of Information Systems*, 32(3), pp.372-389.
- Malhotra, A., Majchrzak, A. and Niemiec, R.M., 2017. Using public crowds for open strategy formulation: mitigating the risks of knowledge gaps. *Long range planning*, 50(3), pp.397-410.
- Manzoor, M. and Vimarlund, V., 2018. Digital technologies for social inclusion of individuals with disabilities. *Health and technology*, 8, pp.377-390.
- Milliken, F.J., Morrison, E.W. and Hewlin, P.F., 2003. An exploratory study of employee silence: Issues that employees don't communicate upward and why. *Journal of management studies*, 40(6), pp.1453-1476.
- Mor Barak, M.E., Cherin, D.A. and Berkman, S., 1998. Organizational and personal dimensions in diversity climate: Ethnic and gender differences in employee perceptions. *The Journal of Applied Behavioral Science*, 34(1), pp.82-104.
- Morton, J.O.S.H., Wilson, A.L.E.X., Galliers, R.D. and Marabelli, M.A.R.C.O., 2019. Open strategy and information technology. *Cambridge Handbook of Open Strategy: Cambridge*, pp.169-185.
- Morton, J., Wilson, A.D. and Cooke, L., 2020. The digital work of strategists: Using open strategy for organizational transformation. *The Journal of Strategic Information Systems*, 29(2), p.101613.
- Motschnig, R. and Hagelkruys, D., 2020. Inclusion of Users with Special Needs in the Human-Centered Design of a Web-Portal. In *Accessibility and Diversity in Education: Breakthroughs in Research and Practice* (pp. 272-290). IGI Global.
- Mount, M.P., Clegg, S.R. and Pitsis, T.S., 2020. Conceptualizing the de-materializing characteristics of internal inclusion in crowdsourced open strategy. *Long Range Planning*, 53(5), p.101986.
- Nishii, L.H., 2013. The benefits of climate for inclusion for gender-diverse groups. *Academy of Management journal*, 56(6), pp.1754-1774.
- O'Connor, C. and Joffe, H., 2020. Intercoder reliability in qualitative research: debates and practical guidelines. *International journal of qualitative methods*, 19, p.1609406919899220.
- Ohunakin, F., Adeniji, A., Ogunnaike, O.O., Igbadume, F. and Akintayo, D.I., 2019. The effects of diversity management and inclusion on organisational outcomes: A case of multinational corporation. *Business: Theory and Practice*, 20(3), pp.93-102.
- Patrick, H.A. and Kumar, V.R., 2012. Managing workplace diversity: Issues and challenges. *Sage Open*, 2(2), p.2158244012444615.
- Patton, M.Q., 2002. Two decades of developments in qualitative inquiry: A personal, experiential perspective. *Qualitative social work*, 1(3), pp.261-283.
- Platts, K. and Tan, K.H., 2004. Strategy visualisation: knowing, understanding, and formulating. *Management decision*, 42(5), pp.667-676.
- Quick, K.S. and Feldman, M.S., 2011. Distinguishing participation and inclusion. *Journal of planning education and research*, 31(3), pp.272-290.
- Strauss, A. and Corbin, J., 1998. Basics of qualitative research techniques.
- de Souza, N.P.R. and Gama, K., 2020. Diversity and Inclusion: Culture and perception in information technology companies. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, 15(4), pp.352-361.
- Roberson, Q.M., 2006. Disentangling the meanings of diversity and inclusion in organizations. *Group & Organization Management*, 31(2), pp.212-236.
- Schlagwein, D., Conboy, K., Feller, J., Leimeister, J.M. and Morgan, L., 2017. "Openness" with and without Information Technology: a framework and a brief history. *Journal of Information Technology*, 32(4), pp.297-305.
- Shore, L.M., Randel, A.E., Chung, B.G., Dean, M.A., Holcombe Ehrhart, K. and Singh, G., 2011. Inclusion and diversity

in work groups: A review and model for future research. *Journal of management*, 37(4), pp.1262-1289.

Tavakoli, A., Schlagwein, D. and Schoder, D., 2015. Open strategy: Consolidated definition and processual conceptualization.

Yakis-Douglas, B., Angwin, D., Ahn, K. and Meadows, M., 2017. Opening M&A strategy to investors: predictors and outcomes of transparency during organisational transition. *Long Range Planning*, 50(3), pp.411-422.

Whittington, R., Cailluet, L. and Yakis-Douglas, B., 2011. Opening strategy: Evolution of a precarious profession. *British journal of management*, 22(3), pp.531-544.

Wirtz, B.W., Daiser, P. and Mermann, M., 2018. Social media as a leverage strategy for open government: an exploratory study. *International Journal of Public Administration*, 41(8), pp.590-603

## Appendix A-Anonymised List of Interviewees

Interviewee	Role	Gender	Interview length (approximately)	Industry	Number Of strategy episode participants	Country of the organisation	Platforms used for OS
P1	Talent manager	F	30mins	Transportation	Approx.1000	US	ThoughtExchange
P2	Digital team leader	F	30mins	Manufacturing	Approx.30	US	Miro
P3	Director of design performance	F	30mins	Construction	Approx.24	US	Miro
P4	Strategy director	M	20 mins	Services	Approx.25	US	Mural
P5	HR manager	F	50 mins	Services	Approx.144	US	Leapsome
P6	Strategy manager	M	35 mins	Services	Approx.144	US	Leapsome
P7	Marketing manager	F	35 mins	Services	Approx.144	US	Leapsome
P8	Client solution manager	F	55 minus	Services	Approx.144	US	Leapsome
P9	Business consultant	M	First interview 40 minus	Services	Approx. 30	US	Mural
			Second interview 20 minutes				
P10	HR director	F	35 mins	Services	Approx.900	Canada	ThoughtExchange
P11	Research and product specialist	M	38 mins	Services	Approx.80	Canada	ThoughtExchange
P12	Managing director	M	60mins	Services	Approx.8	Germany	Mural
P13	Strategy advisor	F	30mins	Services	Approx.7	UK	Spatial
P14	Marketing manager	M	40 mins	Services	Approx.50	UK	Digiworks
P15	Strategy manager	M	50 mins	Services	Approx.20	UK	Mural
P16	Senior director of digital transformation	F	40 mins	Manufacturing	Approx.30	UK	Meta workplace
P17	Strategy manager	M	45 mins	Manufacturing	Approx.200	US	Microsoft Teams
P18	Paralegal manager	F	45 mins	Services	Approx.7	US	Microsoft Teams
P19	Strategy manager	M	40 mins	Services	Approx.200	US	Microsoft Teams
P20	Strategy manager	M	30 mins	Mining	Approx.150	UK	Organisational intranet
P21	Business consultant	M	30 mins	Services	Approx.45	Canada	Miro



P22	Behavioural scientist	M	30 mins	Services	Approx.6	UK	Miro
P23	Project portofolio manager	M	30 mins	Manufacturing	Approx.16	Mexico	Miro
P24	Strtaegy manager	M	35mins	Services	Approx.90	UK	Mural
P25	Agile strategy manager	M	35	Services	Approx.8	El Salvador	Miro
P26	Customer success exectuive	F	45 mins	Services	Approx.40	Spin	Mural
P27	Senior strategy manager	M	35 mins	Services	Approx.32	US	Mural
P28	Strategy manager	M	35 mins	Services	Approx.35	India	Miro
P29	IT resarcher	F	30 mins	Services	Approx.85	UK	Mural
P30	Regional director	F	30 mins	Services	Approx.2000	US	ThoughtExchange
P31	Strategy advisor	M	35 mins	Services	Approx.16	Australia	Mural
P32	Chief culture director	M	35 mins	Services	Approx.8	China	Mural
P33	Marketing manager	F	40	Manufacturing	Approx.25	US	Mural
P34	Operational manager	M	30 mins	Services	Approx.11	Canada	Mural
P35	Strategy manager	M	30 mins	Services	Approx.13	US	Miro
P36	Consumer research engineer	M	40 mins	Manufacturing	Approx.65	US	Miro
P37	Product manager	M	35mins	Services	Approx.18	US	Mural
P38	Product manager	M	45 mins	Services	Approx.6	US	Miro
P39	Head of learning innovation	M	30 mins	Services	Approx.18	UK	Mural
P40	Business analyst	M	30 mins	Services	Approx.60	France	Mural
P41	D&I specialist	M	30mins	Services	Approx.200	UK	Microsoft Teams
P42	Strategy manager	F	45 mins	Services	Approx.24	US	Mural
P43	Innovation manager	F	30mins	Services	Approx.10	Australia	Mural
P44	Strategy manager	F	40 mins	Services	Approx.25	US	Mural
P45	Product manager	F	30 mins	Manufacturing	Approx.165	Milan	Mural

# Uncovering the Landscape of Ethical Artificial Intelligence: A Literature Review

**Itoero Abraham (PhD Researcher)**  
*Lancaster University, UK*

**Dr Ruilin Zhu**  
*Lancaster University, UK*

**Dr Mahsa Honary**  
*Lancaster University, UK*

*Research In Progress*

## **Abstract**

*The proliferation of Artificial Intelligence (AI) systems with adaptive features has provided businesses and Society numerous benefits. Nevertheless, a growing body of literature has discussed the ethical issues these systems pose and the responses of human agents. Through a systematic literature review, we analyse and synthesise the existing literature on AI ethics and suggest new directions for further research in this area. There is a need to investigate the AI developers' stance on ethics, as the lack of transparency and accountability in the AI domain has sparked new debates and tension between AI performance, explainability and accountability. The deployment of AI systems affects Human role identities and reputations. Consequently, there is an increasing demand for explainable AI, interdisciplinary co-design processes, and ethical codes of conduct, and there is an urgent need to identify methods of accelerating AI developments without compromising human rights of autonomy and agency.*

**Keywords:** Artificial Intelligence ethics, Machine Learning, Machine autonomy, Explainable AI,

## 1. Introduction

The proliferation of Artificial Intelligence (AI) systems with learning capabilities and adaptive behaviours has enabled businesses to gain a competitive edge and provided Society with access to more comprehensive information and services at a rapid rate (Qin et al., 2020; Abràmoff et al., 2020). As AI algorithms become increasingly commonplace, making complex decisions that have a significant impact on human life (Qin et al., 2020; Trites, 2019; Winfield et al., 2019), there is a growing discussion surrounding the implications of Ethics and human autonomy (Kordzadeh & Ghasemaghahi, 2021; Qin et al., 2020; Winfield et al., 2019). Despite the demand from government and regulatory bodies for the development of AI to prioritise the welfare of Society (Rinta-Kahila et al., 2021), the detrimental effects of AI-driven decision-making have caused financial losses and distress to human agencies (Rinta-Kahila et al., 2021). AI systems in our societies are characterised by numerous black-boxes, where the inner workings of mechanisms, robotics, and automation enabled by algorithms remain unknown (Innerarity, 2021), thus violating ethical principles or diminishing people's rights and dignity and have a widespread effect on humans and societies (Rességuier & Rodrigues, 2020).

In recent years, AI has had a profound impact on many sectors, such as healthcare, agriculture, automobiles, and defence, due to its ability to operate independently of human intervention (Khan et al., 2021; Qin et al., 2020; Trites, 2019; Winfield et al., 2019). As AI continues to evolve and expand, it increasingly becomes embedded in every industry and Society (Khan et al., 2021), thereby bringing ethical concerns (Berens & Grote, 2020; Katzenbach & Ulbricht; Winfield et al., 2019). However, the study of AI ethics and its application in many practical fields and regions of the world is still in the early stages (Khan et al., 2021; Kordzadeh & Ghasemaghahi, 2021; Koshiyama et al., 2022) and will continue advancing as AI increases its capabilities to handle more critical tasks (Khan et al., 2021). This study focuses on AI systems with self-learning capabilities, otherwise described as black-box algorithms (Topol, 2019). Given the scant nature of the available efforts to examine the AI systems and their resultant effects on human autonomy and agency (Trites, 2019), we seek to aid future studies in AI ethics by answering the following two research questions through a systematic literature review:

*(1) How does the deployment of AI impact human autonomy? And*

*(2) How do humans react to the ethical challenges of using AI systems?*

This study views AI ethics from the rights perspective because AI influences people's rights to privacy, dignity and self-governance (Laitinen & Sahlgren, 2021; Lanzing, 2019; Riley & Bos, 2022). The theory of rights maintains that people have a unique moral position, allowing them to autonomously self-govern themselves, otherwise viewed as self-ownership (Aertsen, 2019). More so, people are the equitable and moral owners of themselves and their powers (Hoag, 1991). However, people may willingly relinquish their power to other people or artificial persons, such as AI, if they have the power

to do so (Treleaven et al., 2019). Importantly, ethics of rights emphasises the right to personal physical and psychological integrity and human autonomy as central to morality (Dewan, 2022; Hoag, 1991).

## 2. Foundations

Kolkman and Kemper (2019) suggested that AI algorithms, which drive systems, are viewed as blank slates and socio-technical assemblages that cannot be held accountable for ethical violations without a critical audience. Despite the limited responsibility of developers in framing the algorithm, Martin (2019) argues that algorithms are not neutral but rather value-laden, as they can lead to moral consequences. As a result, it is necessary to consider the AI algorithm's application and effect on people's right to self-determination holistically. Given the significant role of AI algorithms in ethical decisions (Martin, 2019), it is proposed that the developers of such algorithms should be held responsible for the moral consequences that arise from their use (Kolkman & Kemper, 2019; Martin, 2019) if they are designed to prevent individuals from assuming responsibility for a given decision. Developers have acknowledged the increasing ethical concerns arising from AI developments, yet they have not taken sufficient steps to address them (Kolkman & Kemper, 2019; Martin, 2019; Winfield et al., 2019). Kordzadeh and Ghasemaghaei (2021) have called for developers to incorporate suitable procedures into the system development lifecycle, such as requirements gathering, usability and all testing processes, to gain a more comprehensive understanding of how people in specific contexts perceive and react to ethical issues that may arise from AI implementations.

Hagendorff (2020) postulates that AI ethics is void of reinforcement mechanisms due to the lack of consequences upon failure to uphold various codes of ethics in AI developments and integration. Experiments have suggested that the significance of ethics in the AI development process has been given an alarming amount of attention in the AI community (Kim & Routledge, 2021; Sutrop, 2019). Furthermore, there is a lack of in-depth understanding of the broader societal effects and the substitution of ethical values for economic motivation (Hagendorff, 2020). Additionally, current institutions are failing to guide Society in a direction that produces a common good, as the challenges posed by creating favourable social policy based on AI recommendations are rising rapidly in the era of AI (Lahsen, 2020). Moreover, the political use of AI, such as machine learning, neural networks, deep learning, and robotics, is perceived as a significant risk, and the objectives for which AI systems are developed and integrated often conflict with fundamental rights or social values (Hagendorff, 2020; Lahsen, 2020).

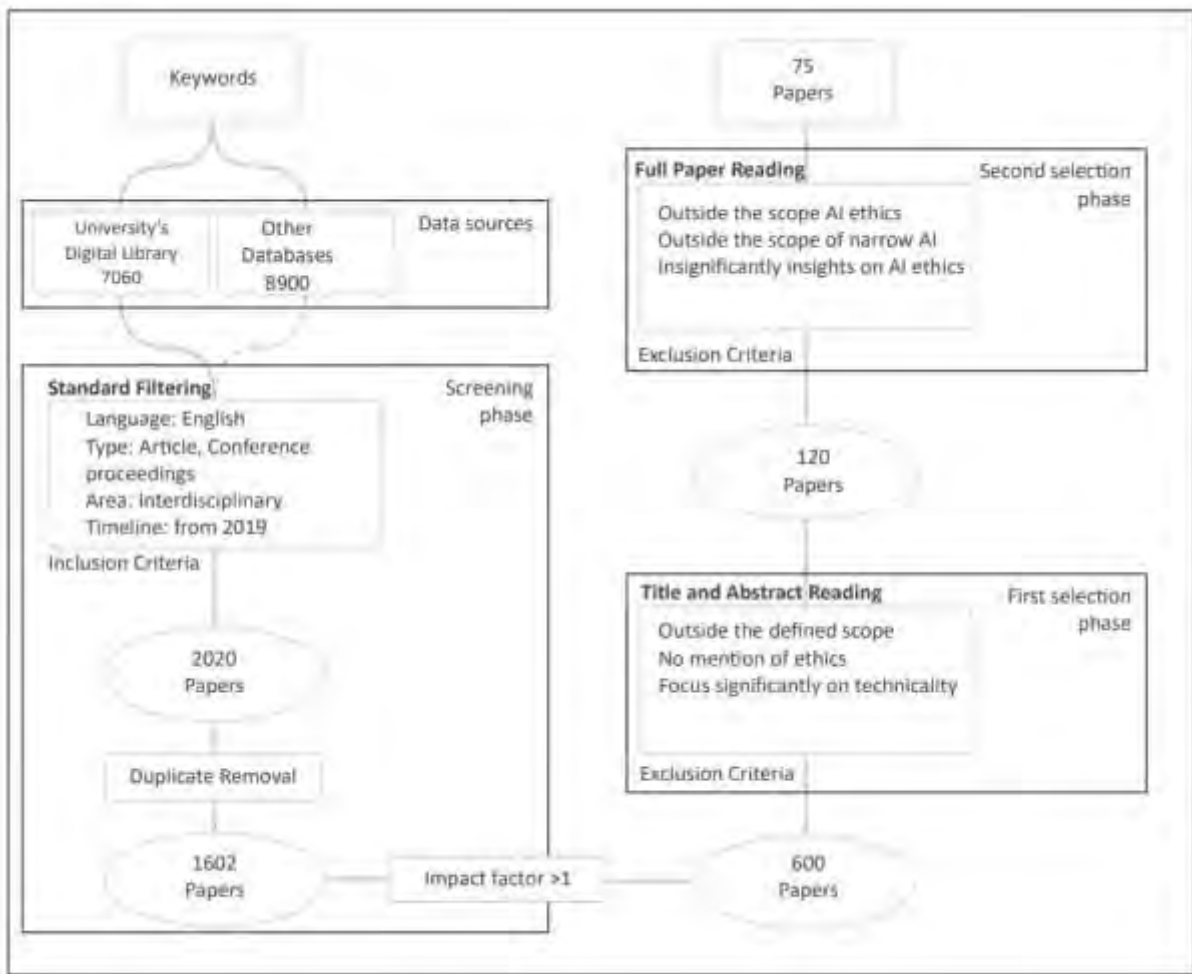
## 3. Methodology

We followed the principles outlined by Templier and Paré (2015) to analyse and synthesise the existing literature on AI ethics and propose new directions for further research. The principle applied included (1) reviewing the current state of AI ethics literature to identify and articulate the issue, (2) defining and applying a literature search strategy, (3) examining our search process for article inclusion, (4) assessing the quality of the articles concerning the information system (IS) field, (5) extracting potentially

relevant articles from a variety of databases across multiple disciplines, as AI ethics is highly interdisciplinary and rapidly evolving (Koshiyama et al., 2022; Qin et al., 2020; Trites, 2019; Winfield et al., 2019), and finally, (6) analysing and synthesising the study findings. We recognised that, due to the ubiquity of AI (Trites, 2019) and its profound socioeconomic impacts (Durán & Jongsma, 2021; Gill, 2019; Trites, 2019), excluding non-IS-related databases from the search would adversely affect the study outcome. Thus, we included articles with AI ethics-related themes from organisational, technological, behavioural, and social spheres for comprehensiveness.

We employed a comprehensive approach to our literature search, utilising the University's digital library's search engine and Elsevier's Scopus, Web of Science, IEEE Xplore Digital Library, ACM, INFORM and ScienceDirect. We included articles from the proceedings of the International Conference on Information Systems (ICIS) and SSRN as they offer useful articles on AI and Ethics and relevant insights. We retrieved articles published in conference proceedings and scholarly journals between August 2019 and August 2022, as the concept of AI ethics has recently gained traction in both the industrial and academic spheres (Kordzadeh & Ghasemaghaei, 2021; Koshiyama et al., 2022). Additionally, we conducted a backward and forward search to identify related articles not captured in the initial search to ensure comprehensiveness. After screening, articles that focused mainly on the technical implementation or deployment of AI technologies were excluded, as well as any duplicates from multiple sources.

We conducted an iterative assessment of the citations of related articles to broaden the scope of our search. Our final keywords included "Artificial Intelligence Ethics", "Human-Centred AI", "Algorithmic Bias", "Accountable AI", "Algorithmic Fairness", "Responsible AI", "Trustworthy AI", "AI Autonomy". This resulted in 75 articles were chosen. Figure 1 below presents an overview of the selection process.



**Figure 1. Stages of the article selection process.**

## 4. Preliminary Finding

### 4.1 Effect of AI Deployment on Human Autonomy

The use of AI systems through hyper-personalisation has been met with ethical dilemmas, such as privacy concerns and the potential for decreased human rights and liberties due to wrongly trained data (Vempati et al., 2020; Hernández-Ramírez, 2019; Floridi & Cows, 2019). This is because businesses employ omnichannel to explore customer behavioural patterns in real-time (Vempati et al., 2020), which helps them generate highly customised user experiences (Maddodi, 2021). Such a practice has been viewed as a threat to human autonomy (Laitinen & Sahlgren, 2021) and identity (Adamson et al., 2021), as people do not have control over how their personal or behavioural data is being modelled, interpreted, and used (Hernández-Ramírez, 2019). This can lead to them being displaced from their behaviours and the knowledge their behavioural data provide (Naughton 2019).

Implementing automation technologies, such as AI, has been seen to optimise workflows to overhead and increase efficiency and revenue (Madakam et al., 2019; Vempati et al., 2020). However, this process

has been observed to limit the decision-making opportunities of individuals (Calvo et al., 2020; Laitinen & Sahlgren, 2021; Väänänen et al., 2021) and potentially disrupt their role identities and reputations (Floridi & Cowls, 2019; Grønsund & Aanestad, 2020; Mayer et al., 2020; Strich et al., 2021). This can cause the loss of human autonomy (Sobczak, 2022; Laitinen & Sahlgren, 2021) and identity crises (Floridi & Cowls, 2019).

A recent study in the medical field has revealed that the use of medical costs as a substitute for patients' health provisions has resulted in unacceptable racial bias in the allocation of healthcare resources. This is because patients from a racially disadvantaged background were mistakenly viewed as having a lower risk than those from a racially advantaged background due to their lower incurred costs for a given health risk status (Obermeyer et al., 2019). This demonstrates that the existing bias can lead to the isolation of individuals with protected characteristics (Hagendorff, 2020; Sarker, 2021; Winfield et al., 2019). To ensure that AI applications uphold fundamental human rights and liberties such as dignity, privacy and the right of self-determination, developers must bridge the knowledge gap between themselves and the users to understand the ethical implications of AI systems (Treleaven et al., 2019; Mittelstadt, 2019). Additionally, there is a need to address the tension between machine and human autonomy (Calvo et al., 2020; Väänänen et al., 2021). This has led to a demand to reformulate how autonomous Technology weighs human welfare and ethical considerations, evaluating if the innovation or services align with end-user values or cause harm before being incorporated into the public or private domain (Adamson et al., 2019).

Figure 2 outlines how the implementation of AI systems influences human autonomy.

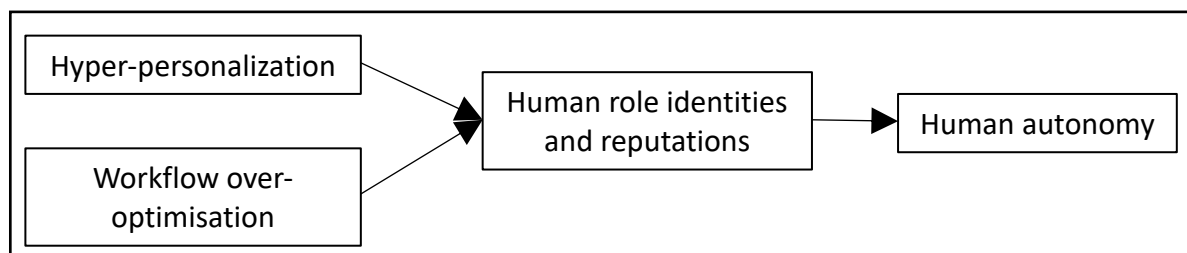


Figure 2. AI systems influence human autonomy.

#### 4.2. Humans respond to the ethical challenges of using AI

The detrimental effects of unethical characteristics in AI systems on the reputations of individuals and professional bodies (Floridi & Cowls, 2019; Grønsund & Aanestad, 2020; Mayer et al., 2020; Strich et al., 2021) have caused a decrease in public trust in the developers of these systems (Adamson et al., 2019). The failure of these developers to address these ethical issues in an accountable and transparent manner further exacerbates this lack of trust (Abràmoff et al., 2020; Winfield et al., 2019). Consequently, there is an increased need for AI systems to be comprehensible to humans (Colaner, 2021; Miller, 2019; Sheth et al., 2021), as the absence of explainability results in an opaque black box

that conceals the inner workings of the algorithms (Colaner, 2021). This has caused a new wave of research in XAI (Alizadeh et al., 2021), which strives to create AI systems that are more transparent and responsible (Alizadeh et al., 2021; Miller, 2019; Rieg et al., 2020).

Given the fact that AI developers may not be able to understand the full scope of the ethical implications of their systems on people (Kazim & Koshiyama, 2021), there has been a shift in the public interest toward interdisciplinary design processes (co-design) for moderating ethical challenges (Colombino et al., 2021; Kazim & Koshiyama, 2021). Although the enforcement of ethical codes of practice for AI remains uncertain (Hagendorff, 2020), many individuals and professional bodies are striving to implement such codes of conduct to protect individuals from potential harm (Hagendorff, 2020; Sarker, 2021). Moreover, AI ethics necessitates a continuous process of examining the ethical implications of AI on humans in order to formulate effective strategies to mitigate the overall impacts (Rességuier & Rodrigues, 2020).

Figure 3 integrates how AI systems affect human autonomy and how people respond to the ethical challenges of using AI systems. Instead of asking for a direct rise in human autonomy, people are calling for the explainability of AI algorithms, an interdisciplinary co-design process to moderate AI ethical challenges, and AI ethical codes of conduct to protect individuals from potential harm. It remains to be seen if this indirect approach will eventually increase human autonomy.

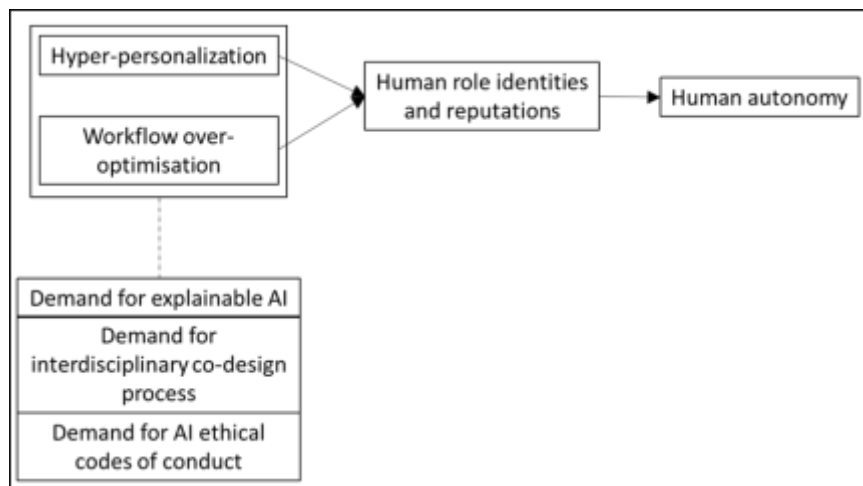


Figure 3. How AI systems affect human autonomy and how people respond to AI systems.

## 5. Discussion and next steps

While it is perceived that the opaqueness of the autonomous or AI systems is shielding their unethical outcomes from criticism (Trites, 2019; Durán & Jongsma, 2021), there is an increasing demand for AI developers to be held accountable for the shortcomings of the systems they develop (Martin, 2019; Kolkman & Kemper, 2019, Winfield et al., 2019) given that the systems may exhibit or re-emphasise the biased inclinations of their creators through their outputs (Kolkman & Kemper, 2019). Unfortunately, no regulatory framework justifies what represents a good or bad AI developer



(Mittelstadt, 2019). In addition, it is progressively more challenging for AI developers to remain ethically conscious without losing their jobs due to their management and other influential stakeholders' interests (Mittelstadt, 2019). Thus suggesting the need to explore the developers' viewpoints

AI systems may not form a different society or disengage themselves from human supervision. However, their evolutionary agency would still be shaped by humans who relinquish control and delegate decision-making to AI systems (Rozendaal et al., 2019; Kaptelinin, 2022). This suggests that AI's unethical actions originate from, and could be reflections of, specific forces and tensions in Society (Kaptelinin, 2022) and the people who create them (Geis et al., 2019). There is a burning need for research in AI to understand the interaction between resources, power relations, and the social norms which influence the development, integration, and use of AI systems (Kaptelinin, 2022). It is commonly agreed that expanded studies would provide new knowledge on dealing with AI technologies and understanding social phenomena and changes that are prerequisites to harnessing AI's power, as well as unravelling the root causes of unethical outcomes (Durán and Jongsma, 2021; Carabantes, 2019; Rozendaal et al., 2019; Kaptelinin, 2022).

Though the literature heightened the call for AI developers to be accountable for unethical systems (Martin, 2019; Kolkman & Kemper, 2019; Winfield et al., 2019), no significant study had highlighted the developers' concerns and perceptions. Given that each implementation of AI may represent a distinctive set of circumstantial and ethical issues specific to its processes and environment (Calvo et al., 2020; Qin et al., 2020) and the lack of significant research and development in the developing world (in particular, Africa) (Oxford Insights, 2019; Nakalembe & Kerner, 2023), it is valuable to advance this study to explore AI developers' perception of ethics in the developing world in regards to AI ethical development to contribute new knowledge AI development while safeguarding the Society from technological harm.

## References

- Abràmoff, M.D., Tobey, D. & Char, D.S., 2020. Lessons Learned About Autonomous AI: Finding a Safe, Efficacious, and Ethical Path Through the Development Process. *American Journal of Ophthalmology*, 214, pp.134–142.
- Adamson, G., Havens, J.C. & Chatila, R., 2019. Designing a value-driven future for ethical autonomous and Intelligent Systems. *Proceedings of the IEEE*, 107(3), pp.518–525.
- Aertsen, J.A. (2019) 'Thomas Aquinas on the good: The relation between Metaphysics and Ethics', *Aquinas's Moral Theory*, pp. 235–253. doi:10.7591/9781501728365-012.
- Alizadeh, F., Stevens, G. & Esau, M., 2021. I don't know, is Ai also used in airbags? *I-com*, 20(1), pp.3–17.
- Berens, P. & Grote, T., 2020. On the ethics of algorithmic decision-making in healthcare. *On the ethics of algorithmic decision-making in healthcare*. Available at: <https://www.ethicalpsychology.com/2020/04/on-ethics-of-algorithmic-decision.html> [Accessed February 9, 2022].
- Calvo, R.A. et al., 2020. Supporting human autonomy in AI systems: A Framework for Ethical Enquiry. *Philosophical Studies Series*, pp.31–54.

- Carabantes, M., 2019. Black-Box Artificial Intelligence: An epistemological and critical analysis. *AI & SOCIETY*, 35(2), pp.309–317.
- Colaner, N., 2021. Is explainable artificial intelligence intrinsically valuable? *AI & SOCIETY*, 37(1), pp.231–238.
- Colombino, T. et al., 2021. Ethical design of a robot platform for disabled employees: Some practical methodological considerations. *Frontiers in Robotics and AI*, 8.
- Dewan, L. (2022) 'Chapter 2. wisdom as foundational ethical theory in St. Thomas Aquinas', *Wisdom, Law, and Virtue*, pp. 32–57. doi:10.1515/9780823293551-005.
- Durán, J.M. & Jongsma, K.R., 2021. Who is afraid of Black Box Algorithms? on the epistemological and ethical basis of trust in Medical Ai. *Journal of Medical Ethics*.
- Floridi, L. & Cows, J., 2019. A unified framework of five principles for AI in society. *SSRN Electronic Journal*.
- Geis, J.R. et al. (2019) 'Ethics of Artificial Intelligence in Radiology: Summary of the joint European and North American multisociety statement', *Radiology*, 293(2), pp. 436–440. doi:10.1148/radiol.2019191586. Gill, K.S., 2019. Dance of the artificial alignment and ethics. *AI & SOCIETY*, 35(1), pp.1–4.
- Grønsund, T. & Aanestad, M., 2020. Augmenting the algorithm: Emerging human-in-the-loop work configurations. *The Journal of Strategic Information Systems*, 29(2), p.101614.
- Hagendorff, T., 2020. The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, 30(1), pp.99–120.
- Hoag, R.W. (1991) 'Ethical theory and social issues', *Teaching Philosophy*, 14(4), pp. 476–479. doi:10.5840/teachphil199114476
- Innerarity, D., 2021. Making the Black Box Society Transparent - ai & society. *SpringerLink*. Available at: <https://link.springer.com/article/10.1007/s00146-020-01130-8> [Accessed February 7, 2022].
- Kaptelinin, V., 2022. The social production of Technological Autonomy. *Human–Computer Interaction*, 37(3), pp.256–258.
- Katzenbach, C. & Ulbricht, L., 2019. Algorithmic governance. *Internet Policy Review*, 8(4).
- Kazim, E. & Koshiyama, A.S., 2021. A high-level overview of AI ethics. *Patterns*, 2(9), p.100314.
- Khan, F.H., Pasha, M.A. & Masud, S., 2021. Advancements in microprocessor architecture for ubiquitous AI—an overview on history, evolution, and upcoming challenges in AI implementation. *Micromachines*, 12(6), p.665.
- Kim, T.W. & Routledge, B.R., 2021. Why a Right to an Explanation of Algorithmic Decision-Making Should Exist: A Trust-Based Approach. *Business Ethics Quarterly*, pp.1–28.
- Kolkman, D. & Kemper, J., 2019. Transparent to whom? No algorithmic accountability without a critical audience. *Taylor & Francis*. Available at: <https://www.tandfonline.com/doi/full/10.1080/1369118X.2018.1477967> [Accessed February 9, 2022].
- Kordzadeh, N. & Ghasemaghaei, M., 2021. Algorithmic bias: Review, synthesis, and future research directions. *European Journal of Information Systems*, pp.1–22.
- Koshiyama, A. et al., 2021. Towards algorithm auditing: A survey on managing legal, ethical and technological risks of AI, ML and associated algorithms. *SSRN Electronic Journal*.
- Koshiyama, A., Kazim, E. & Treleaven, P., 2022. Algorithm auditing: Managing the legal, ethical, and technological risks of Artificial Intelligence, machine learning, and associated algorithms. *Computer*, 55(4), pp.40–50.
- Lahsen, M., 2020. Should AI be designed to save us from ourselves? Artificial Intelligence for Sustainability. *IEEE Technology and Society Magazine*, 39(2), pp.60–67.

- Laitinen, A. & Sahlgren, O., 2021. AI systems and respect for human autonomy. *Frontiers in Artificial Intelligence*, 4.
- Lanzing, M., 2019. "Strongly recommended" revisiting decisional privacy to judge hypernudging in self-tracking technologies. *Philosophy & Technology*, 32(3), pp.549–568.
- Martin, K., 2019. Ethical implications and accountability of algorithms. *SpringerLink*. Available at: <https://link.springer.com/article/10.1007/s10551-018-3921-3> [Accessed February 9, 2022].
- Madakam, S., Holmukhe, R.M. and Jaiswal, D.K. (2019) The Future Digital Work Force: Robotic Process Automation (RPA), *JISTEM - Journal of Information Systems and Technology Management*.
- Maddodi, S. (2021) Artificial intelligence and hyper-personalisation for improving customer experience, SSRN. Available at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3805904](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3805904) (Accessed: October 11, 2022).
- Mayer, A.-S., Strich, F. & Fiedler, M., 2020. Unintended consequences of introducing AI systems for decision making. *MIS Quarterly Executive*, pp.239–257.
- Miller, T., 2019. Explanation in artificial intelligence: Insights from the Social Sciences. *Artificial Intelligence*, 267, pp.1–38.
- Mittelstadt, B., 2019. Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), pp.501–507.
- Nabi, M.K. (2019) "The Impact of Artificial Intelligence (AI) on Workforce in Emerging Economies," *Global Journal of Management And Business Research*, pp. 51–59.
- Obermeyer, Z. et al., 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), pp.447–453.
- Qin, F., Li, K. & Yan, J., 2020. Understanding user trust in artificial intelligence-based educational systems: Evidence from China. *British Journal of Educational Technology*, 51(5), pp.1693–1710.
- Rességuier, A. & Rodrigues, R., 2020. AI ethics should not remain toothless! A call to bring back the teeth of Ethics. *Big Data & Society*, 7(2), p.205395172094254.
- Rieg, T. et al., 2020. Demonstration of the potential of white-box machine learning approaches to gain insights from cardiovascular disease electrocardiograms. *PLOS ONE*, 15(12).
- Riley, S. & Bos, G., Human Dignity. *Internet encyclopedia of philosophy*. Available at: <https://iep.utm.edu/human-dignity/> [Accessed May 8, 2022].
- Rinta-Kahila, T. et al., 2021. Algorithmic decision-making and system destructiveness: A case of automatic debt recovery. *European Journal of Information Systems*, 31(3), pp.313–338.
- Rozendaal, M.C., Boon, B. & Kaptelinin, V., 2019. Objects with intent. *ACM Transactions on Computer-Human Interaction*, 26(4), pp.1–33.
- Sarker, I.H., 2021. Machine learning: Algorithms, real-world applications and Research Directions. *SN computer science*. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7983091/> [Accessed February 7, 2022].
- Sheth, A. et al., 2019. Shades of knowledge-infused learning for enhancing deep learning. *IEEE Internet Computing*, 23(6), pp.54–63.
- Sheth, A. et al., 2021. Knowledge-intensive language understanding for explainable ai. *IEEE Internet Computing*, 25(5), pp.19–24.
- Sobczak, A. (2022) Robotic Process Automation as a digital transformation tool for increasing organisational resilience in polish enterprises, MDPI. *Multidisciplinary*

- Digital Publishing Institute. Available at: <https://www.mdpi.com/2071-1050/14/3/1333> (Accessed: October 11, 2022).
- Strich, F., Mayer, A.-S. & Fiedler, M., 2021. What do I do in a world of artificial intelligence? investigating the impact of substitutive decision-making AI systems on employees' professional role identity. *Journal of the Association for Information Systems*, 22(2), pp.304–324.
- Sutrop, M., 2019. Should we trust artificial intelligence? *Trames. Journal of the Humanities and Social Sciences*, 23(4), p.499.
- Templier, M. & Paré, G., 2015. A framework for guiding and evaluating literature reviews. *Communications of the Association for Information Systems*, 37.
- Topol, E.J., 2019. High-performance medicine: The convergence of human and Artificial Intelligence. *Nature medicine*. Available at: <https://pubmed.ncbi.nlm.nih.gov/30617339/> [Accessed February 7, 2022].
- Treleaven, P., Barnett, J. & Koshiyama, A., 2019. Algorithms: Law and regulation. *Computer*, 52(2), pp.32–40.
- Trites, A., 2019. Black Box Ethics: How Algorithmic Decision-Making is Changing How We View Society and People: Advocating for the Right for Explanation and the Right to be Forgotten in Canada. *Global Media Journal*, 11(2), pp.18–30.
- Väänänen, K. et al., 2021. Editorial: Respecting human autonomy through human-centered AI. *Frontiers in Artificial Intelligence*, 4.
- Vempati, S. et al., 2020. Enabling hyper-personalisation: Automated AD creative generation and ranking for fashion e-commerce. *Lecture Notes in Social Networks*, pp.25–48.
- Winfield, A.F. et al., 2019. Machine Ethics: The Design and Governance of Ethical AI and Autonomous Systems [Scanning the Issue]. *Proceedings of the IEEE*, 107(3), pp.509–517.

# What are the Drivers of Non-Fungible Tokens (NFTs) Community Engagement?

*Research In progress*

## **Abstract**

*Despite the growing importance and popularity of non-fungible tokens (NFTs), little research has examined the importance of the NFT community to brands and marketing strategies. Furthermore, there is a notable lack of empirical studies that examine the various factors that influence NFT engagement. In an effort to address this research gap, we propose a conceptual model that allows for the examination of the effects of various social factors on NFT engagement. The research model therefore draws on social capital theory and social exchange theory. We also highlight how the data will be collected and the expected results. We expect this study to contribute to the current NFT and customer engagement literature and provide a better understanding of how certain social factors can encourage NFT users to engage within the community and how managers can use NFT communities more effectively.*

**Keywords:** Non-Fungible Token (NFTs), NFTs Community, users Engagement, Social Capital Theory, Social Exchange Theory.

## 1.0 Introduction

Non-fungible tokens (NFTs) represent digital credentials with an innate resistance to duplication (Popescu, 2021). Essentially, these tokens serve as unique digital identifiers that cannot be duplicated, substituted, or altered. They are securely recorded on a blockchain and can verify ownership and authenticity (Ali et al., 2023; Chandra, 2022). Originating from Ethereum's smart contracts, NFTs have rapidly evolved into a distinct type of cryptocurrency (Fairfield, 2021). They include various digital assets such as videos, images, photos, GIFs, virtual avatars, and audio content, extending their application to real-world physical objects. Through the secure features of blockchain technology, NTF ownership records remain accessible and immutable, ensuring that there is a single legitimate owner at any given time (Chainalysis, 2021; Popescu, 2021).

NFTs have attracted significant interest from artists, brands, and marketing investors as they have become a thriving marketplace for artists and creators. This dynamic eco-systems provides artists and creators with an exceptional avenue to enhance self-promotion and increase revenue. In 2022, the trading volume of NFTs across various platforms of digital artwork reached approximately \$24.7 billion (Molenaar, 2023), with a projected value expected to reach \$6.2 billion by 2028 (Molenaar, 2023). In addition, market forecasts point to a promising projection that 19.31 million NFT users will be active participants by 2027 (Statista, 2023). Hence, the value of NFTs as an effective tool for creating unique marketing strategies has been widely acknowledged by brands, as seen through the highly successful Adidas NFT collection, Nike's "Cryptokicks" release for verifying athletic sneakers, and the CEO of Instagram's firm stance on NFT integration (Boag and Rich, 2020). NFTs are gaining recognition for their potential to represent individual status and digital identity within specific communities, indicating collective confidence in their long-term value and growing significance beyond monetary gain (Jaipuria, 2021; Kelly, 2021).

NFT communities and online communities have distinct characteristics, although there is some overlap between them. NFT communities specifically center around the creation, exchange, and utilization of of NFTs, primarily to display digital art, collectibles, and analog digital objects (Wilson, 2023). These communities develop on

specific platforms such as forums, social media channels, and NFT marketplaces, where they serve as discussion centers for NFT projects, artists, market trends, and transactions. The dynamic nature of NFT communities involves the buying, selling and showcasing of NFTs, providing unique monetization opportunities. In contrast, online communities, accessible through various platforms, pursue broader objectives include socializing, sharing information, providing support, and discussing various topics, with monetization as a secondary focus (Martin, 2023). NFT communities, influenced by blockchain and cryptocurrency advances, have a unique character that differs from the technologically diverse but less blockchain-focused online communities (Behl et al., 2024).

While NFT communities have become increasingly popular and important to individuals, companies, and brands, there remains a dearth in this area, as highlighted by Nadini et al. (2021). Distinct from conventional online social communities, connecting with an NFT community signifies a heightened inclination towards owning an NFT collection. This is indicative of a deliberate effort to integrate into the community and signals a profound level of interest. Engagement with an NFT community results in increased personal investment and fosters a stronger connection with both NFTs and fellow enthusiasts, thereby intensifying the desire for continued involvement within that specific community.. Hence, exploring the drivers of these communities can play a crucial role in shaping individual perspective and interactions.

Recent studies have made significant progress in advancing the understanding of NFTs (e.g., Ali et al., 2023; Ante, 2021; Evans, 2019; Nadini et al., 2021; Van Haaften-Schick and Whitaker, 2021; Wang et al., 2021). However, these studies primarily focus on the technical dimensions of NFTs. For example, Wilson et al. (2022) have extensively documented the definition of NFTs, their interaction with blockchain and other cryptocurrencies, and their diverse applications across industries. They have also explored the facilitators and risks associated with NFTs. Nevertheless, some researchers (e.g., Zbinden, 2021; Wang et al., 2021) suggest that NFT technology is still in its early stages. Conversely,, Bile (2023) emphasizes the essential role of community support for the success of NFTs. Despite its growing importance of NFTs communities, there is a notable lack of literature dedicated to NFT communities (Colicev, 2023). This calls for further research to understand how

consumer engagement in NFT communities, compared to social media, affects different stages of the marketing process. Therefore, This paper aims to answer three main research questions:

RQ1: What is the effect of social factors (social capital and social exchange) on users' engagement with NFTs community?

RQ2: How do user's attitude effect their engagement with NFTs community?

To address these questions, we have developed a research model grounded in social capital theory (Nahapiet and Ghoshal, 1998) and social exchange theory (Blau, 1964; Homans, 1958). This model is developed to explore how social capital and social exchange influence users' attitudes toward NFTs and their engagement in NFT communities. We used social capital theory since NFTs are considered as a new form of social capital (Haque, 2021), which can explain the reasons based on social capital for the people involved in NFT community. On another side, our research uses Social Exchange Theory (SET) to understand the motivations of NFT users to participate in these NFT communities since NFT community is also a social community where members share their knowledge and help each other. Based on this, we will examine the impact of social capital factors—namely, structural capital, relational capital, cognitive capital, and social exchange beliefs, including reciprocity, reputation, trust, and altruism, on users' attitudes toward NFT communities and their engagement within these communities.

## **2.0 Theoretical Framework**

### **2.1. Social Capital Theory**

Social capital theory (SCT), as defined by Nahapiet and Ghoshal (1998), encompasses the tangible and potential resources within an individual's or group's relationships. Various pro-social behaviours have been elucidated using this concept, including the exchange of information and knowledge (Chang and Hsu, 2016). SCT enables the sharing and integration of resources within a social group and also provides a justification for the existence of the organisation (Chen and Hung, 2014). Previous studies have investigated social capital in different contexts, such as social networking sites (Phua et al., 2017), gaming (Teng, 2018), and job performance (Swanson, 2020). However, no study has yet explored the contribution of social capital to NFTs, despite the fact that NFTs are considered a new form of social capital (Haque, 2021).



Therefore, this study seeks to address this research gap by examining the role of social capital in NFT community engagement.

## 2.2. Social Exchange Theory

Social Exchange Theory (SET), which originated in the 1950s and is rooted in psychology, provides a basic framework for analysing human behaviour and relationships in order to understand social structures (Blau, 1964; Homans, 1958). Initially applied to the study of human behaviour, SET was later extended to the study of organisational behaviour, emphasising the importance of norms, social institutions and formal exchanges between organisations. An essential aspect of SET is to understand what drives exchange (Emerson, 1981). It postulates that a person engages in an exchange process after evaluating costs and benefits, and ultimately enters into relationships to minimise costs and maximise benefits (Cloarec et al., 2022). SET has been applied in various domains, including service robots (Kim et al., 2022), AI-powered intelligence intention (Jiang et al., 2022), family-firm relationships (Hayward et al., 2022), and online social communities (Qin et al., 2011; Shiau and Luo, 2012). Following these studies, our research adopts SET to understand what motivates NFT users to engage with these NFT communities.

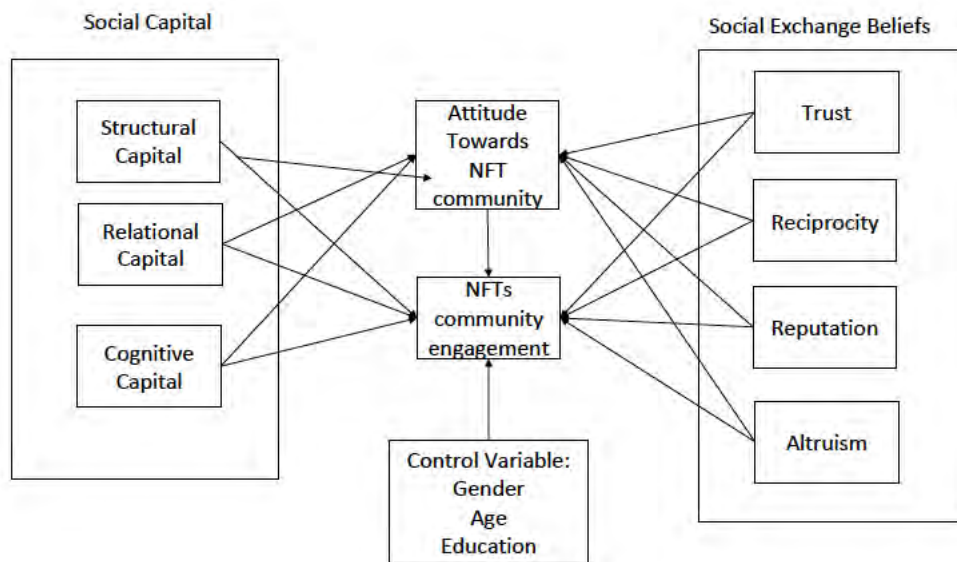


Fig.1. Proposed Conceptual Model

### **3.0. Proposed Methodology and Expected Results**

The research model comprises nine constructs: Structural, cognitive, and relational capital, which will be evaluated based on Chang and Hsu's (2016) study. Reciprocity, as defined by Shiau and Luo (2012), will be used to assess social exchange beliefs, and trust will also be evaluated based on their research. Altruism and attitude will be examined in Hsu and Lin's (2008) study. Finally, the construct of NFT community engagement is adapted from Kang et al.'s (2016). To evaluate the proposed research model, we will conduct a survey-based study to investigate the impact of social capital and social exchange related factors on user's attitude toward NFTs community and NFTs community engagement. To collect the data, we will invite users who have experience with NFTs and are active members of NFT communities on LinkedIn and specific WhatsApp groups. This study will use a purposive sampling method with a non-probability approach, carefully selecting samples that meet specific criteria and share similar characteristics (Uma Sekaran & Bougie, 2013). The analysis will consist of two stages; in the first stage we will use partial least square (PLS)-based structural equation modelling (SEM). In this second stage, as a complementary analysis, we will use fuzzy qualitative comparative analysis to determine the conditions that are sufficient for NFT community engagement (e.g., Al-Emran et al., 2023). Based on our analysis, we expect that social capital factors and social exchange beliefs will have a positive impact on users' attitudes towards NFT communities and NFT community engagement. Furthermore, we expect a positive effect of users' attitudes towards the NFT community on NFT community engagement.

### **4.0 Expected Contribution and Conclusion**

The dynamic landscape of NFTs presents a promising frontier for exploring the intersection between technology, finance, and the human desire for interaction and community. In this paper, we have introduced a research model that delineates the influence of social factors on users' attitudes and engagement within NFT communities. To the best of our knowledge, this proposed research model represents one of the first studies dedicated to investigating the role of social factors in the context of NFT communities. The results of this study expected to contribute to both academia and practical applications. From academic perspectives, we extend the body of knowledge on customer engagement and NFTs context by proposing a research model that integrates the impact of social factors on users' attitude and engagement.

From practical perspective, given that NFTs represent emerging community commerce platforms, it becomes imperative to gain insights into how the social dynamics of these platforms influence the intensity of engagement.

## References

- Ali, O., Momin, M., Shrestha, A., Das, R., Alhaji, F. and Dwivedi, Y. K. (2023) *A review of the key challenges of non-fungible tokens*, Technological Forecasting and Social Change, 187 122248.
- Al-Emran, M., AlQudah, A. A., Abbasi, G. A., Al-Sharafi, M. A. and Iranmanesh, M. (2023) *Determinants of using AI-based chatbots for knowledge sharing: evidence from PLS-SEM and fuzzy sets (fsQCA)*, IEEE Transactions on Engineering Management, 1-15.
- Ante, L. (2021) *Non-fungible token (NFT) markets on the Ethereum blockchain: Temporal development, cointegration and interrelations*, Available at SSRN 3904683.
- Behl, A., Pereira, V., Nigam, A., Wamba, S. and Sindhvani, R. (2024) *Knowledge development in non-fungible tokens (NFT): a scoping review*, Journal of Knowledge Management, 28(1), 232-267.
- Bile, J. (2023) *The Importance of Creating a Community for your New NFT Project*, Accessed at: <https://www.linkedin.com/pulse/importance-creating-community-your-new-nft-project-jay-ar-bile/>
- Blau, P. M. (1964) *Exchange and power in social life*, Exchange and Power in Social Life, 1-352.
- Boag, D. and Rich, J. (2020) *Nike's Dec 2019 patent reveals revolutionary NFT use*. Medium, Accessed at: <https://nftnyc.medium.com/nikes-dec-2019-patent-reveals-revolutionary-nft-usea74c115bd0c>.
- Chandra, Y. (2022) *Non-fungible token-enabled entrepreneurship: A conceptual framework*, Journal of Business Venturing Insights, 18 e00323.
- Chang, C. M. and Hsu, M. H. (2016) *Understanding the determinants of users' subjective well-being in social networking sites: An integration of social capital theory and social presence theory*, Behaviour & Information Technology, 35 720-729.
- Chainalysis (2021) *Report Preview: The 2021 NFT Market Explained*, Accessed at: <https://www.chainalysis.com/blog/nft-market-report-preview-2021/>.
- Chen, P. C. and Hung, S. W. (2014) *Collaborative green innovation in emerging countries: a social capital perspective*, International Journal of Operations & Production Management, 34 347-363.
- Cloarec, J., Meyer-Waarden, L. and Munzel, A. (2022) *The personalization privacy paradox at the nexus of social exchange and construal level theories*, Psychology & Marketing, 39 647-661.
- Colicev, A. (2023) *How can non-fungible tokens bring value to brands*, International Journal of Research in Marketing, 40 30-37.
- Evans, T. M. (2019) *Cryptokitties, cryptography, and copyright*, AIPLA QJ, 47 219.

- Emerson, R. M. (1981) *Social exchange*, In M. Rosenberg and R. Turner (Eds.), *Social psychology: Sociological perspective* (pp. 3-24). New York: Basic Books.
- Fairfield, J. A.T. (2022) *Tokenized: The Law of Non-Fungible Tokens and Unique Digital Property*, *Indiana Law Journal*, 97 4.
- Haque, F. (2021) *NFTs are the new social capital*, Accessed at: <https://erlystage.com/nfts-are-the-new-social-capital/>
- Hayward, M., Hunt, R. and Miller, D. (2022) *How vulnerability enriches family firm relationships: A social exchange perspective*, *Journal of Family Business Strategy*, 13 100450.
- Homans, G. C. (1958) *Social behavior as exchange*, *American Journal of Sociology*, 63 597-606.
- Jaipuria, T. (2021) *NFTs as financial assets, status, identity, community and fun*, Accessed at: <https://tanay.substack.com/p/nfts-as-financial-assets-status-identity>.
- Jiang, H., Cheng, Y., Yang, J. and Gao, S. (2022) *AI-powered chatbot communication with customers: Dialogic interactions, satisfaction, engagement, and customer behavior*, *Computers in Human Behavior*, 134, 107329.
- Qin, Y. S. (2020) *Fostering brand–consumer interactions in social media: the role of social media uses and gratifications*, *Journal of Research in Interactive Marketing*, 14, 337-354.
- Kelly, D. (2021) *Instagram Is "Actively Exploring" NFTs, According to CEO Adam Mosseri*, Accessed at: <https://hypebeast.com/2021/12/instagram-actively-exploring-nfts-ceo-adam-mosseri>
- Kim, H., So, K. K. F. and Wirtz, J. (2022) *Service robots: Applying social exchange theory to better understand human–robot interactions*, *Tourism Management*, 92, 104537.
- Martin, M. (2023) *The Top 17 Best Online Community Platforms in 2024*, Accessed at: <https://www.thinkific.com/blog/best-online-community-platforms/>
- Molenaar, K (2023) *NFTs Statistics – Sales, Trends and More [2023]*, Accessed at: <https://influencermarketinghub.com/nfts-statistics/>
- Nadini, M., Alessandretti, L., Di Giacinto, F., Martino, M., Aiello, L. M. and Baronchelli, A. (2021) *Mapping the NFT revolution: market trends, trade networks and visual features*. arXiv preprint arXiv:2106.00647.
- Nahapiet, J. and Ghoshal, S. (1998) *Social capital, intellectual capital, and the organizational advantage*, *Academy of Management Review*, 23 242-266.
- Penston, J. (2007) *Patients' preferences shed light on the murky world of guideline-based medicine*, *Journal of Evaluation in Clinical Practice*, 13 154-159.
- Popescu, A. D. (2021) *Non-fungible tokens (nft)–innovation beyond the craze*, In 5th International Conference on Innovation in Business, Economics and Marketing Research, 32.
- Phua, J., Jin, S. V. and Kim, J. J. (2017) *Uses and gratifications of social networking sites for bridging and bonding social capital: A comparison of Facebook, Twitter, Instagram, and Snapchat*, *Computers in Human Behavior*, 72 115-122.
- Shiau, W. L. and Luo, M. M. (2012) *Factors affecting online group buying intention and satisfaction: A social exchange theory perspective*, *Computers in Human Behavior*, 28 2431-2444.
- Statista (2023) *NFT – Worldwide*, Accessed at: <https://www.statista.com/outlook/dmo/fintech/digital-assets/nft/worldwide>.

- Swanson, E., Kim, S., Lee, S. M., Yang, J. J. and Lee, Y. K. (2020) *The effect of leader competencies on knowledge sharing and job performance: Social capital theory*, Journal of Hospitality and Tourism Management, 42 88-96.
- Teng, C. I. (2018) *Look to the future: Enhancing online gamer loyalty from the perspective of the theory of consumption values*, Decision Support Systems, 114 49-60.
- Van Haaften-Schick, L. and Whitaker, A. (2021) *From the artist's contract to the blockchain ledger: New forms of artists' funding using equity and resale royalties*, Available at SSRN 3842210.
- Wang, Q., Li, R., Wang, Q. and Chen, S. (2021) *Non-fungible token (NFT): Overview, evaluation, opportunities and challenges*, arXiv preprint arXiv:2105.07447.
- Wilson, A. (2023) *The Rise of NFTs and the Need for NFT Marketplace Platforms on Different Niche*, Accessed at: <https://medium.com/illumination/nft-marketplace-platforms-on-different-niche-842fbf0ad124>
- Wilson, K. B., Karg, A. and Ghaderi, H. (2022) *Prospecting non-fungible tokens in the digital economy: Stakeholders and ecosystem, risk and opportunity*, Business Horizons, 65 657-670.
- Zbinden, F. (2021) *NFTs – a new era of crypto adoption or just another fad?*, Accessed at: <https://www.linkedin.com/pulse/nft-new-era-crypto-adoption-just-another-fad-fabian-zbinden/>

# Self-Determination Theory in Crowdfunding: Data from Vietnam

*Research In progress*

## **Abstract**

*This research aims to establish a comprehensive and validated model of backer motivations in crowdfunding, serving as a foundational framework for future studies in this domain. Leveraging the Self-Determination Theory to categorize crowdfunding backer motivations into intrinsic and extrinsic categories, our study seeks to unveil disparities in the impact of intrinsic and extrinsic motivation on crowdfunding. Furthermore, it endeavors to identify the primary motivating factor that drives individuals to contribute to crowdfunding projects. Based on the data collected from Vietnam, this research has the potential to offer valuable guidance to both scholars and crowdfunding practitioners, particularly for the emerging markets.*

**Keywords:** crowdfunding, intrinsic motivation, extrinsic motivation, self-determination theory, emerging markets.

## **1.0 Introduction**

Crowdfunding helps to generate a certain amount of money from a diverse group of people on the Internet to finance various projects (Harms, 2007; Goldfarb, 2013; Nevin et al., 2017; Rodriguez-Ricardo et al., 2018; Brem et al., 2019; Mendes-Da-Silva et al., 2019; Kim et al., 2020a). A variety of factors are known to influence crowdfunding intention for would-be backers, including those backers' personality traits (Cox et al., 2017; Rodriguez-Ricardo et al., 2018), platform choice and campaign type (Gleasure & Feller, 2016; Cox & Nguyen, 2018), information provided by the fund-seeker (Hornuf & Schwienbacher, 2015; Alcántara-Pilar et al., 2018; Foster, 2019), identity narratives constructed by fund-seekers (Herzenstein et al., ), the use of social media (Polzin et al., 2018; Moritz et al., 2015; Nevin et al., 2017), and herding effects (Herzenstein et al., 2011). Hence, understanding backer motivation can help project owners to launch worthy crowdfunding campaigns and successfully realize important projects.

Self-Determination Theory is an influential theory to understand motivation and it prescribes intrinsic and extrinsic motivation as the determinants for people to get motivated to do something (Deci & Ryan, 1985). A limited number of articles have



examined crowdfunding through Self-Determination Theory. However, prior research in this area is meagre (Chen et al., 2019) and in the incipient stages, e.g., focused exclusively on charity-based crowdfunding (Chen et al., 2021; Pitchay et al., 2022; Wang et al., 2019), or, in other crowdfunding contexts, focusing on creating, but not yet verifying, lists of motivations for crowdfunding intention (Yao et al., 2016), or on verifying the linkages among intrinsic and extrinsic cues and crowdfunding participation, while not yet addressing the underlying backer motivation (Allison et al., 2015). Therefore, this research seeks to address this gap by posing the research question: “*How do intrinsic and extrinsic values contribute to promote crowdfunding project participation?*”. In this study, it will be applied to design a framework to categorize backer motivations in crowdfunding and sets a base for comparisons with other similar funding and crowd participation instruments.

We plan to collect data from Vietnam, a developing country in Southeast Asia where crowdfunding initiatives are rapidly increasing in number. Prior research in crowdfunding has already investigated similar developing country contexts such as ASEAN-5 (Indonesia, Malaysia, Thailand, Singapore, and Philippines; Dikaputra et al., 2019), China (Chen et al., 2021; Wang and Xue., 2019) and Oman (Pitchay, 2022), though each of these are limited a single type of crowdfunding and there is no prior research that we know of in Vietnam, specifically. Pietro and Buttice (2020) would predict that crowdfunding is curtailed in countries where regulatory environments are not very business-friendly and which embrace collectivistic cultures, such as developing countries in Asia; but at the same time, developing countries are likely to benefit from crowdfunding since traditional finance is lacking in many regards.

## **2.0 Literature review**

### **2.1. Backer motivation to participate in crowdfunding.**

Several studies have been conducted to understand backer motivation in crowdfunding. These studies have found that collecting rewards, helping others, being part of a community, and supporting a cause important to oneself are key to motivating investors (Gerber et al., 2012). A number of studies (Aitamurto, 2011; Burtch et al., 2012; Choy & Schlagwein, 2015; 2016; Gleasure & Feller, 2016; Hui et al., 2012; Jian & Shin (2015); Li et al., 2017; 2018; Mariani et al., 2017; Schwienbacher & Larralde, 2010) note other motivations such as the need for approval, interactions with other people, and generating attention for one’s organization, as well as altruism, fun, satisfaction, contributing to the community, and creating social change. There is some inconsistency

in the findings, with some scholars emphasizing externally provided benefits (Bretschneider and Leimeister, 2017; Cholakova & Clarysse, 2015; Gerber & Hui, 2013; Gerber & Hui, 2016; Yao et al., 2016; Mariani et al., 2017) while others emphasize the backers' own interest to participate (Allison et al., 2015; Gerber & Hui, 2013; Zhang & Chen, 2018a).

## **2.1. Self-Determination Theory**

Self-Determination Theory (Deci & Ryan, 1985) is an established theory of motivation that attempts to explain people's interest to take actions through the concepts of intrinsic and extrinsic motivation (Ryan & Deci, 2000). A limited number of articles have examined crowdfunding through the Self-Determination Theory. A few studies focus exclusively on charity-based crowdfunding contexts (Chen et al., 2021; Pitchay et al., 2022; Wang and Xue, 2019). In other contexts, Yao et al. (2016) propose a model of crowdfunding participation drawn from qualitative research which rests on Self-Determination Theory; however, they do not verify the relationships in this model through quantitative research. Allison et al. (2015) applied several intrinsic and extrinsic cues derived from the crowdfunding proposition to examine crowdfunding participation through a quantitative study, but they did not directly examine participant motivations. Given that some crowdfunding campaigns promise tangible prizes to backers, and others mainly rely on backers' goodwill, the two types of motivation in SDT seem very relevant to explain backing motivation. In summary, the application of Self-Determination Theory has intrigued scholars, but this stream of research has not yet resulted in a tested and verified model of backer motivations, prompting the present research.

## **3.0 Hypothesis development**

### **3.1. Intrinsic motivation**

Based on Self-Determination Theory (Deci & Ryan, 1985), the three innate psychological needs – competence, autonomy, and relatedness – are integral to any decision-making process. Competence is the feeling of self-efficacy when fulfilling a task, autonomy is the self-determination of a person's behavior, and relatedness is the level of connection of an individual to the issue. Our review of the determinants to crowdfunding intention found that intrinsic determinants can be divided into three groups: altruism, self-worth and personal development, which will be discussed below in turn.

#### **3.1.1. Altruism**



Altruism will prompt users to contribute to online communities, such as when members of online communities offer help, experiences, support, and reviews to others in online travel communities (Yoo & Gretzel, 2008), or when software developers seek to help the open-source community (Lukkarinen et al., 2018). Indeed, such altruistic behaviours stem from enjoyment in helping, satisfaction in the action itself, and hedonic motivation (Wang & Fesenmaier, 2004). It is obvious that altruistic motivations are particularly critical in crowdfunding that relies mainly on the goodwill of backers, such as is the case with donation crowdfunding. Indeed, research on charity-based crowdfunding demonstrates that altruism is an important motivator (Jian & Shin, 2015; Gleasure & Feller, 2016; Chen et al., 2021); however, it should not be assumed that altruism is unimportant for other types of crowdfunding (Cholakova & Clarysse, 2015; Bretschneider & Leimeister, 2017). On this basis, altruism is a contributing factor to the backing decision in crowdfunding. Hence, the first hypothesis is:

*H1. Altruism positively motivates backers to contribute to crowdfunding projects.*

### **3.1.2. Self-worth**

Individuals will adjust their beliefs about themselves by comparing themselves with referents such as peers (Festinger, 1954). Furthermore, social comparisons will cause individuals to calibrate their actions (Festinger, 1954). In a crowdfunding platform, the profile of a backer with their social media connections will be highlighted along with the contributed amount, which exposes the backer to social comparisons. If an online community member does not possess sufficient self-esteem, they will end up contributing less to online fora (Kim et al., 2011). Self-worth is relevant when individuals back initiatives which align with their life goals. For example, backers can be motivated by supporting a cause important to them, contributing to the community, or creating social change (Aitamurto, 2011). We propose this mechanism can also be extended to crowdfunding decisions. The hypothesis is:

*H2: Self-worth positively motivates backers to contribute to crowdfunding projects.*

### **3.1.3. Personal development**

According to Moysidou (2016), crowdfunding backers are considered rational investors who are constantly striving for maximum benefits, and they consequently develop personal needs to influence the outcome of the project. This need has been investigated by Bretschneider et al. (2015) in innovative projects, as customers think that through their participation in the project, they can influence the product features that are valuable to them. By active contribution in the product development process backers could enhance their knowledge in several aspects from financial knowledge, product

development to management skills. In addition, engagement in community activities and interacting with other members could bring their enjoyment of playfulness (Yoo & Gretzel, 2008). This perception of pleasure and joyfulness further deepens members' involvement and sense of belonging to the community (Koh & Kim, 2003). Since members can be motivated by this self-satisfaction, they are likely to contribute more efforts in collective work in online travel communities (Wang & Fesenmaier, 2004; Yoo & Gretzel, 2008). As a result, the following hypothesis is proposed:

*H3: Personal Development positively motivates backers to contribute to crowdfunding projects.*

### **3.2. Extrinsic Value**

Self-Determination Theory holds that external rewards can act as motivators for individuals' actions alongside intrinsic rewards if they do not contradict the individual's autonomy (Deci & Ryan, 1985). Such external rewards can be many; for example, the equity or other rewards received from an investment, or more intangible rewards such as verbal praise from the project owner. Our review found that extrinsic motivation for crowdfunding can be divided into three groups: reward, recognition, and social connection, which will be discussed below in turn.

#### **3.2.1. Reward**

Many investors in crowdfunding anticipate a tangible return (Ferreira, 2018). The return can be in the form of a limited-edition product, a physical prize of interest to the backer, or financial profit. This incentive does not have to be material; it might just simply be an act of acknowledgment (Gerber and Hui, 2013). It is understandable that rewards are critical in equity and reward-based crowdfunding, where participants expect to receive something tangible (Cholakova & Clarysse, 2015; Bretschneider & Leimeister, 2017; Cox et al., 2017; Dikaputra et al., 2019; Allon & Babich, 2020). However, it is possible though not necessarily the norm that backers will receive a tangible reward of some kind also in charity-based crowdfunding. Therefore, a reward is one of the primary reasons why investors participate in crowdfunding, and leads to the hypothesis:

*H4: Backers in crowdfunding projects are positively motivated when there are financial returns and rewards.*

#### **3.2.2. Recognition**

The comment section in the crowdfunding project is used by backers to express their desires for the project outcome and simultaneously used for recognizing the backer's status, achievements, and merit. This desire for recognition stems from a person's need

for fame and esteem. Recognition is commonly researched in online open-source communities, and users on these platforms look for explicit recognition to be more active in participation and contribution via the feedback culture (Hars and Ou, 2002). Backers in charity-based crowdfunding, in particular, are able to satisfy their need of approval by society by participating in the initiative (Collins and Pierrakis, 2012; Hui et al., 2012; Li et al., 2017; Schwienbacher and Larralde, 2010; Jian & Shin, 2015); but this mechanism may well operate in other types of crowdfunding, too. Therefore, we believe that recognition acts as a motive that drives backers to make an online contribution, which gives them positive reactions from other backers and the project creators. The research hypothesis is:

*H5: Backers in crowdfunding projects are positively motivated if there is recognition from other people.*

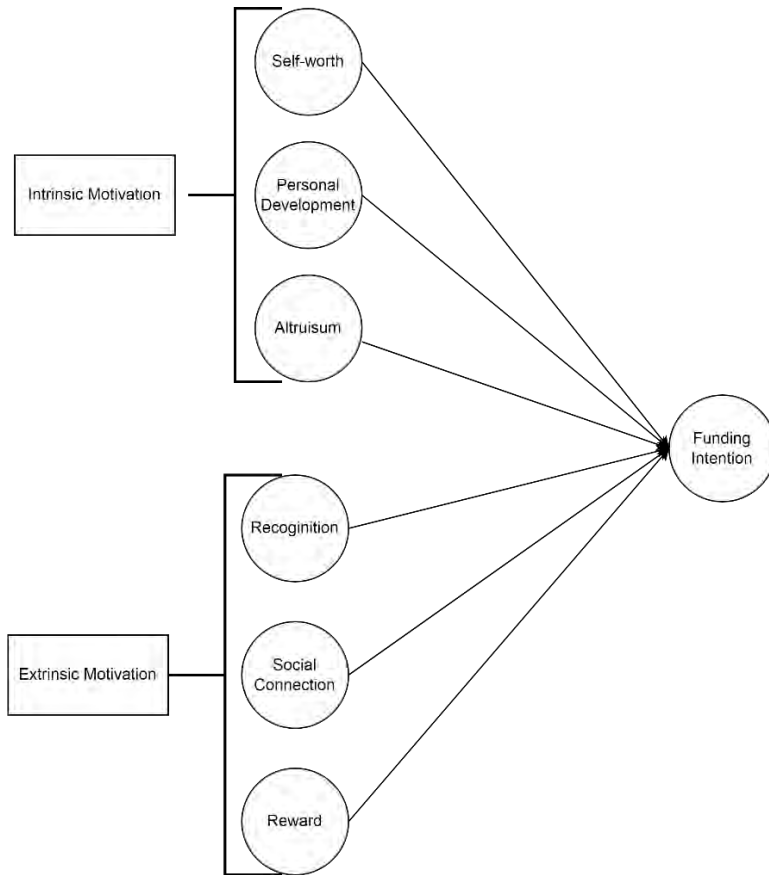
### **3.2.3. Social Connection**

According to Moysidou (2016), crowdfunding backers may have various types of personal affiliations to project founders and/or project issues. Such connections can be a motivation to support the crowdfunding project. Social connection can also be applied to other potential backers, as individuals observe others' investing decisions and are more likely to contribute towards a project when others they know also make similar contributions (Frey & Meier, 2004). Research in reward and charity-based crowdfunding has recognized people's interest in supporting a campaign when it includes their family and friends (Jian & Shin, 2015), when they personally know the campaign creator and beneficiaries (Choy & Schlagwein, 2016), or just helping out (Dai & Zhang, 2019); following social influence such as norms and morality (Chen et al., 2021; Pitchay et al., 2021), contributing to the community through the campaign (Aitamurto, 2011; Gerber et al., 2012; Jian & Shin, 2015), and engaging socially through the campaign (Choy and Schlagwein, 2015; Gleasure and Feller, 2016; Hui et al., 2012; Jian & Shin, 2015). Thus, we have the following hypothesis:

*H6: Backers in crowdfunding projects are positively motivated if there is social connection with other people.*

## **3.4 Research Model**

Based on the preceding literature review, the conceptual framework of this study is depicted in Figure 1. This proposed integrated model was validated by the data collected from the mobile banking services in Vietnam.



## **4.0 Research Method**

### **4.1. Development of measurement instruments**

The questionnaire was designed based on the research model and the study of Bretschneider (2017). The independent variables have two parts, which are the intrinsic and extrinsic values of backers in making investment decisions. Both the intrinsic and extrinsic value consist of 3 distinct motivations. Each value is regarded as a latent variable measured by 3, 4 or 5 observed variables respectively. The dependent variable is the investment decision, which focuses on whether the backer has intention to participate in crowdfunding projects. The questionnaire was constructed using the Likert-type scale, ranging from strongly disagree (1) to strongly agree (5), measuring the logical consensus of backers based on the observed variables. In addition, it was validated by ten experts with master's or doctorate degrees in the field to ensure the clarity and aim of each question towards the research goal.

### **4.2. Data collection**

This study utilized both probability sampling and snowball sampling methods. It welcomed participation from Vietnamese individuals across all regions through various social media platforms. Surveys were distributed using a combination of online and offline approaches, provided that respondents are over 18-year-old, possessed knowledge of crowdfunding and/or had previously made a backing decision in a crowdfunding project. We started our survey with a screening question of whether the participant had heard or known about crowdfunding or participated in any crowdfunding project.

## **5.0 Expected Contributions**

This research is expected to provide a tested and verified model of backer motivations for crowdfunding that sets a standard for future research in the area. Given that Self-Determination Theory is an accepted and widely adopted theory for understanding motivation, its application to categorize crowdfunding backer motivations into intrinsic and extrinsic types will enable the creation of a useful instrument. Prior literature applying STD in crowdfunding is limited to single types of crowdfunding or unverified constructs. In particular, as altruistic and egoistic motivations often intermingle in crowdfunding (Gleasure & Feller, 2016), SDT is expected to answer to the particular characteristics of crowdfunding. In a practical sense, the research will provide suggestions for crowdfunding campaign owners to determine how they wish to appeal

to would-be backers, thus enabling more popular and successful campaigns.

## References

- Aitamurto, T., 2011. The impact of crowdfunding on journalism. *Journalism Practice* 5 (4), 429–445.
- Alcántara-Pilar, J. M., Francisco, J. B. E., Tanja, A., and Del Barrio-García, S. (2018). The antecedent role of online satisfaction, perceived risk online, and perceived website usability on the affect towards travel destinations. *Journal of Destination Marketing Management* 9, 20–35.
- Allison, T.H., Davis, B.C., Short, J.C., Webb, J.W., 2015. Crowdfunding in a prosocial microlending environment: examining the role of intrinsic versus extrinsic cues. *Entrepreneurship Theory and Practice* 39 (1), 53–73.
- Allon, G. & Babich, V. (2020). Crowdsourcing and Crowdfunding in the Manufacturing and Services Sectors. *Manufacturing & Service Operations Management* 22(1), 102-112. <https://doi.org/10.1287/msom.2019.0825>
- Belleflamme, P., & Lambert, T. (2014). Crowdfunding: Some Empirical Findings and Microeconomic Underpinnings. *Electronic Journal*. Available at: <http://dx.doi.org/10.2139/ssrn.2437786>
- Benjamin, G. & Margulis, J. (2000). Angel financing: *How to find and invest in private equity*. New York: Wiley.
- Brem, A., Bilgram, V., and Marchuk, A. (2019). How crowdfunding platforms change the nature of user innovation – from problem solving to entrepreneurship. *Technological Forecasting and Social Change*, 144, 348–360.
- Bretschneider, U., Leimeister, J.M., 2017. Not just an ego-trip: exploring backers' motivation for funding in incentive-based crowdfunding. *Journal of Strategic Information Systems* 26 (4), 246–260.
- Burtch, G., Ghose, A., Wattal, S., 2012. An empirical examination of the antecedents and consequences of investment patterns in crowd-funded markets. *Information Systems Research* 1–48.
- Chen, Y., Dai, R., Wang, L., Yang, S., Li, Y. and Wei, J. (2021). Exploring donor's intention in charitable crowdfunding: intrinsic and extrinsic motivations. *Industrial Management & Data Systems*, Vol. 121 No. 7, 2021, pp. 1664-1683.
- Cholakova, M., Clarysse, B., 2015. Does the possibility to make equity investments in crowdfunding projects crowd out reward-based investments? *Entrepreneurship Theory and Practice* 39 (1), 145–172.
- Choy, K., Schlagwein, D., 2015, May. IT Affordances and Donor Motivations in Charitable Crowdfunding: The “Earthship Kapita” Case. ECIS.
- Choy, K., and Schlagwein, D. (2016). Crowdsourcing for a better world on the relation between IT affordances and donor motivations in charitable crowdfunding. *Information Technology and People* 29, 221–247.
- Cox, J. & Nguyen, T. (2018). Does the crowd mean business? An analysis of rewards-based crowdfunding as a source of finance for start-ups and small businesses. *Journal of Small Business and Enterprise Development*, 25 (1), 147-162.
- Dai, H. and Zhang, D. J. (2019). Prosocial Goal Pursuit in Crowdfunding: Evidence from Kickstarter. *Journal of Marketing Research*, 56(3), 498-517.
- Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. New York: Penguin.
- Dikaputra, R., Sulung, L.A.K. & Kot, S. (2019). Analysis of Success Factors of Reward-Based Crowdfunding Campaigns Using Multi-Theory Approach in ASEAN-5 Countries. *Social Sciences* 8, 293.
- Ferreira, F., & Pereira, L. (2018). Success Factors in a Reward and Equity Based

- Crowdfunding Campaign. *2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*, 17-20. June, 2018, Stuttgart, Germany.
- Festinger, L., 1954. A theory of social comparison processes. *Human Relations* 7(2), 117–140.
- Frey, B., Meier, S., 2004. Social comparisons and pro-social behavior: testing conditional cooperation in a field experiment. *Am. Econ. Rev.* 94 (5), 1717–1722.
- Füller, J., 2010. Refining virtual co-creation from a consumer perspective. *California Management Review* 52 (2), 98–122.
- Gerber, E.M., Hui, J.S., & Kuo, P.Y. (2012). Crowdfunding: Why people are motivated to post and fund projects on crowdfunding platforms. *In Proceedings of the International Workshop on Design, Influence, and Social Technologies: Techniques, Impacts and Ethics*, 2, 11.
- Gerber, E.M., Hui, J., 2013. Crowdfunding: motivations and deterrents for participation. *ACM Transactions on Computer-Human Interaction* 20 (6), 34.
- Gleasure, R., Feller, J., 2016. Does heart or head rule donor behaviors in charitable crowdfunding markets? *International Journal of Electronic Commerce* 20 (4), 499–524.
- Gorczyca, M., Hartman, R.L., 2017. The new face of philanthropy: the role of intrinsic motivation in millennials' attitudes and intent to donate to charitable organizations. *Journal of Nonprofit & Public Sector Marketing* 29 (4), 415–433.
- Hars, A., Ou, S., 2002. Working for Free? Motivations for participating in open source projects. *Int. J. Electron. Commerce* 6 (3), 25–39
- Herzenstein, M., Dholakia, U. M., and Andrews, R. L. (2011). Strategic herding behavior in peer-to-peer loan auctions. *Journal of Interactive Marketing* 25, 27–36.
- Herzenstein, M., Sonenshein, S. & Dholakia, U.M. (2011). Tell Me a Good Story and I May Lend You Money: The Role of Narratives in Peer-to-Peer Lending Decisions. *Journal of Marketing Research* 48, 138-149.
- Hornuf, L., and Schwienbacher, A. (2015). “Funding Dynamics in Crowd investing,” in *Research Papers in Economics*. Muenster: Universität Trier.
- Jabr, W., Mookerjee, R., Tan, Y., Mookerjee, V.S., 2014. Leveraging philanthropic behavior for customer support: the case of user support forums. *MIS Quarterly*, 38 (1), 187–208.
- Jian, L., Shin, J., 2015. Motivations behind donors' contributions to crowd-funded journalism. *Mass Communication and Society* 18 (2), 165–185.
- Kim, H.-W., Zheng, J.R., Gupta, S., 2011. Examining knowledge contribution from the perspective of an online identity in blogging communities. *Computers in Human Behavior*, 27, 1760-1770.
- Kim, M. J., Bonn, M., and Lee, C. K. (2020a). The effects of motivation, deterrents, trust, and risk on tourism crowdfunding behavior. *Asia Pacific Journal of Tourism Research* 25, 244–260.
- Kuppuswamy, V. & Bayus, B.L. (2017). Does my contribution to your crowdfunding project matter? *Journal of Business Venturing* 32, 72–89.
- Li, J.J., Chen, X., Kotha, S., Fisher, G., Li, J.J., Fisher, G., 2017. Catching fire and spreading it: a glimpse into displayed entrepreneurial passion in crowdfunding campaigns. *Journal of Applied Psychology* 1–16.
- Li, Y.Z., He, T.L., Song, Y.R., Yang, Z., Zhou, R.T., 2018. Factors impacting donors' intention to donate to charitable crowd-funding projects in China: a UTAUT-based model. *Information, Communication and Society* 21 (3), 404–415.
- Lukkarinen, A., Wallenius, J., & Seppälä, T. (2017). Investor Motivations and Decision Criteria in Equity Crowdfunding. *Electronic Journal*.

- Mariani, A., Annunziata, A., Aprile, M.C., Nacchia, F., 2017. Crowdfunding and wine business: some insights from Fundovino experience. *Wine Econ. Policy* 6 (1), 60–70.
- Mason, C.M. & Harrison, R.T. (2008). Measuring business angel investment activity in the United Kingdom: A review of potential data sources. *Venture Capital*, 10(4), 309–330.
- Mendes-Da-Silva, W., Rossoni, L., Francisco, E., and Gattaz, C. C. (2019). “Geographic distance and reward crowdfunding donations,” in *Individual Behaviors and Technologies for Financial Innovations*, ed. Mendes-Da-Silva, W (Heidelberg: Springer International Publishing), 189–214.
- Mollick, E. (2012). The dynamics of crowdfunding: determinants of success and failure. *Journal of Business Venturing* 29, 1–16.
- Moritz, A., Block, J. & Lutz, E. (2015). Investor communication in equity-based crowdfunding: a qualitative-empirical study. *Qualitative Research in Financial Markets*, 7 (3), 309-342.
- Moysidou, K., & Spaeth, S. (2016). Cognition, emotion and perceived values in crowdfunding decision-making. Paper presented at the *Open and User Innovation Conference*, 31<sup>st</sup> August, 2016, Boston, USA.
- Nevin, S., Gleasure, R., O’Reilly, P., Feller, J., Shanping, L., and Cristoforo, J. (2017). “Social identity and social media activities in equity crowdfunding,” in *Proceedings of the 13th International Symposium on Open Collaboration* August 23 - 25, 2017 (Galway: Association for Computing Machinery).
- Pichler, F., and Tezza, I. (2016). “Crowdfunding as a new phenomenon: origins, features and literature review”, in *Crowdfunding for SMEs: a European perspective*, eds. R. Bottiglia and F. Pichler, (London: Macmillan Publishers Ltd), 5–44.
- Pitchay, A.A., Eliz, N.M.A., Ganesan, Y., Mydin, A.A., Ratnasari, R.T. and Thaker, M.A.M.T. (2021). Self-determination theory and individuals’ intention to participate in donation crowdfunding. *International Journal of Islamic and Middle Eastern Finance and Management*, 15 (3), 506-526.
- Polzin, F., Toxopeus, H., and Stam, E. (2018). The wisdom of the crowd in funding: information heterogeneity and social networks of crowdfunders. *Small Business Economics* 50, 251–273.
- Popescu, D., Radu, L.R., Pavaloaia, V.D., and Georgescu, M.R. (2020). Psychological Determinants of Investor Motivation in social media-Based Crowdfunding Projects: A Systematic Review. *Frontiers in Psychology*
- Rodriguez-Ricardo, Y., Sicilia, M., and López, M. (2018). What drives crowdfunding participation? The influence of personal and social traits. *Spanish Journal of Marketing* 22, 163–182.
- Ryan, R.M. & Deci, E.L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54–67.
- Wang, Y. and Fesenmaier, D.R. (2004). Towards understanding members’ general participation in and active contribution to an online travel community. *Tourism Management*, 25, 709-722.



- Wang, Z., and Xue, Y. (2019). Understanding backers' funding intention in reward crowdfunding: an elaboration likelihood perspective. *Technology in Society* 58:101149.
- Yao, H., Wang, Z, Zhou, T. (2016). Study on the investors' participation motivations in award-based crowdfunding. *International Business and Management* 12(2), 29–33. <https://doi.org/10.3968/8352>.
- Yoo, K.H. and Gretzel, U. (2008). What motives consumers to write online travel reviews? *Information Technology & Tourism*, 10, 283-295.
- Zhang, H., Chen, W., 2018a. Backer motivation in crowdfunding new product ideas: is it about you or is it about me? *Journal of Product Innovation Management* 36 (2), 241–262.

# The Impact of Streamer's Speaking Signalling on Live Streaming Sales and Post-Purchase Behaviour

## Abstract

*Live streaming commerce has emerged as a distinctive form of e-commerce, wherein products or services are promoted and transacted through real-time video broadcasts. This evolving paradigm leverages the immediacy and interactivity fostered between streamers and their audience during live video content. While there is a widespread recognition that the effectiveness of live sales is contingent upon the presentation skills of streamers in articulating product value and stimulating audience purchase, a critical examination of the influence of streamers' linguistic signals on audience purchasing behaviour remains insufficiently explored. This study, grounded in Signalling Theory, endeavours to investigate whether specific attributes of streamers' speech, including speaking speed, pitch, and vocabulary richness, exert an impact on the live sales. Furthermore, the research delves into the potential extension of charismatic speaking to post-purchase phenomena, encompassing transactions subsequent to the conclusion of live video broadcasts.*

**Keywords:** live-streaming commerce, streamers, speaking charisma, Signalling Theory, live sales, post purchase.

## 1.0 Introduction

In the era dominated by ubiquitous Internet technology, the development of live streaming commerce has witnessed a marked and significant growth in recent years. Recognised as a pivotal marketing instrument, this mode of commerce facilitates the broadening of businesses' outreach to a more expansive customer base, thereby enabling the exploitation of substantial business opportunities (Global Live Streaming Market Report, 2023). The escalating scholarly focus on this domain is discernible through the large body of research papers dedicated to unravelling the intricacies

inherent in contemporary e-commerce practices (Bao & Zhu, 2023; Zhang, Cheng, & Huang, 2023; Zhou & Huang, 2023).

The pivotal role played by streamers is evidently underscored as a foundational prerequisite for the operations of live streaming commerce (Wongkitrungrueng, Dehouche, & Assarut, 2020). Positioned as key facilitators, they contribute significantly to enhancing product presentations, instigating synchronous communication, and cultivating dynamic social interactions with potential buyers (Lu & Chen, 2021; Xu, Wu, Chang, & Li, 2019). The evolution of interactions between streamers and viewers has transcended the traditional framework of live video broadcasts as a mere conduit for product sales. Instead, it has evolved into a forum where participants engage in bidirectional information exchange and communication, fundamentally altering the nature of engagement between streamers and viewers (Hu, Zhang, & Wang, 2017; Xu et al., 2019). Notwithstanding the presence of prior research examining the influence exerted by streamers in orchestrating on-site video broadcasts, there has been a limited focus on scrutinising the linguistic signals employed by streamers, encompassing elements such as speaking pitch and vocabulary richness. This deficiency has resulted in an insufficient exploration of the implications of communication signals on sales performance.

This study aims to address this gap by investigating the streamers' linguistic signals and their potential influence on audience purchasing behaviour both during and after live streams. Specifically, the research interrogates the extent to which streamers' speaking signals, notably speed, pitch and vocabulary richness, impact the audience's purchasing behaviour. Furthermore, the study delves into the potential continuum of the influence of streamers' charismatic speaking into post-purchase phenomena,

encompassing transactions occurring after the culmination of live video broadcasts. In pursuit of these objectives, this study addresses this research question: **To what extent and in what manner do the linguistic signals employed by streamers impact their audience's purchasing behaviour, both during and after live streams?**

To address the research question, we employ Signalling Theory (Connelly, Certo, Ireland, & Reutzel, 2011) and operationalise a model that integrates the speaking signals of streamers, as the signal senders, within the live streaming context. The study investigates the resultant effects on the purchasing behaviour of the audience, who serve as the receivers of these signals. This conceptual framework is anticipated to yield valuable insights into the intricate interplay between streamers' linguistic attributes and the consumer purchasing behaviour. Moreover, it seeks to elucidate the mechanisms through which streamers can strategically construct approaches to enhance sales performance within the domain of live streaming commerce. We convert linguistic signals of streamers from unstructured data (MP3 files) into quantifiable data (speech rate and pitch variation). Then, by using quantitative methods to analyse its relation with our pre-processed sales data, we infer the relationship between these signals and the outcomes. We hope that by doing so, streamers can refer to this and improve these linguistic signals to achieve optimized sales performance.

## **2.0 Literature Review**

### **2.1 Speaking Signals in Live Streaming Commerce**

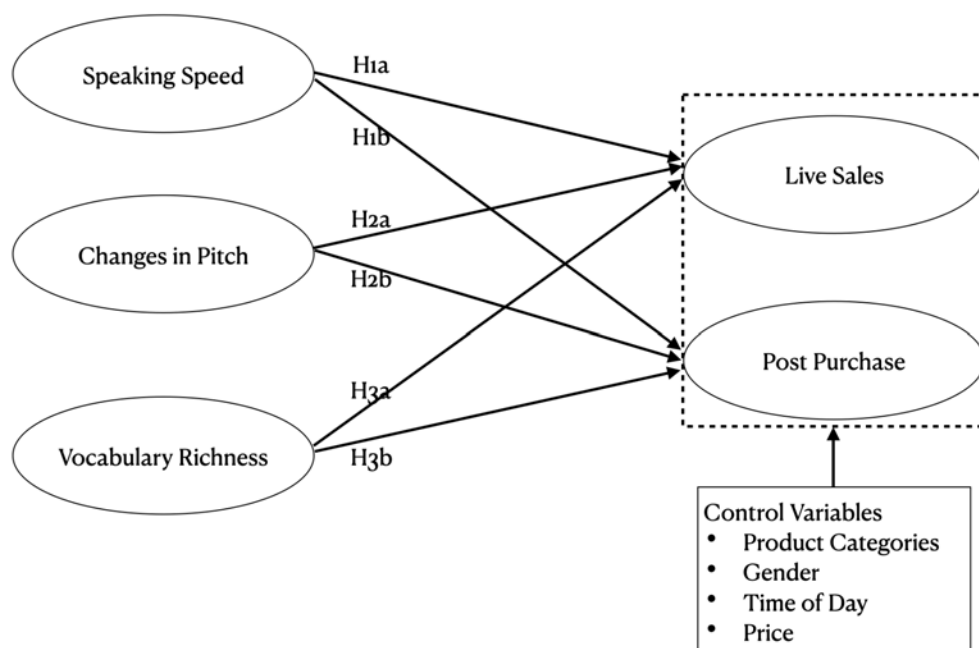
The acquisition of information is pivotal in decision-making across various contexts, ranging from individual purchase decisions to organisational management, and entrepreneurship. Signalling Theory, at its core, addresses the imperative of mitigating information asymmetry among involved parties, where varying amounts and qualities

of information come into play (Spence, 2002). Consequently, one party, termed the *signaller*, must deliberate on whether and how to communicate, or signal, pertinent information; Simultaneously, the other party, as the *receiver*, confronts the task of interpreting the signal (Connelly et al., 2011). In this theoretical concept, signals encompass a diverse array of information related to social cues, actions, emotions, attitudes, and relationships. Examples of such signals span voices, gestures, messages conveyed in conversations, facial expressions, emotions, and the dynamics of relationships between individuals or groups (Poggi & Francesca, 2010). In essence, Signalling Theory elucidates the complexities of communication strategies in situations where the exchange of information significantly influences decision-making processes.

In the realm of live streaming commerce, the effectiveness of video broadcasts hinges largely on the performance of streamers who demonstrate products through active presentations, providing potential buyers with information crucial for their decision-making (Lu & Chen, 2021). Streamers play a central role in delivering detailed and vivid information to the audience, thereby fostering their interests, encouraging participation, eliciting feedback, and ultimately prompting their purchasing intentions and subsequent actions. The communication process involves a multitude of signals conveyed from streamers to viewers. While a predominant share of signals pertains to non-verbal phenomena, such as facial expressions and eye contact, it is imperative to note that verbal indicators also wield considerable influence, exemplified by the concept of verbal intensity and variation (Brunet & Cowie, 2012). Despite the significance of signals in the context of live streaming, there exists a notable dearth of studies examining the impact of streamers' speaking signals on audience purchase behaviours.

## 2.2 Research Model and Hypotheses

To address the identified research gap, our study introduces a research model (see Figure 1) that focuses on the linguistic signals exhibited by streamers during live broadcasts and their subsequent impact on both live sales and post-purchase activities. Specifically, our investigation encompasses distinct attributes of streamers' speech, such as speaking speed, pitch, and vocabulary richness, seeking to ascertain their substantial influence on the phenomenon of live sales and post-purchase transactions.



**Figure 1. The Proposed Model of Streamers' Linguistic Signals**

In the proposed model, we first posit that the precise modulation of speaking speed holds paramount significance for streamers in facilitating effective communication with their audience. Right speaking speed is deemed essential to ensure that the audience comprehends the conveyed message, remains engaged, and responds positively to product presentations, and, in turn, manifests as immediate order placement during the live video presentation or in the post-purchase phase. Previous research has found that the pace at which sellers deliver informative signals directly affects customers'

cognitive processes, thereby influencing their capacity to comprehend and process content—a pivotal determinant in consumer purchasing decisions (Peterson, Cannito, & Brown, 1995). Given the nature of live streaming commerce to physical on-site sales, characterised by a vibrant and enthusiastic atmosphere, the imperative lies in actively engaging (potential) consumers and persuading them towards making a purchase. We hypothesise that a fast speaking speed contributes positively to both live sales and post-purchase transactions.

H1a: Fast speaking speed positively associated with live sales.

H1b: Fast speaking speed is positively associated with post-purchase transactions.

Secondly, because intonation, pitch and pausing are indicators within the realm of communication signals, we will also take these factors into consideration.

H2a: Variations in Pitch are positively related to Live Sales.

H2b: Variations in Pitch are positively related to Post Purchase.

Finally, in recognition of the pivotal role linguistic expression plays within the live streaming commerce, we incorporate the richness of vocabularies into our framework. Although a substantial portion of communication signals relies on non-verbal cues, the nuanced interplay between these cues and verbal indicators is of significance (Brunet & Cowie, 2012), particularly in the context of live streaming, where researchers assert that comprehensive and articulate product descriptions provided by streamers play a pivotal role in influencing sales performance (Wang, Liu, & Fang, 2021). Empirical evidence underscores the considerable engagement of more than 30% of consumers in live streams, driven by a primary objective of acquiring profound insights into specific

product details. Johnson, Safadi, and Faraj (2015) further accentuate that an augmented richness of vocabulary in communication significantly contributes to the perception of individuals as significant influencers. In alignment with these considerations, we posit the following hypotheses:

H3a: Vocabulary Richness is positively related to Live Sales.

H3b: Vocabulary Richness is negatively related to Post Purchase.

#### **4.0 Research Methodology**

We gathered a substantial dataset from a Taiwanese live streaming company, comprising MP3-formatted video clips from 1,004 streamers. Each streamer's file contains records of their live sales from 2021 to 2022, with each host conducting between 10 and 100 live streaming sessions. Therefore, we have a total of 10,000 to 100,000 live streaming sessions, categorized into several main categories, including food, apparel, cosmetics, furniture, and others, available for analysis. The average total sales per live streaming session range from approximately £1,500 to £8,000. With the company's consent, we acquired data on both live sales from their historical records and post-purchase transactions from their e-commerce platform.

In this research, speaking speed is operationalised as “the number of utterances per minute.” We utilised OpenAI Whisper<sup>1</sup> to transcribe the collected audio files into text. Variation in pitch is defined as “the highness/lowness of a streamer’s voice.” We employed the methodology outlined in Roohi, Mekler, Tavast, Blomqvist, and Hämäläinen (2019) research and utilised the Librosa library package<sup>2</sup> to calculate the root-mean-square of pitch<sup>3</sup>. Additionally, for assessing “vocabulary richness,” we

---

1 Whisper is a general-purpose speech recognition model. It is trained on a large dataset of diverse audio and is also a multitasking model that can perform multilingual speech recognition, speech translation, and language identification.

2 Librosa is a Python package for analyzing and processing audio signals.

3 Root Mean Square is a measure of the average magnitude of a set of values. In the context of audio signals, the value is used to describe the "loudness" of a sound waveform, taking into account the



adapted Johnson et al. (2015)'s definition, considering it as the average number of unique words per text file in a speaking context. To investigate the influence of these linguistic elements on sales performance more precisely, we adopted the control variables proposed, including gender, product categories, time of day and price. Chi-square tests will be conducted to explore the relationships between these variables within the proposed model.

## **5.0 Conclusions**

This study is currently in the data analysis phase, focused on cleaning the dataset and conducting statistical evaluations. Future efforts in this study will focus on expanding the dataset to enhance the generalizability of the research findings. We envisage that by the time of the conference at Kent University, we will have garnered additional research findings, affording us the opportunity to disseminate these results to the scholarly audience in attendance.

The forthcoming results are poised to yield two substantive contributions. Firstly, from a theoretical standpoint, we aim to articulate a mechanism that offers valuable insights for streamers and live streaming managers. Specifically, our research endeavours to elucidate how the construction of linguistic signals can be strategically leveraged to enhance sales outcomes.

Secondly, from a methodological perspective, the analytical approaches we employ bear significance, particularly in the context of analysing linguistic elements in Chinese—a domain where comprehensive research is notably limited. By presenting and applying robust analysis methods, we aspire to contribute methodological

---

varying amplitudes of the sound over time.

advancements that extend the current understanding of linguistic dynamics in Chinese live streaming contexts.

## References

- Bao, Z., & Zhu, Y. (2023). Understanding customers' stickiness of live streaming commerce platforms: An empirical study based on modified e-commerce system success model. *Asia Pacific Journal of Marketing and Logistics*, 35(3), 775-793.
- Brunet, P. M., & Cowie, R. (2012). Towards a conceptual framework of research on social signal processing. *Journal on Multimodal User Interfaces*, 6, 101-115.
- Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of management*, 37(1), 39-67.
- Hu, M., Zhang, M., & Wang, Y. (2017). Why do audiences choose to keep watching on live video streaming platforms? An explanation of dual identification framework. *Computers in Human Behavior*, 75, 594-606.
- Johnson, S. L., Safadi, H., & Faraj, S. (2015). The emergence of online community leadership. *Information systems research*, 26(1), 165-187.
- Lu, B., & Chen, Z. (2021). Live streaming commerce and consumers' purchase intention: An uncertainty reduction perspective. *Information & Management*, 58(7), 103509.
- Peterson, R. A., Cannito, M. P., & Brown, S. P. (1995). An exploratory investigation of voice characteristics and selling effectiveness. *Journal of Personal Selling & Sales Management*, 15(1), 1-15.
- Poggi, I., & Francesca, D. E. (2010). *Cognitive modelling of human social signals*. Paper presented at the Proceedings of the 2nd international workshop on Social signal processing.
- Roohi, S., Mekler, E. D., Tavast, M., Blomqvist, T., & Hämäläinen, P. (2019). *Recognizing emotional expression in game streams*. Paper presented at the Proceedings of the annual symposium on computer-human interaction in play.
- Spence, M. (2002). Signaling in retrospect and the informational structure of markets. *American economic review*, 92(3), 434-459.
- Wang, Y., Liu, J., & Fang, Y. (2021). Storytelling in E-marketplace Live Streaming: The Effects of Narrative Content and Delivery on Sales Performance.
- Wongkitrungrueng, A., Dehouche, N., & Assarut, N. (2020). Live streaming commerce from the sellers' perspective: implications for online relationship marketing. *Journal of Marketing Management*, 36(5-6), 488-518.
- Xu, X., Wu, J.-H., Chang, Y.-T., & Li, Q. (2019). The investigation of hedonic consumption, impulsive consumption and social sharing in e-commerce live-streaming videos.
- Zhang, X., Cheng, X., & Huang, X. (2023). "Oh, My God, Buy It!" Investigating impulse

buying behavior in live streaming commerce. *International Journal of Human-Computer Interaction*, 39(12), 2436-2449.

Zhou, Y., & Huang, W. (2023). The influence of network anchor traits on shopping intentions in a live streaming marketing context: The mediating role of value perception and the moderating role of consumer involvement. *Economic Analysis and Policy*, 78, 332-342.

# Responsible Digital Innovation with Micro-Businesses: A Resource Orchestration Perspective

**Rayen Jui-Yen Chang**

National Chengchi University, Taiwan  
& Lancaster University Management  
School, UK  
j.chang9@lancaster.ac.uk

**Niki Panteli**

Lancaster University Management  
School, UK  
n.panteli1@lancaster.ac.uk

*Research In progress*

## Abstract

*While the increasing focus on responsible innovation (RI) highlights its growing significance in the context of digital innovation, most RI research is centred on large organizations. Scant research has explored how resources are interconnected within the context of micro-businesses (MBs). To analyse the impact of MBs on resource configuration, this study takes a theoretical lens from resource orchestration theory (ROT) to understand how organizations effectively structure, bundle, and leverage resources. This work-in-progress research involves a qualitative case study of an MB in Taiwan, named Royalty (a pseudonym). The expected findings aim to reveal insights into how MBs orchestrate their resources for RI implementation, highlighting the challenges and opportunities they face. In this study, we delve deeper into this aspect by examining how MBs leverage their resource portfolios to facilitate innovation with a strong emphasis on responsibility.*

**Keywords:** Responsible innovation, Digital innovation, Micro-businesses, Resource orchestration theory, Qualitative case study

## 1.0 Introduction

With the growing interconnection of science and innovation, discussions regarding responsible innovation (RI)<sup>1</sup> have broadened. Stilgoe et al. (2013) state that “RI means taking care of the future through collective stewardship of science and innovation in the present” (p.1570). The academic discourse has addressed different aspects of RI, from conceptual questions (e.g., Raman, 2014) to integrating RI into organizations (e.g., Ambos & Tatarinov, 2022), and to assessing outcomes and consequences (e.g., Yaghmaei & Poel, 2021). Notably, this evolving discourse on RI has become increasingly intertwined with the field of digital innovation. Scholars call upon businesses to engage in responsible digital innovation, entailing the development and adoption of digital products and services that do no harm, do good, and are governed in a responsible manner (Scherer & Voegtlin, 2020).

---

<sup>1</sup> Here, ‘RI’ refers both to ‘responsible innovation’ and ‘responsible research and innovation’.

Despite extensive discourse, most RI literature mainly centres on larger companies involved in commercialising innovation, with a focus on addressing ethical concerns and responsibilities in the innovation decision-making process (Lukovics et al., 2017). This emphasis on larger companies creates a notable gap in understanding, as micro-businesses (MBs) inherently lack the extensive resources and capabilities found in larger counterparts (Shirish et al., 2023). MBs, characterized by their small scale, limited workforce of normally fewer than 10 employees, and significantly fewer resources (OECD, 2021). Therefore, the environment of RI within the MBs sector remains an understudied research area (Moore & Spence, 2006).

The idea of RI has relevance to the MBs sector for several reasons. First, due to their flexibility and adaptability, MBs are more agile in innovation compared to their larger companies (Dabić et al., 2021). Second, focusing on sustainable product, service, and business model innovation can be more effective for promoting responsibility in MBs than expecting them to follow the formal strategic management practices of larger companies (Nasiri et al., 2022). Third, while MBs face resource constraints compared to large companies (Parker et al., 2009), research indicates that resource scarcity can actually be beneficial for innovation (Gibbert et al., 2007). Therefore, we argue that responsible MBs have the potential to address resource limitations in their own strategies. Although researchers have recognized the resources needed for RI within MBs (Halme & Korpela, 2014), they lack clarity on the configuration and interconnection of these resources to implement RI. This study looks to close this gap in the RI literature by addressing the following research question: **How can MBs configure, integrate, and leverage their limited resources in order to foster RI in their digital initiatives?** This paper aims to understand the relationship between resources in MBs for RI.

To address this question, we employed the theoretical lens of resource orchestration theory (ROT), as it is emphasised that organisational outcomes are achieved through the use of resources (Lopez et al., 2022). According to the ROT, an organization can realize the value of its resources only through effective structuring, bundling, and management of those resources (Asiaei et al., 2020; Sirmon et al., 2011). Therefore, we argued that the perspective on ROT can offer an appropriate way to conceptualise the intricate interplay of resources in MBs for the purpose of RI (Sirmon et al., 2011). This theory allows for an examination of how MBs configure, integrate, and leverage

their limited resources to drive innovation with a focus on responsibility which aligns with the aim of our study.

For this study, we use a qualitative case study research design. The research setting involved one rural MB, named *Royalty* (a pseudonym), founded by a married couple deeply rooted in tea farming, has evolved into a family enterprise, with their team expanded to include their three children (college students). The core business of *Royalty* centres around the promotion and sale of locally cultivated agricultural products sourced from Lishan<sup>2</sup> region. They strategically leverage the region's agricultural limited resources, effectively aligning their strategies with the area's pronounced seasonal changes. This case provided us with an unparalleled opportunity to acquire an in-depth understanding of how resources are configured and connected to implement RI within MBs sector.

This study is currently in progress. The expected findings propose a mechanism that involves the structuring, bundling, and management of these resources for configuration. These resources collectively support MBs in achieving resource integration for RI.

## **2.0 Theoretical Foundations**

RI is an innovation process where the emphasis lies not only on the product but also on the purpose and method of innovation (Von Schomberg, 2012). For MBs, RI may be their core product, service, or business model, constituting their primary focus rather than being an additional part within their business or product portfolio (Halme & Korpela, 2014). This makes it even more important to investigate the combination and configuration of resources needed for implementing RI in the context of MBs. However, previous studies have focused on larger companies, leaving a significant knowledge gap in our understanding of how MBs structure and connect their resources for RI.

Given the influence of resource configuration on RI within MBs, the research paradigm of ROT captures the intricate interplay of resources (Sirmon et al., 2011). Based on the ROT, organisations can exploit the full potential of their resources and capabilities by deploying them in a complementary manner (Liu et al., 2016). As

---

<sup>2</sup> Lishan, nestled in the heart of Taiwan, owes its unique and rare agricultural products to the high-altitude climate, including Lishan tea, fragrant pears, and honey apples.

highlighted by Nevo and Wade (2010), resources should be bundled and managed effectively to create innovations. However, the major obstacle in implementing resource orchestration is understanding how managers can mobilize and structure resources (Asiaei et al., 2020; Miao et al., 2017). This challenge is particularly relevant for MBs, given their unique constraints and limited resources. In this study, exploring how MBs orchestrate their resources becomes a crucial point, as it can reveal insights into RI within their specific context.

### **3.0 Research Methods**

A qualitative case study method (Myers, 2020) and a narrative inquiry approach (Pentland, 1999) were selected to investigate the relationship between resources within MBs during the implementation of RI processes. For this study, we have so far selected one MB, called Royalty (a pseudonym), as it represents a typical MB within our research context, providing insights into resource configuration strategies and emphasizing a resource-driven approach to encourage RI. For example, Royalty's strategic connections to the Lishan region's agricultural resources during each season illustrate the dynamic resource orchestration necessary for RI in an MB.

In the primary data collection of this study, to date, we have done one informal interview and one formal interview. The informal interview provided an opportunity to establish rapport and gain a preliminary understanding of Royalty's operations and their approach to RI. This interview lasted approximately 2.5 hours and served as a foundation for the subsequent formal interview. Specifically, we have gained an understanding of the current situation of Royalty, the challenges they have encountered in the past, and how they have addressed and resolved them. During the formal interview (approximately 1.5 hours), we delved deeper<sup>3</sup> into specific aspects related to resource configuration, innovation processes, and their alignment with RI practices within Royalty. However, while we've made significant progress in data collection, we acknowledge that there's more work ahead. In the upcoming phases (around three months), we plan to conduct additional interviews, archival data collection, and potentially observations to enrich our data. We also plan to undertake interviews with several other MBs in the same region to enrich our data collection.

---

<sup>3</sup> One of the questions raised at the formal interview: can you describe how you manage and organize your resources, especially in the context of developing new innovative products or services?



## 4.0 Implications & Conclusions

The findings are expected to reveal the aspects of resource orchestration in MBs within the context of RI. It is expected that these findings will contribute to a deeper understanding of how MBs configure, integrate, and leverage their limited resources to drive innovation with a focus on responsibility. Moreover, this study aims to uncover the challenges and opportunities faced by MBs in their pursuit of RI and provide insights into how ROT can be applied in this specific context.

In conclusion, this study addresses a significant gap in understanding the interconnection and management of resources for responsible digital innovation in MBs. By taking the theoretical lens of ROT and conducting a qualitative case study on Royalty, a Taiwanese MB, has offered an opportunity into the practical application of ROT in fostering RI practices among MBs.

We believe that this topic provides useful and interesting insights into ROT. Currently, our study is a work in progress, with more exploration needed in understanding how resources configure for RI implementation in MBs. Our study is progressing positively and will be in a position to present further significant research findings at the upcoming conference.

## References

- Ambos, T. C., & Tatarinov, K. (2022) *Building responsible innovation in international organizations through intrapreneurship*, Journal of Management Studies, 59(1) 92-125.
- Asiaei, K., Barani, O., Bontis, N. and Arabahmadi, M. (2020) *Unpacking the black box: how intrapreneurship intervenes in the intellectual capital-performance relationship?* Journal of Intellectual Capital, 21(6) 809-834.
- Dabić, M., Stojčić, N., Simić, M., Potocan, V., Slavković, M., & Nedelko, Z. (2021) *Intellectual agility and innovation in micro and small businesses: The mediating role of entrepreneurial leadership*, Journal of Business Research, 123 683-695.
- Gibbert, M., Hoegl, M., & Välikangas, L. (2007) *In praise of resource constraints*, MIT Sloan management review, 48(3) 15.
- Halme, M., & Korpela, M. (2014) *Responsible innovation toward sustainable development in small and medium-sized enterprises: A resource perspective*, Business Strategy and the Environment, 23(8) 547-566.
- Liu, H., Wei, S., Ke, W., Wei, K.K. and Hua, Z. (2016) *The configuration between supply chain integration and information technology competency: a resource orchestration perspective*, Journal of Operations Management, 44(1) 13-29.
- Lopez, C., Haddoud, M. Y., and Kasturiratne, D. (2022) *Revisiting the innovation-export entry link through a configuration approach*, Journal of Business Research, 149 927-937.

- Lukovics, M., Flipse, S. M., Udvari, B., & Fisher, E. (2017) *Responsible research and innovation in contrasting innovation environments: Socio-Technical integration research in Hungary and the Netherlands*, *Technology in Society*, 51 172-182.
- Miao, C., Qian, S. and Ma, D. (2017) *The relationship between entrepreneurial self-efficacy and firm performance: a meta-analysis of main and moderator effects*, *Journal of Small Business Management*, 55(1) 87-107.
- Moore, G., & Spence, L. (2006) *Responsibility and small business*, *Journal of Business Ethics*, 67 219-226.
- Myers, M. D. (2020) *Qualitative Research in Business and Management*, Sage, London.
- Nasiri, M., Saunila, M., Rantala, T., & Ukko, J. (2022) *Sustainable innovation among small businesses: The role of digital orientation, the external environment, and company characteristics*, *Sustainable Development*, 30(4) 703-712.
- Nevo, S., & Wade, M. R. (2010) *The formation and value of IT-enabled resources: antecedents and consequences of synergistic relationships*, *MIS quarterly*, 163-183.
- OECD. (2021) *Enterprises by Size*. Retrieved 5/10/2023 from <https://data.oecd.org/entrepreneur/enterprises-by-business-size.htm>
- Parker, C. M., Redmond, J., & Simpson, M. (2009) *A review of interventions to encourage SMEs to make environmental improvements*, *Environment and planning C: Government and policy*, 27(2) 279-301.
- Pentland, B. T. (1999) *Building process theory with narrative: From description to explanation*, *Academy of management Review*, 24(4) 711-724.
- Raman, S. (2014) *Responsible innovation: from concept to practice*, *Journal of Responsible Innovation*, 1(3) 329-331.
- Scherer, A. G., & Voegtlin, C. (2020) *Corporate governance for responsible innovation: Approaches to corporate governance and their implications for sustainable development*, *Academy of Management Perspectives*, 34(2) 182-208.
- Shirish, A., Srivastava, S. C., & Panteli, N. (2023) *Management and sustenance of digital transformations in the Irish microbusiness sector: examining the key role of microbusiness owner-manager*, *European Journal of Information Systems*, 1-25.
- Sirmon, D. G., Hitt, M. A., Ireland, R. D., & Gilbert, B. A. (2011) *Resource orchestration to create competitive advantage: Breadth, depth, and life cycle effects*, *Journal of Management*, 37(5) 1390-1412.
- Stilgoe, J., Owen, R., & Macnaghten, P. (2013) *Developing a framework for responsible innovation*, *Research policy*, 42(9) 1568-1580.
- Von Schomberg, R. (2012) *Prospects for technology assessment in a framework of responsible research and innovation*, In M. Dusseldorp and R. Beecroft (Eds), *Technikfolgen Abschätzen Lehren: Bildungspotenziale Transdisziplinärer Methoden*, Wiesbaden, Springer, pp. 39-61.
- Yaghmaei, E., & Poel, I. V. D. (2021) *Assessment of responsible innovation: Methods and practices*, Taylor & Francis.

# Mapping and Visualising the Digital Economy in The Context of Developing Countries: A Bibliometric Analysis

## Authors

Ubongabasi Kingsley Omon<sup>1</sup>

Gordon Fletcher<sup>2</sup>

Mohammed Ali<sup>3</sup>

## Affiliations

The University of Salford, Manchester, UK<sup>1,2,3</sup>

## Complete Research Work

(With a short response to Peer Reviewers' comments on the first draft inserted in Appendix 2)

## Abstract

*The digital economy wields profound transformative potential, particularly in the context of developing nations. Yet, a notable void in systematic and comprehensive knowledge persists within this domain. This paper bridges the gap by conducting a meticulous bibliometric analysis of scholarly publications within the field of inquiry, drawing upon papers retrieved from the Scopus database and published exclusively in the English language. Employing a combination of simple metadata aggregations and illustrative visualisations, this study scrutinises publication trends, research trajectories, influential authors, institutions, and prominent journals within this field of investigation. Furthermore, it identifies the research gaps and opportunities for future studies. This paper contributes significantly to the existing literature by providing a novel, all-encompassing overview of knowledge concerning the digital economy in developing countries between 2003 and 2023 by presenting the identified research trends discovered in papers published in the English language. The output also offers valuable insights for policymakers, practitioners, and researchers who are interested in advancing the digital economy, thus fostering digital transformation in developing nations.*

**Keywords:** Digital Economy, Digital Transformation, Developing Countries, Bibliometric Analysis, VOSviewer, Visualisation.

## 1. Introduction

Bibliometric analysis is a quantitative methodology that employs mathematical and statistical techniques to assess the interconnectedness and impacts of publications, authors, institutions, and nations within a designated field of research – including in business, management, and social science domains (Donthu et al., 2021).

The digital economy has emerged as a pervasive and transformative force that has engendered substantial shifts in industries, economies, and societies on a global scale. Nevertheless, while a considerable body of research scrutinises its effects on *developed nations*, a notable void persists concerning its intricate dynamics within *developing nations and emerging markets*. This knowledge gap hinders the development of effective strategies for sustainable growth and inclusive progress. Evident within this scholarly lacuna is a scarcity of empirical data addressing essential aspects of the digital economy in developing countries that researchers will require to narrow down subsequent studies. To rectify this limitation and propel the field of inquiry forward, the execution of a meticulous bibliometric analysis is unequivocally imperative.

### 1.1. Justification: Why the Digital Economy and Why Developing Countries?

The outcome of this research is poised to furnish essential insights that address current informational insufficiencies and offer foundational groundwork for robust scholarly discourse and informed decision-making. Specifically, this bibliometric analysis will facilitate the identification of underexplored research areas and gaps within the digital economy domain specific to developing countries. By analysing the outcome of this study, future researchers can pinpoint themes and aspects that require more comprehensive investigation. Relatedly, the resulting insights produced at the end of this analysis will empower scholars to focus their efforts on topics that are crucial for understanding and addressing the unique challenges faced by developing nations attempting to leverage technology to build a viable digital economy.

Furthermore, the mapping of the scholarly landscape related to the digital economy in developing countries will help researchers to visualise insights stemming from scholarly contributions of authors, institutions, and journals to the field of inquiry. These visual representations will aid scholars to comprehend the influential nodes within the field, thereby fostering potential partnerships and collaborations. The study also identifies influential authors contributing significantly to the discourse on the digital economy in developing countries, and

by evaluating citation volumes and publication frequency, researchers can recognise thought leaders who shape the direction of research. Studying the emerging research trends and shifting foci presented in this study will allow scholars to stay updated and responsive to the evolving landscape.

Similarly, since policymakers, governmental bodies, and development organisations rely on data-driven insights to formulate effective policies and strategies (van Ooijen et al., 2019), the outcome of this research will provide empirical evidence that informs policy decisions related to the digital economy in developing nations. By understanding the areas of research that are gaining traction and the ones that require more attention, policymakers can craft targeted interventions for fostering innovation, economic growth, and digital inclusivity. By doing so, government agencies, developmental institutions, and funding partners will be aided in making informed decisions about allocating resources strategically, including funding initiatives, grants, and research projects. In effect, this ensures that investments align with the areas that hold the most promise for advancing the digital economy in developing countries, thereby resulting in optimal resource utilisation.

Finally, an initial exploration of the Scopus database yielded no ongoing or previously conducted bibliometric analyses published in the English language pertaining to the digital economy within the context of developing countries, hence this study with the aim of addressing the gap in the current research landscape. Said otherwise, the primary objective of this study is to discover the current research trends identifiable in scholarly papers published in the English language concerning the digital economy in developing countries spanning the period between 2003 and 2023. The empirical evidence produced is intended to aid the understanding of the landscape and then propose future research directions.

## **2. Research Questions**

Considering the broad nature of the concepts, *digital economy*, and *developing countries*, respectively, the following research questions (RQ) were developed to narrow the focus of this study:

**RQ 1:** What are the current research trends identifiable in scholarly papers published in the English language concerning the digital economy in developing countries?

To assist in answering the main question above, other related questions were developed. They are listed below:

**RQ 2:** Which publication outfit publishes the most papers in the English language regarding the digital economy in developing nations from 2003 to 2023?

**RQ 3:** Which country has the highest number of research publications in the English language regarding the digital economy in developing nations from 2003 to 2023?

**RQ 4:** Which organisations fund research into the digital economy in developing nations, and what country are they located?

**RQ 5:** Who are the top contributing authors, and most cited authors relating to research published in the English language about the digital economy in developing nations?

**RQ 6:** What are the most frequently used keywords and terminologies common in papers published in the English language regarding the digital economy in developing nations from 2003 to 2023?

### **3. Research Method**

This section outlines the research methodology employed to conduct the bibliometric analysis. The bibliometric analysis method was utilised to quantify, aggregate, and evaluate the patterns of scholarly publication, citation, and collaboration, offering insights into the evolution and impact of research in this domain. Be that as it may, the scope of this study does not cover some bibliometric analysis techniques, including co-authorship, bibliographic coupling, and citation networks and maps.

#### **3.1. Data Collection**

The data collection process involved retrieving scholarly literature from reputable sources. Consequently, the Scopus database was adopted due to its extensive coverage of journals, conference proceedings, and similar types of scientific research outputs spanning a wide range of disciplines. The search query was formulated to capture publications that discuss the digital economy within the context of developing countries but exclusively in the English language.

### 3.2. Inclusion and Exclusion Criteria

Publications included in the analysis are those that explicitly address the digital economy's impact, trends, challenges, and opportunities within developing countries. The Scopus database was searched using the keyword phrases: “Digital Economy” AND “Developing country” and their common variations (see appendix 1 for the full search string). Journal articles and conference papers were incorporated, while editorials, letters, and short communications were excluded. Non-English publications were also excluded to ensure consistency in data analysis and interpretation, considering that the three researchers involved in this study only read and write in the English language. Furthermore, the time frame considered for data retrieval covers January 2003 to July 2023.

### 3.3. Data Processing and Analysis

The data extraction process involved collecting essential information for each publication, including author names, affiliations, publication year, keywords, citation counts, and references, among others. This data was then organised in a structured dataset to ensure the subsequent analysis in Microsoft Excel and VOSviewer were seamless and accurate. Microsoft Excel was chosen because of its inherent ability to record the count of publication entries in rows and columns, aggregate and return subsequent results in pivot tables, and present the data in clear, simple visualisation charts and graphs to aid the assimilation of the insights derived. VOSviewer was chosen because of its programmed ability to quantify, aggregate, and identify not-so-obvious patterns inherent in the retrieved data.

The analysis encompassed several key dimensions:

- **Journal and Publication Patterns:** Preferred journals for publishing research in the English language related to the digital economy in developing countries were determined. This analysis revealed the dissemination channels for scholarly output in this domain.
- **Document Types and Subject Area Distribution:** The documents retrieved were analysed to determine the types of scholarly output researchers are publishing in the English language within the field, as well as the subject areas that they are focusing their work.
- **Geographical Distribution:** The geographic distribution of publications in the English language was examined to identify countries and regions with significant contributions

to the field. This analysis highlights regions that play a pivotal role in shaping research discourse in the domain.

- **Publication and Author Impact and Trends:** An analysis was conducted to observe the progression of publications in the English language across time, thereby discerning prevailing patterns in research output. Concurrently, the scrutiny extended to discovering prominent authors and their corresponding citation volumes. This endeavour illuminates the extent of influence and impact that specific papers hold within the discourse surrounding the digital economy, thus, effectively identifying the contributors who, by virtue of their publication frequency and citation impact, substantively enrich the field.
- **Keyword and Term Co-occurrence Analyses:** Keywords and terms associated with publications in the English language were analysed to uncover prevailing research themes and shifts in focus over time. This analysis provides insights into the evolving topics within the digital economy domain.

### **3.4. Validity and Limitations**

The validity of the findings is reinforced by the utilisation of a comprehensive and reputable database – Scopus. On the other hand, the analysis' limitations include potential database biases, publication lag, and the exclusion of non-English publications. These limitations might restrict the generalisability of the findings.

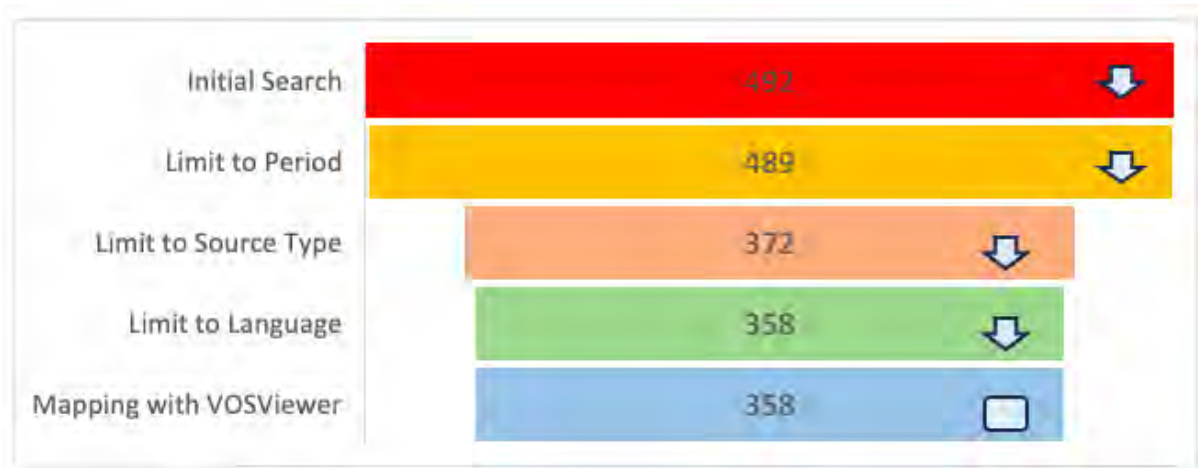
### **3.5. Ethical Considerations**

This study adhered to ethical guidelines by using publicly available data from reputable sources. Proper citation and acknowledgment of authors' contributions are maintained throughout the analysis.

### **3.6. Search Strategy**

Figure 1 outlines the strategy adopted to retrieve the data from Scopus. The search was conducted on the 21<sup>st</sup> of August 2023. The search terms, “Digital Economy” and “Developing Countries” – and their respective variations, were used to trigger the database to produce all the papers that fit the inclusion criteria.





**Figure 1: Bibliometric Analysis Search Strategy**

To summarise, the bibliometric analysis is designed to provide a thorough examination of the digital economy's scholarly landscape within the context of developing countries. By scrutinising publication trends, authorship patterns, document types and preferred subject areas, keyword evolution, journal preferences, and geographical distribution, this analysis seeks to uncover insights that contribute to a holistic understanding of the field's dynamics and impact.

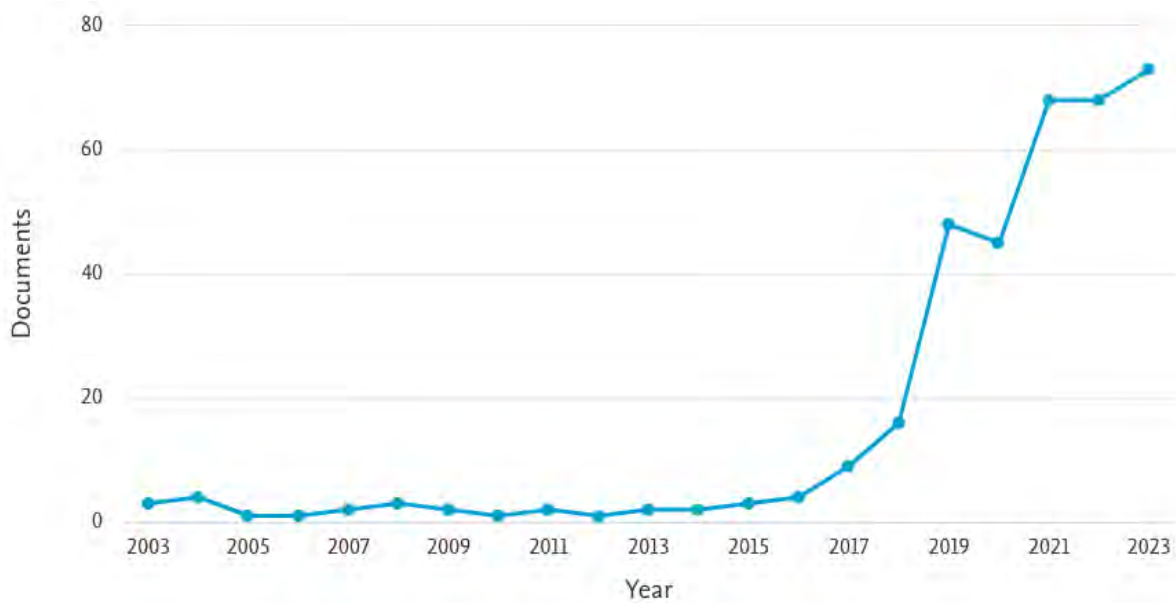
## **4. Findings**

After conducting a search on Scopus, a total of 358 papers meeting the predefined inclusion criteria were identified and retrieved. Subsequently, a combination of software tools, including Microsoft Excel, Scopus, SCImago Journal Ranking, and VOSviewer were utilised to conduct comprehensive statistical analyses of the dataset. Additionally, these tools were used to generate pertinent network maps and visualisations tailored to specific analytical requirements.

### **4.1. Annual Publication Trend in Scopus**

Commencing in 2003, a consistent and noteworthy upward trajectory in the volume of scientific articles was observed. This encompassed both journal and conference papers published in the English language and dedicated to examining the impact of the digital economy on developing countries. In the inaugural year, 2003, a modest tally of three papers marked the inception of this scholarly journey. Subsequently, there was a remarkable surge in 2019, with an impressive count of 48 articles published within a twelve-month span.

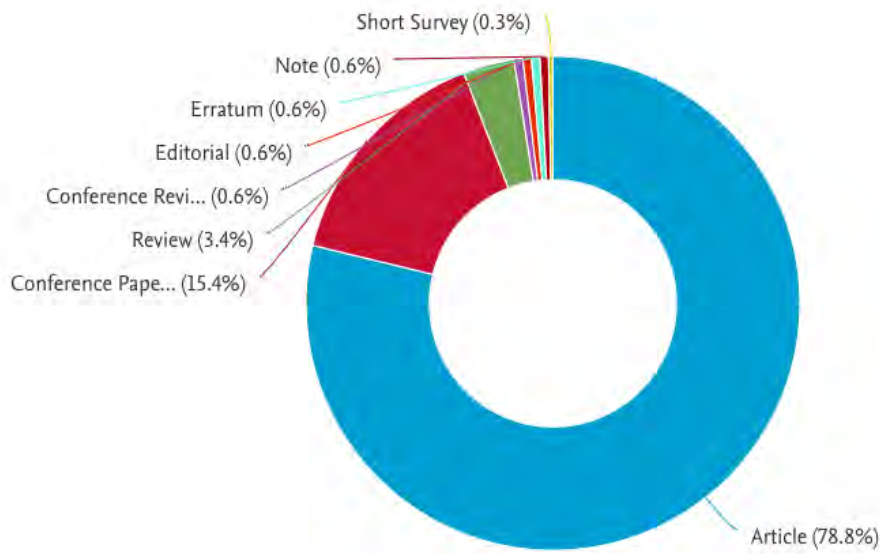
This ascent continued undeterred in the ensuing years, exemplified by the publication of 68 papers in both 2021 and 2022, respectively, thereby signifying a sustained and robust interest in research within this domain. The trend remains resolute in 2023, as evident from the output of 73 documents published from January 1<sup>st</sup> to July 31<sup>st</sup>, underscoring the enduring momentum of scientific contributions. The time series chart in Figure 2 provides a visual representation of this upward trajectory.



**Figure 2: Document Publication Trend from 2003 to 2023**

#### **4.2. Document Types**

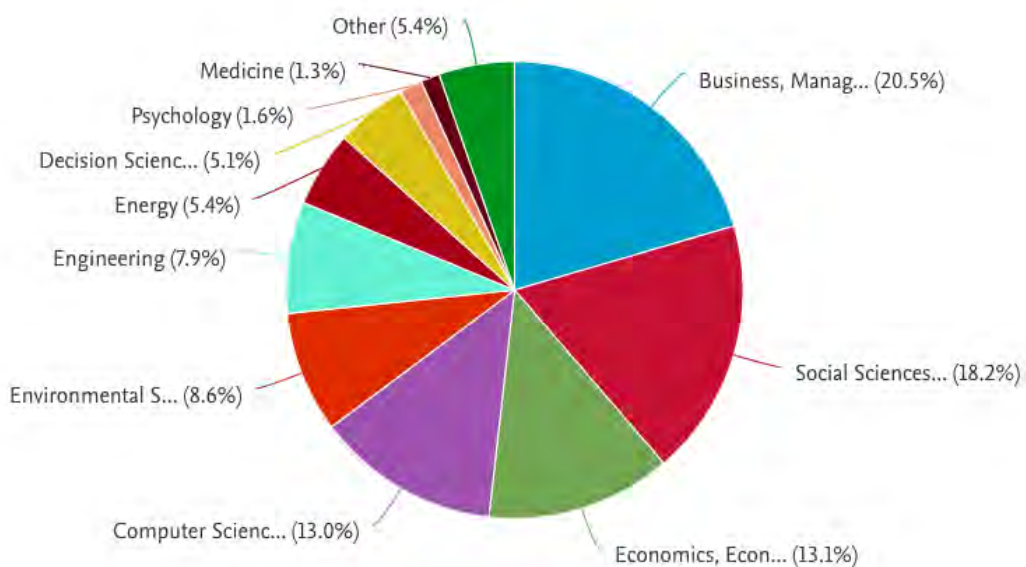
Among the 358 documents obtained from the Scopus database, a significant proportion – 78.8%, or 282 papers – comprise journal articles. On the other hand, 15.4% of the retrieved documents, totalling 55 papers, fall under the category of conference proceedings. These two primary document classifications collectively account for over 90% of the total papers published in the English language identified as dedicated to investigating the digital economy in developing nations. Figure 3 illustrates the distribution in a pie chart.



**Figure 3: Types of Documents Published**

### 4.3. Subject Area Distribution

The extracted data was subjected to a more granular examination based on subject area classifications, with the aim of gaining deeper insights and enhancing our understanding of the specific domains that the authors researching the digital economy in emerging markets are inclined towards. The resulting analysis is visually represented in Figure 4, a pie chart, which succinctly conveys the distribution of subject areas.



#### **Figure 4: Subject Area Distribution of Published Papers**

According to the chart, the predominant subject areas of interest for researchers in this field are as follows:

- Business, Management, and Accounting, constituting 20.5% of the research.
- Social Sciences, encompassing 18.2% of the research.
- Economics, Econometrics, and Finance, accounting for 13.1% of the research.

These findings underscore the multidisciplinary nature of digital economy research in emerging markets, with a pronounced emphasis on subjects related to business, management, social sciences, and economics.

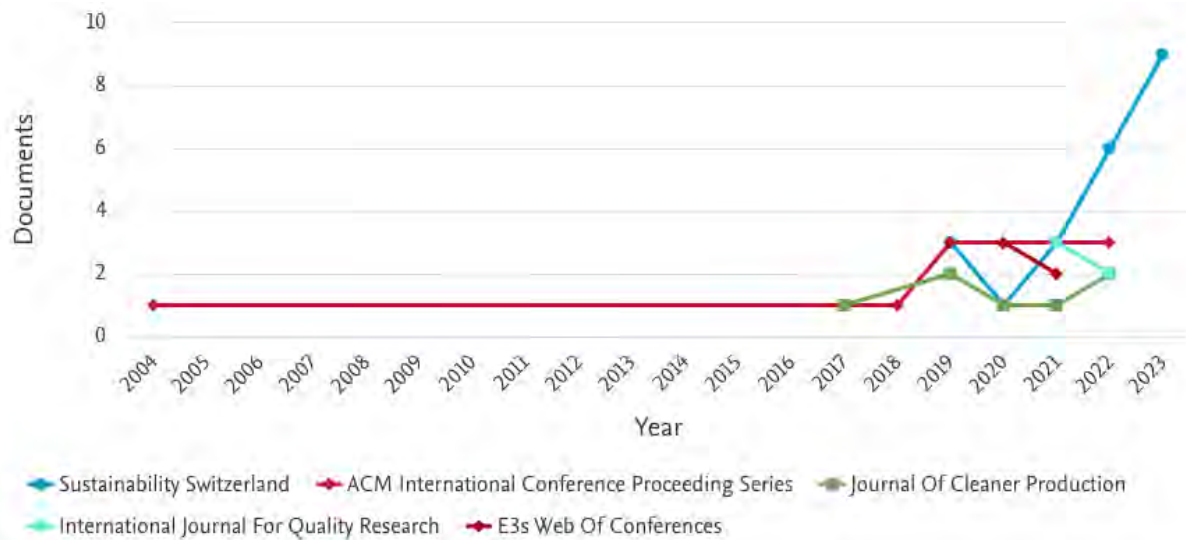
#### **4.4. Top Publication Sources**

In alignment with the previously identified trend, there is a discernible surge in research and scholarly publications concerning the digital economy within the context of developing countries. As highlighted in Figure 5, the journal 'Sustainability Switzerland' serves as a notable example of this burgeoning interest.

Remarkably, there were no recorded publications in 'Sustainability Switzerland' until the year 2019, when the journal featured three papers addressing this subject matter. This trend, however, has shown remarkable growth, considering that the number of papers published in the journal surged to nine by July 31<sup>st</sup>, 2023, for that half-calendar year alone. In total, between 2019 and 2023, there have been 22 papers published in 'Sustainability Switzerland' on this topic.

It is noteworthy that, of the 22 papers published in 'Sustainability Switzerland' within this timeframe, a substantial portion, precisely 54% or 12 papers, have centered their focus on the digital economy in China. This underscores the prominence of China within the discourse on the digital economy's role in developing countries.

Furthermore, it is essential to acknowledge that several other journals have also contributed to the dissemination of research on this subject matter. These journals include, but are not limited to, the 'ACM International Conference Proceeding Series,' 'Technological Forecasting and Social Change,' and the 'International Journal for Quality Research.'



**Figure 5: Journal Publication Trend Comparison**

Subsequently, a SCImago Journal Rank (SJR) analysis was used to assess the standing and potential impact of the journals and conference proceedings publishers through whom authors disseminate their research in the English language. The SJR metric serves as an evaluative instrument, adept at gauging the prestige and influence of scholarly journals through a meticulous examination of their citation patterns. Its overarching goal lies in providing a nuanced and contextually sensitive evaluation of a journal's prominence, differentiating itself from conventional metrics like the Journal Impact Factor (JIF).

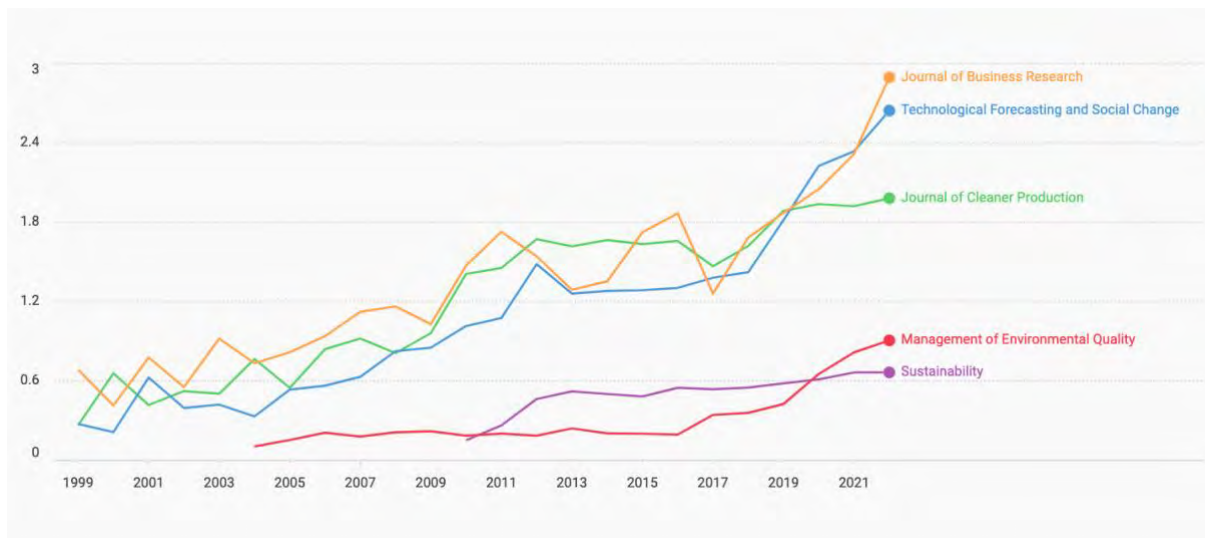
The calculation of the SJR indicator extends its purview beyond a mere tally of citations attributed to a journal. Conversely, it takes into account the caliber and relevance of the sources that cite the journal. In essence, the SJR metric offers a comprehensive lens through which we can holistically grasp the intricate interplay between a journal's citation impact and the intrinsic quality of the sources referencing it. The following table presents the top 10 journals and conference proceedings publishing entities that were subjected to evaluation in this research. These entities are ranked based on the number of papers they have published in the domain under review, ranked in descending order from the highest to the lowest count.

<b>S/N</b>	<b>Source</b>	<b>2022 SCImago Journal Rank (SJR)</b>	<b>Country of Publication</b>	<b>Number of Papers Published</b>
<b>1</b>	Sustainability Switzerland	0.664	Switzerland	22
<b>2</b>	ACM International Conference Proceeding Series	0.209	United States of America	8
<b>3</b>	Journal Of Cleaner Production	1.981	United Kingdom	7
<b>4</b>	Technological Forecasting and Social Change	2.644	United States of America	7
<b>5</b>	E3s Web of Conferences	0.182	France	5
<b>6</b>	International Journal for Quality Research	0.296	Serbia	5
<b>7</b>	International Journal of Trade and Global Markets	0.218	Switzerland	5
<b>8</b>	Journal of Business Research	2.895	United States of America	4
<b>9</b>	Management Of Environmental Quality	0.906	United Kingdom	4
<b>10</b>	Bulletin For International Taxation	0.141	Netherlands	3

**Table 1: Journal Ranking on The SCImago Scale**

Notably, scientific discourse concerning the digital economy within the context of developing nations predominantly gravitates towards publication outlets hosted in developed countries. However, there exists an exception to this prevailing pattern depicted in Table 1, This outlier is exemplified by the International Journal for Quality Research, which is headquartered in Serbia – a nation generally categorised as 'developing.'

Moreover, to facilitate a succinct overview of the prestige and potential impact of the leading five journals, a comparative representation in Figure 6 is provided. The analytical criterion employed herein rests upon the most current SJR scores (at the time of researching and writing) attributed to the respective publishing entities. The trend line below corroborates Table 1, showing clearly that the Journal of Business Research, published in the USA, with an SJR score of 2.895, is the most impactful among the journals in the dataset used for conducting this bibliometric analysis.

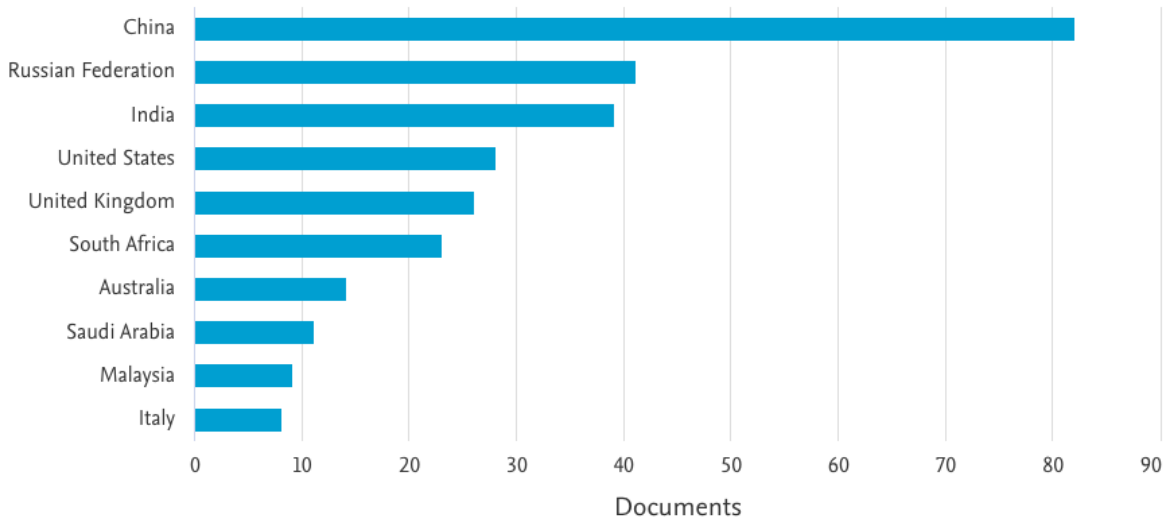


**Figure 6: SCImago Journal Rank Scores and Trends of Respective Journals**

#### **4.5. Country or Territory with the Highest Number of Research Publications**

The analysis of documents extracted from the database included a segmentation based on the affiliations of the lead authors. This segment aims to identify the country or territory responsible for the highest volume of scientific publications concerning the digital economy within the context of developing nations. Figure 7 provides a comprehensive depiction of the top 10 countries where these lead authors are based. This chart not only serves as an indicator

of the regions contributing significantly to the scholarly discourse but also sheds light on the locales where primary data collection and empirical investigations have been conducted to elucidate the intricate relationship between the digital economy and developing countries.



**Figure 7: Country or Territory with The Most Research Publications**

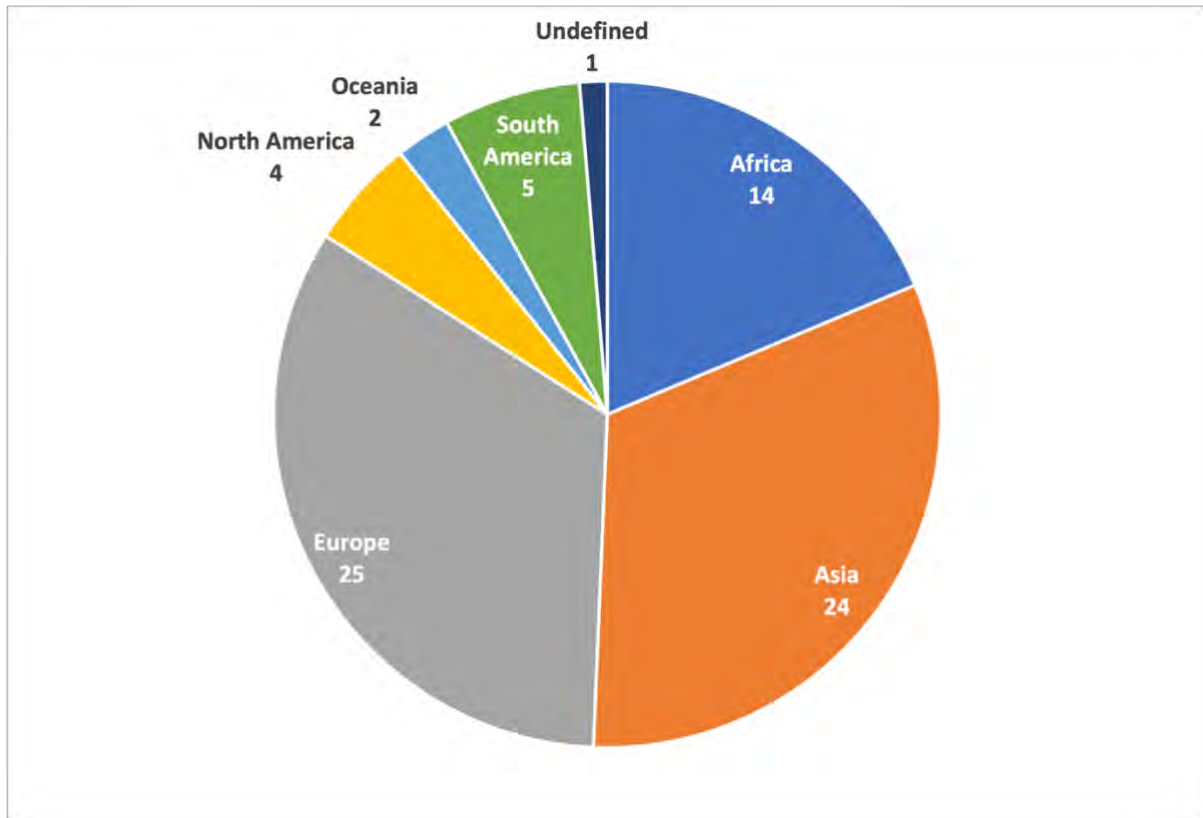
China emerges as the foremost contributor in this endeavour, with a notable total of 83 documents. This figure slightly surpasses the combined contributions of the subsequent two leading nations, namely Russia (41 documents) and India (39 documents). Rounding out the top five positions are the United States and the United Kingdom, with 28 and 26 documents, respectively. The result shows some interesting outcome: China and Russia predominantly speak and write in Chinese Mandarin and Russian, respectively, yet they produce more papers published in the English language than other English-speaking developing countries, like India, for example.

#### **4.6. Regional Distribution of Publications**

The extracted data underwent further scrutiny, with a meticulous country-level analysis aimed at identifying the geographic regions and count of countries of respective regions where empirical research pertaining to the digital economy in developing nations has been actively conducted. Figure 8, presented in the form of a pie chart, provides a succinct overview of the cumulative regional contributions to the body of knowledge in this field.

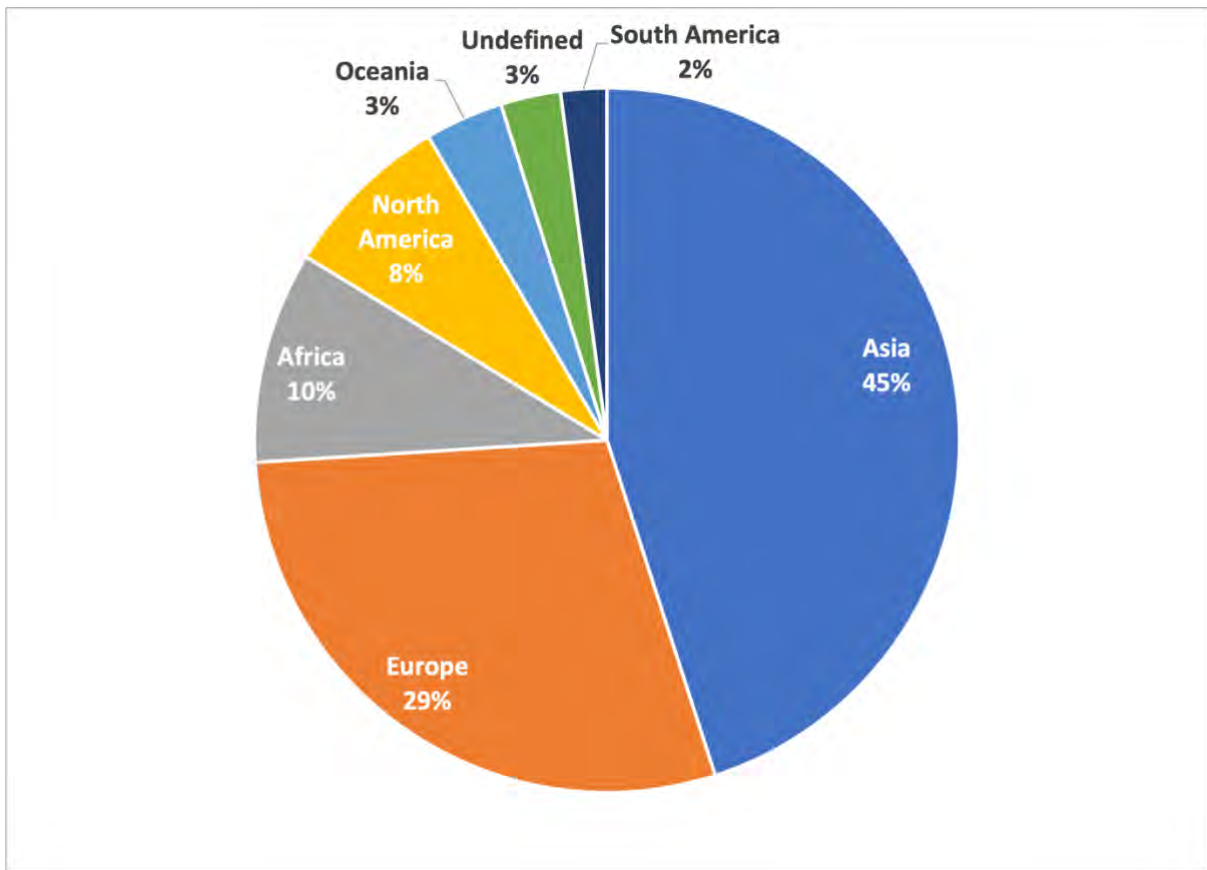


The chart reveals that the majority of countries engaged in this research area are situated within Europe (25 countries) Asia, (24), and Africa (14), respectively. In contrast, Oceania exhibits the lowest number of participating countries, represented by only 2 countries. North America and South America exhibit relatively similar patterns, represented by 4 and 5 of the total countries, respectively.



**Figure 8: Geographic Regions and Count of Countries with Publications**

Additionally, a detailed analysis was conducted to determine the regions wielding the most significant influence in terms of the aggregate volume of papers published throughout the reviewed period. Figure 9 serves as a visual representation of this analysis, portraying the percentage distribution of published papers categorised by continental regions. It is imperative to acknowledge that a higher volume of empirical research in this domain translates into a more extensive reservoir of data, ultimately enriching our comprehension of the intricate interplay between the digital economy and the diverse developmental factors and challenges encountered by respective regions.



**Figure 9: Regional Distribution of Aggregated Volume of Published Papers**

Figure 9 illustrates the distribution across various continental regions. Asia emerges as the dominant contributor, accounting for 45% of all publications between 2003 and 2023. Europe follows closely, albeit with a lower share of publications, at 29%. Africa registers a notable but comparatively smaller contribution, standing at 10%. North America, Oceania, and South America collectively exhibit fewer publications, constituting 8%, 3%, and 2%, respectively, indicating a lesser research focus on this topic within these regions.

In summary, with respect to regional spread, Asia leads in both the volume of publications and the assortment of countries involved, underscoring its prominence as a key region in terms of research output in the domain. Europe also maintains a relatively high volume of publications, although it encompasses fewer countries than Asia. Africa, despite being home to over 50 countries, about half of which are recognised as English-speaking, and the vast majority classified as developing countries, exhibits a smaller representation in terms of publication count. Meanwhile, North America, Oceania, and South America contribute fewer publications from fewer nations when juxtaposed with the paper count from Asia and Europe.

#### 4.7. Top Contributing Authors

The examination of the digital economy within the framework of developing countries has attracted scholarly attention from both individual authors and collaborative co-authorship teams. To ascertain the author(s) wielding the most substantial impact within this field, an analysis of documents published in the English language sourced from the Scopus database was conducted, categorising them by author names. The findings revealed that Bogoviz Aleksei V., an independent researcher from Moscow, Russia, has emerged as the most prolific contributor, having authored the highest number of articles – four in total. Notably, three of these papers were published in the year 2021 alone, with an additional one published in 2023. Specifically, the author engaged in collaborative efforts, co-authoring three papers (Alekseev et al., 2021; Bogoviz et al., 2021; Popkova et al., 2023), and singlehandedly authored one (Bogoviz, 2021), thereby reaffirming their noteworthy impact within this scholarly domain. Figure 10 illustrates the top 10 contributors to the subject – ranked by highest to lowest volume.

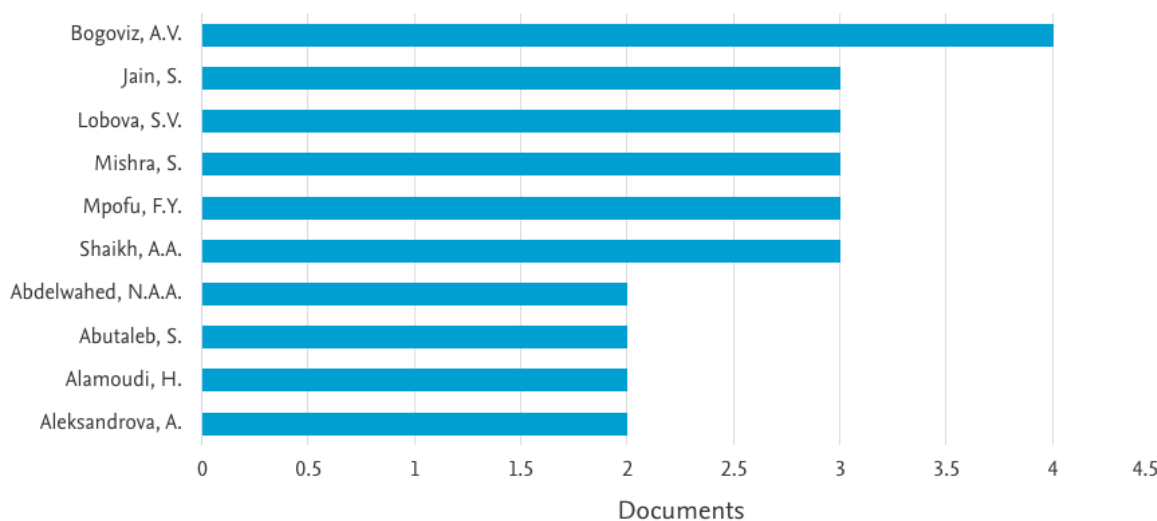


Figure 10: Top Ten Contributing Authors

#### 4.8. Most Cited Publications

In pursuit of a comprehensive understanding of the impact that authors have made in the realm of the digital economy within developing nations, the analysis also sought to identify authors whose paper(s) have garnered the highest number of citations. As elucidated by Bihari et al. (2023), the H-index emerges as a reliable indicator of an author's impact within their specific domain of expertise, shedding light on the influence their work has exerted in the scholarly community.

The analysis unveiled that the paper authored by Kumar et al. (2018) has garnered a remarkable 240 citations, underscoring its significant impact within the field. Regardless, it is noteworthy that out of the 358 documents considered, only 24, comprising 6.7% of the total, have crossed the threshold of 50 or more citations, signifying their noteworthy influence while, on the other hand, also indicating that many papers in this field are yet to gain co-citation traction. The top 10 most cited publications are visually presented in Table 2, offering a succinct overview of the contributions that have attracted substantial recognition within the scientific discourse. The table is sorted by the highest number of citations to the lowest.

<b>S/N</b>	<b>Document Title</b>	<b>Author(s) and Year of Publication</b>	<b>Source</b>	<b>Number of Citations</b>
1	A strategic framework for a profitable business model in the sharing economy	Kumar et al. (2018)	Industrial Marketing Management	240
2	Ride-hailing, travel behaviour and sustainable mobility: an international review	Tirachini (2020)	Transportation	149
3	Innovation in emerging economies: Research on the digital economy driving high-quality green development	Ma and Zhu (2022)	Journal of Business Research	135
4	The Global Platform Economy: A New Offshoring Institution Enabling Emerging-Economy Microproviders	Lehdonvirta et al. (2019)	Journal of Management	127
5	Upsides and downsides of the sharing economy: Collaborative consumption business models'	Dreyer et al. (2017)	Technological Forecasting and Social Change	117

	stakeholder value impacts and their relationship to context			
6	First-mover firms in the transition towards the sharing economy in metallic natural resource-intensive industries: Implications for the circular economy and emerging industry 4.0 technologies	Chiappetta Jabbour et al. (2020)	Resources Policy	85
7	Rapid expansion of international new ventures across institutional distance	Deng et al. (2018)	Journal of International Business Studies	85
8	Chinese travelers' behavioral intentions toward room-sharing platforms: The influence of motivations, perceived trust, and past experience	Wu et al. (2017)	International Journal of Contemporary Hospitality Management	83
9	Determinants of consumers' participation in the sharing economy: A social exchange perspective within an emerging economy context	Boateng et al. (2019)	International Journal of Contemporary Hospitality Management	75
10	Chinese culture and e-commerce: An exploratory study	Efendioglu and Yip (2004)	Interacting with Computers	74

**Table 2: Top 10 Most Cited Authors and Their Publication Titles**

#### 4.9. Top Funding Sponsors

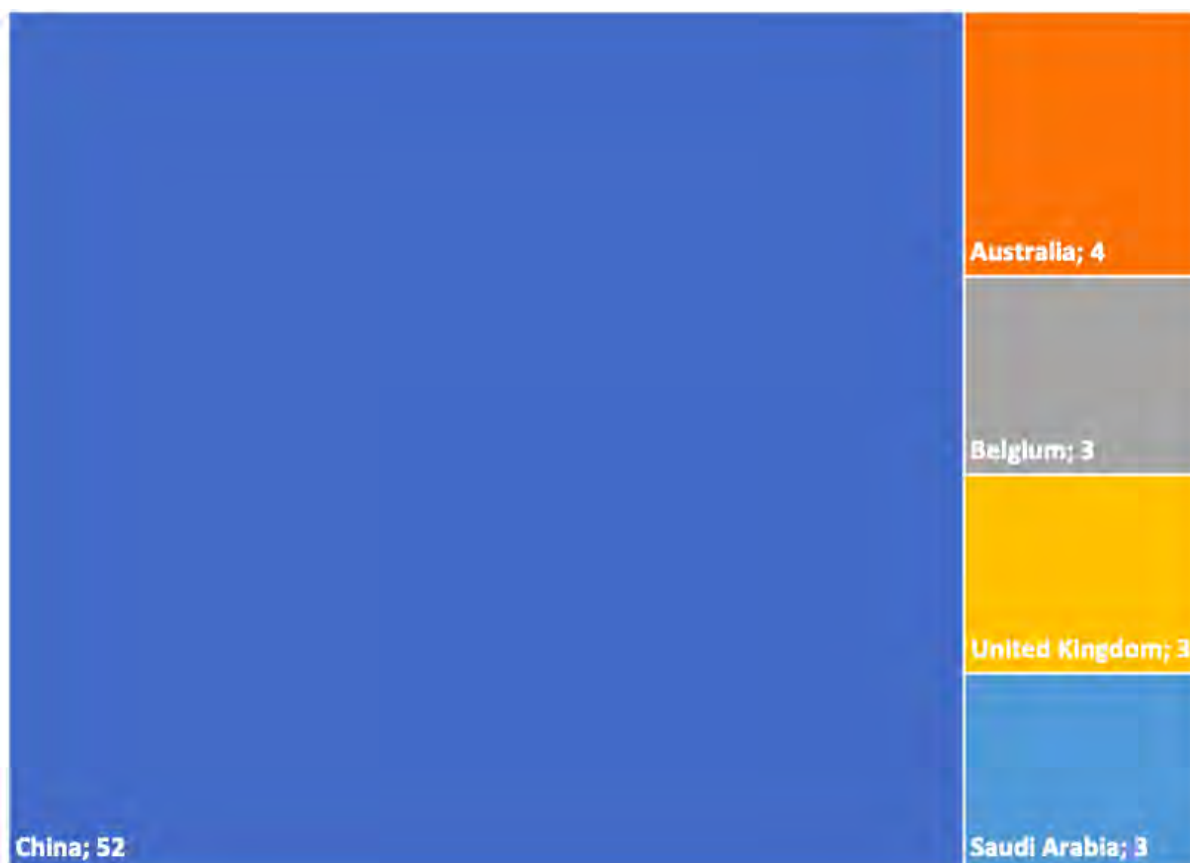
In a manner akin to the examination of influential authors and impactful papers concerning the digital economy within the context of developing countries, an analysis to unveil the funding sponsor(s) that have played a significant role in facilitating research endeavours was conducted. The outcomes, clearly ranked in descending order based on the number of publications they have supported, reveal that the ‘National Natural Science Foundation of China’ emerges as the preeminent funding sponsor, having commissioned or supported the research and publication of 24 documents in just seven years – 2017 to 2023. It is noteworthy that, among the top 10 funding sponsors, institutions based in China occupy the top four positions. Table 3 provides an illustrative overview of the funding sponsors that have earned a spot on the list.

S/N	Funding Sponsor	Country of Sponsor	Number of Publications
1	National Natural Science Foundation of China	China	24
2	National Office for Philosophy and Social Sciences	China	12
3	Fundamental Research Funds for the Central Universities	China	9
4	Ministry of Education of the People's Republic of China	China	5
5	Deanship of Scientific Research, King Saud University	Saudi Arabia	3
6	Economic and Social Research Council	United Kingdom	3
7	European Commission	Belgium	3
8	Australian Research Council	Australia	2
9	China Postdoctoral Science Foundation	China	2

10	Commonwealth Scientific and Industrial Research Organisation	Australia	2
----	--	-----------	---

**Table 3: Top 10 Funding Sponsors and Their Locations**

To provide a more contextual understanding, a tree map analysis to ascertain the collective number of publications commissioned by the top 10 funding sponsors was carried out. This was done by aggregating the funding sponsors by their respective country-locations as single entities. Figure 11 presents the outcome of this analysis, indicating that research funding sponsors situated in China, in combination, commissioned a total of 52 documents. This figure notably surpasses the cumulative total of the next four countries, combined, with a ratio four times greater. This underscores the substantial lead of China-based funding sponsors in terms of research support in this domain – even though the country’s major language is Mandarin and not English language.



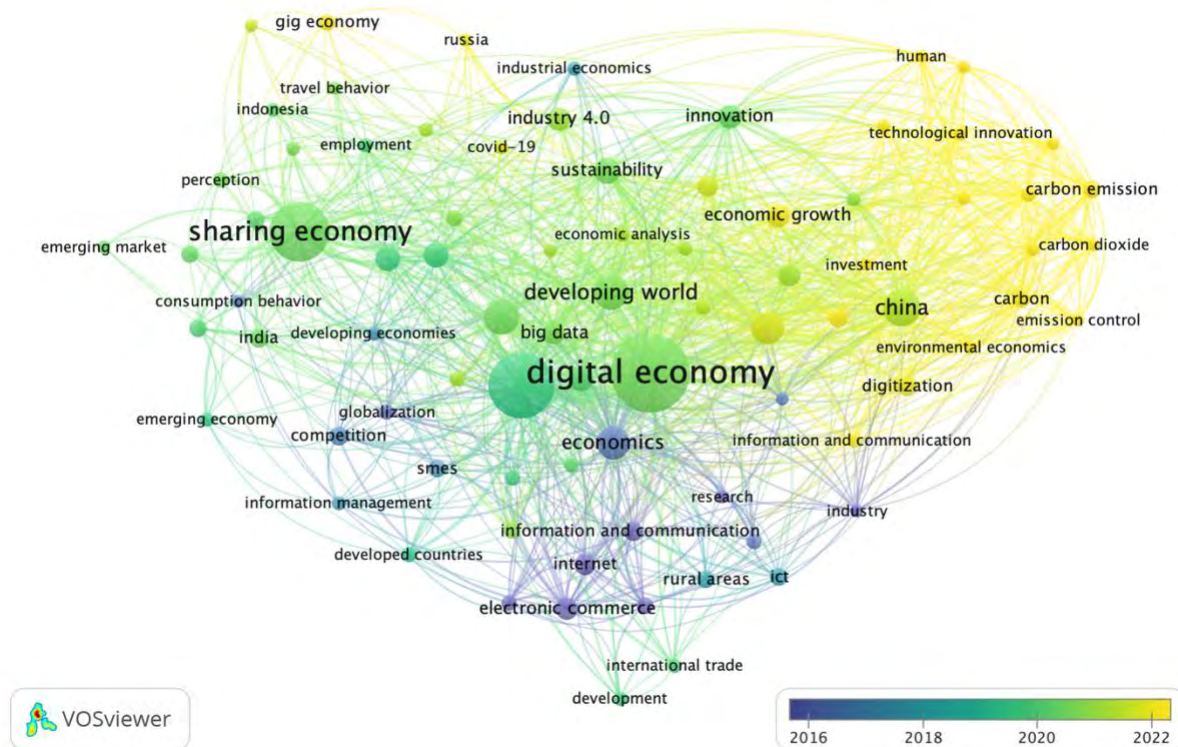
**Figure 11: Country-Analysis of Aggregate Count of Papers Commissioned by Funding Sponsors**



#### 4.10. Keyword Co-occurrence Map

In order to visualise the prevalent author keywords and index keywords chosen by both authors and publishers, respectively, and to delineate the interrelationships between them, a Keyword Co-occurrence analysis was undertaken. This analysis not only offers insights into the prevailing trends in keyword selection but also sheds light on the evolving popularity of specific keywords within this field.

For the purpose of this analysis, a minimum threshold was set to focus on keywords that appeared at least 5 times across the 358 papers downloaded. Out of the 2,096 identified keywords, only 82 met this threshold. In Figure 12, the color scheme employed signifies that keywords and phrases shaded in purple and blue are waning in popularity, while those tinted green and yellow are gaining prominence among authors in this domain.



**Figure 12: Keyword Co-occurrence Map**

The findings from the analysis indicates a substantial interest in research at the intersection of the digital economy and sustainability. Keywords and phrases highlighted in yellow, such as ‘carbon,’ ‘environmental economics,’ ‘emission control,’ ‘digitisation,’ and ‘gig economy,’ among others, are indicative of the growing areas of focus for researchers in the years 2021/22



and beyond. Conversely, keywords and phrases like ‘competition,’ ‘SMEs,’ and ‘electronic commerce,’ among others, which were more prevalent before 2017, appear to be waning in popularity.

Additionally, Table 4 shows the most frequently occurring keywords, along with the number of times they appear in the titles and/or abstracts of the 358 papers sourced from the search results.

<b>S/N</b>	<b>Keyword</b>	<b>Number of Occurrences</b>
<b>1</b>	Digital economy	118
<b>2</b>	Developing countries	87
<b>3</b>	Sharing economy	75
<b>4</b>	China	33
<b>5</b>	Developing world	30
<b>6</b>	Sustainable development	30
<b>7</b>	Economics	27
<b>8</b>	Economic development	26
<b>9</b>	Sustainability	17
<b>10</b>	Economic and social effects	16

**Table 4: Count of the Most Frequently Occurring Keywords**

#### **4.11. Term Co-occurrence Map**

In Figure 13, a Term Co-occurrence map is presented, which is a visual aid to facilitate the comprehension of the prevailing themes appearing in the authors' titles and abstracts, and the interconnections between them. To ensure a focused representation, a minimum threshold for term inclusion was established, i.e., only terms that appeared at least 10 times within the 358 documents extracted were considered. Out of the 8,723 identified terms, 209 met this threshold.

Notably, the VOSviewer software, by default, incorporates 60% of the most pertinent terms into the visualisation, thus ensuring optimal visibility and legibility for human observers. Consequently, the software analysed a total of 125 terms.

This analysis serves to illuminate the thematic trends characterising research into the digital economy within the context of developing nations. The color scheme employed segregates the terms into distinct clusters. The weight (defined as the degree to which the term is mostly used), as determined by the VOSviewer software, serves as the unit of analysis. A higher weight, tending towards 1.4, signifies that the given word or phrase is becoming more common in the texts. As such, terms colored in green and yellow represent emerging themes that authors are increasingly embracing. Noteworthy themes emerging from this analysis include ‘China,’ ‘panel data,’ ‘technological innovation,’ ‘effect,’ ‘platform,’ ‘sustainability,’ and ‘sharing economy,’ reflecting the contemporary focal points of current research. It is imperative to note that while words and phrases shaded in purple and blue with weights leaning towards 0.6, such as ‘education,’ ‘competitiveness,’ ‘transition,’ and ‘infrastructure,’ among others, may exhibit a relatively declining trend, they continue to maintain relevance.

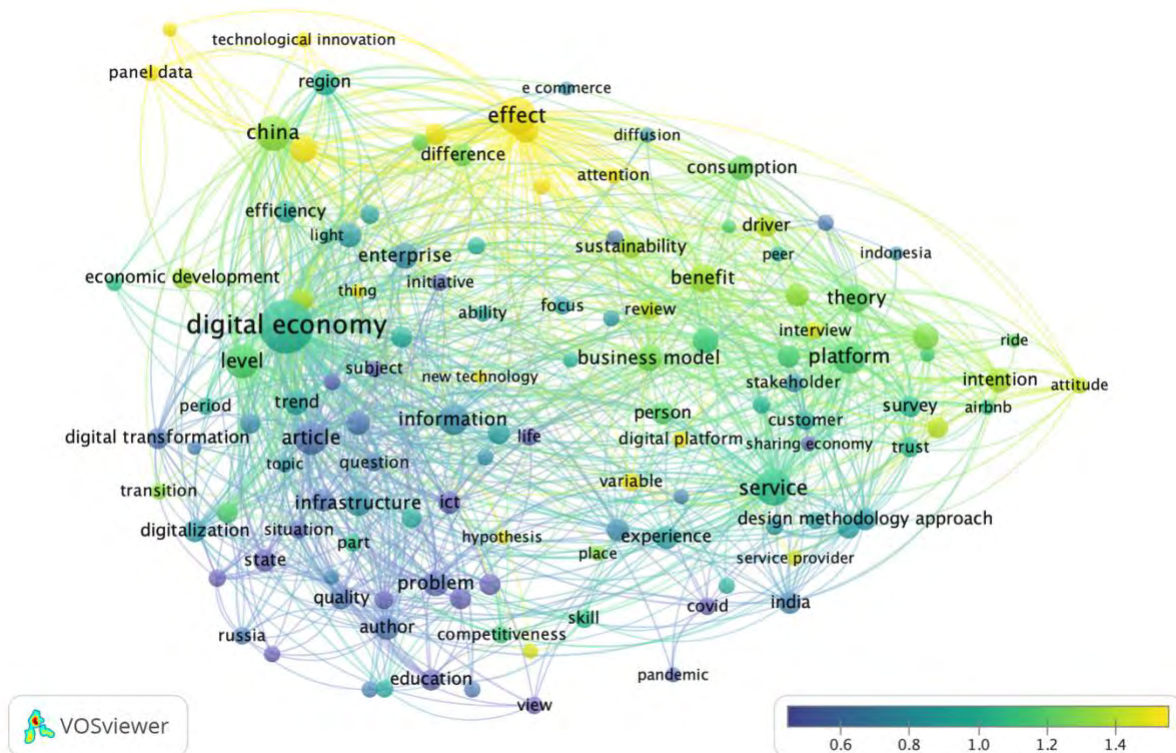


Figure 13: Term Co-occurrence Map

## 5. Discussion

This study conducted a comprehensive bibliometric analysis of scientific publications in English language related to the digital economy within the context of developing countries, yielding several noteworthy insights and implications for researchers, policymakers, and practitioners.

The analysis illuminated the relatively young and rapidly evolving nature of research in this domain, with a significant surge in publications, particularly from 2019 onwards. The dominance of journal articles (78%) over and above conference proceedings papers (15%) suggests a preference for disseminating research outcomes through academic journals. This outcome potentially indicates limited avenues for collaborative discussions regarding the digital economy in the framework of developing nations. Furthermore, the analysis revealed that researchers predominantly approach the digital economy in developing countries through the lenses of business and management, social sciences, and economics, demonstrating a clear understanding of the profound relationship between digital technologies and society at large.

Furthermore, a notable observation is that the most productive and influential authors, institutions, and journals primarily originate from developed countries in the Global North, including the USA, UK, and Switzerland, with China being a notable exception. This highlights the imperative for greater involvement and contributions from researchers in developing countries who possess firsthand knowledge of local contexts and challenges. Additionally, this accentuates the call to action for funding sponsors, governmental and development bodies, universities, and research institutions to allocate resources for data-driven empirical fieldwork that can deepen our understanding of the interplay between digital technologies, societal constructs, labour dynamics, and economic growth.

The study also shed light on funding sponsors in this domain, revealing a prominent role played by China-based institutions, with the Asian giant sponsoring 14.5%, (or 52) of the 358 publications examined. This mirrors China's leadership and experience in digital technology development, deployment, and evaluation, as well as its commitment to sustainability – as evidenced by the research interests of the publications sponsored. Curiously, though, while people in China predominantly speak and write in Chinese Mandarin, their researchers have produced more digital economy-related papers published in the English language than any other English-speaking developing country, like India, for example.

Additionally, the study identified prevalent keywords and topics within the field, highlighting the importance of subjects like innovation, carbon emissions, environmental economics, and investment. Nevertheless, it drew attention to underexplored areas requiring equal consideration, including digital inclusion, digital divide, digital literacy, digital policy, digital governance, digital ethics, and digital security. These topics are crucial for ensuring equitable distribution of digital economy benefits and mitigation of potential risks, respectively.

Relatedly, the visual representations provided insights into research trends and emerging topics. For instance, the Term Co-occurrence analysis showcased the evolution of research themes, from early foci on information management and consumption behaviour to contemporary exploration of complex topics such as digital transformation, innovation, sustainability, and energy efficiency.

Regardless, the geographical analysis pointed to an uneven distribution of publications across different regions and countries within developing nations, potentially exacerbating disparities in research capacity, funding, infrastructure, policy, and culture. To address this, there is a pressing need for more inclusive and regionally balanced research that caters to the specific needs and contexts of diverse regions and countries.

Finally, the low citation impact and visibility of numerous publications and journals within this field, as revealed in this study, may suggest several plausible factors, including a potential lack of quality or relevance in some publications or journals, as well as a possible deficit in awareness and recognition of the research contributions and implications of the digital economy within the context of developing nations. Considering the pivotal role that digital technologies can play in the economic development of developing nations, there arises a compelling need for research that not only adheres to rigorous academic standards but also addresses pertinent issues of relevance.

In conclusion, this bibliometric analysis offers valuable insights into the dynamic landscape of research published in the English language on the digital economy in developing countries, highlighting opportunities for more equitable and impactful research, collaboration, and dissemination of knowledge. Consequently, the field would benefit immensely from more robust, pertinent research that aligns with the expectations of the academic community, while

simultaneously ensuring more effective dissemination and communication of research findings and recommendations to key stakeholders and other interested audiences.

## **6. Limitations of The Study**

This study possesses several limitations that not only merit acknowledgment, but more importantly, serve as avenues for further exploration in future research endeavours.

Firstly, this study exclusively relied on Scopus as the primary data source for conducting the bibliometric analysis. While Scopus represents one of the most comprehensive and widely utilised databases for scientific publications, it may not encompass the entirety of publications in the English language or journals within the field of digital economy in developing countries. Future research efforts may seek to broaden the scope by incorporating other databases or sources, including but not limited to Web of Science, Google Scholar, and regional or localised databases, to complement and cross-verify the outcomes presented here.

Secondly, this study employs a predetermined set of keywords and search terms, as detailed in Appendix 1, to identify publications associated with the digital economy in developing countries. However, it is worth recognising that these keywords and search terms may not encompass all conceivable variations or synonyms pertaining to the digital economy concept nor do they capture the diverse definitions of the phrase, developing countries. Subsequent research endeavours may explore alternative or supplementary keywords and search terms to refine or expand the scope of investigation.

Thirdly, the analysis within this study leverages a blend of metadata aggregation and qualitative methods and techniques for data analysis and visualisation. Yet, it is essential to acknowledge the potential limitations or biases that may arise concerning data quality, validity, reliability, accuracy, or interpretation inherent to these methods and techniques. Future research initiatives might consider alternative or additional approaches to corroborate and augment the analysis and visualisation of retrieved data.

Fourthly, this study exclusively employs the VOSviewer software for keyword and term occurrence analyses only. Consequently, it does not focus on co-authorship, co-occurrence, co-citation, or bibliographic coupling analyses often executed utilising the tool. Subsequent researchers could delve into these dimensions, potentially conducting meta-analyses to offer further depth and richness to scholarship in this domain.

Lastly, the study exclusively considers research papers published in the English language. This limitation is significant, given that several regions of the world hosting developing countries often communicate in languages other than English. For instance, South America predominantly employs Spanish and Portuguese, while a good number of African countries communicate in Swahili, French, Portuguese, Arabic, and other languages. As a result, the study may have overlooked relevant papers published in languages other than English. Future research endeavours should consider multilingual approaches to ensure a more comprehensive exploration of the subject matter.

## **7. Future Research Direction**

Building upon the findings and implications elucidated within this study, there arises several avenues for prospective research inquiries. These directions are intended to contribute to a more nuanced and comprehensive understanding of this multifaceted domain.

Firstly, forthcoming research should embark on a thorough and extensive exploration of the underexplored or burgeoning topics identified in this investigation. Notably, areas such as digital inclusion, digital divide, digital literacy, digital policy, digital governance, digital ethics, and digital security merit heightened attention and scrutiny. Delving into these subjects with depth and breadth is crucial for grasping the intricacies and addressing the opportunities and challenges inherent in the digital economy within developing countries.

Secondly, future studies should embrace a diverse array of innovative theoretical frameworks and perspectives to dissect and explain phenomena and issues pertinent to the digital economy in developing nations. The adoption of multifaceted frameworks, such as institutional theory, stakeholder theory, resource-based view, dynamic capabilities, social network theory, social capital theory, sociotechnical framework, diffusion of innovations theory, technology acceptance models, technology-organisation-environment frameworks, or other relevant theories and models, can facilitate a more comprehensive examination of the influences exerted by various factors or actors and their reciprocating impact.

Thirdly, forthcoming research should employ other rigorous and robust empirical methodologies and techniques to collect and analyse data germane to the digital economy in developing countries. The application of versatile methodologies, such as mixed methods, systematic literature review approaches, case studies, surveys, experiments, interviews, focus

groups, observations, or other apt methods and techniques, should be harnessed to amass and scrutinise quantitative and qualitative data gleaned from diverse sources and stakeholders.

Fourthly, prospective research endeavours should delve into the contextual and comparative facets of the digital economy within developing countries. For instance, comparative analyses between different regions or countries within the developing world, or between developing and developed nations, can unveil commonalities and disparities in terms of opportunities, challenges, strategies, policies, practices, outcomes, and impacts of the digital economy.

Lastly, future research should intensify its engagement with practitioners and policymakers directly involved in the sphere of the digital economy within developing countries. Collaborative endeavours with industry partners, governmental agencies, development organisations, and civil society groups hold promise for co-creating knowledge and devising solutions to advance and enhance the digital economy's landscape in developing nations. This synergy between academia and the field can foster practical, real-world impact, and facilitate the translation of research insights into actionable strategies and policies for the betterment of these regions.

## **8. Conclusion**

This paper presents a novel and comprehensive bibliometric analysis of the digital economy in the context of developing countries, using data from the Scopus database. The paper applies various metadata aggregation methods to evaluate and assess diverse aspects of scientific publications, revealing the trends, patterns, themes, gaps, and keyword network structure of the field, as well as the most productive and influential authors, institutions, journals, and countries. The paper has also provided some visual representations of the research landscape, using various tools and techniques such as trend analysis, top publication sources, authors, and funding sponsors, among others.

To address Research Question 1 (RQ1), the findings underscore a discernible shift in research focus towards ‘sustainability’ within the field. Evidently, researchers are displaying an increasing interest in probing and comprehending how digital technology can foster sustainable economic growth in developing nations, thereby avoiding environmental harm.

In response to RQ2, RQ3, and RQ4, respectively, the study's outcomes illuminate several key facets: Sustainability Switzerland emerges as the preeminent journal of choice among authors,

the majority of papers published within the reviewed timeframe originate from China, and the National Natural Science Foundation of China turns out to be the foremost funding sponsor in this domain.

Furthermore, the investigation reveals that Bogoviz Aleksei V. has made the most substantial contribution in terms of the number of papers published, firmly addressing RQ5. Notably also, Kumar et al. (2018) emerges as the most cited author in the field.

In addressing RQ6, the study identifies prevalent keywords and recurring themes. Prominent keywords include ‘China,’ ‘sharing economy,’ and ‘sustainable development,’ while recurrent themes include ‘technological innovation,’ ‘efficiency,’ ‘digital platforms,’ and ‘sustainability.’ These insights encapsulate the evolving landscape of research within the reviewed period.

The paper further discussed the implications of the results and proposed directions for future research, such as learning from China's best practices and challenges in developing and implementing digital technologies and platforms, establishing more collaborations and partnerships with researchers and institutions in respective developing countries, and addressing the crucial topics that are essential for ensuring the equitable distribution and mitigation of the benefits and risks of the digital economy, among others. The paper contributes to the literature by providing an overview of the current state of knowledge about the digital economy in developing countries, specifically focusing on papers published in the English language. It also offers valuable insights for policymakers, practitioners, and researchers who are interested in advancing the digital economy and fostering digital transformation in developing nations. The paper is significant, relevant, and timely as it addresses a crucial topic that has profound implications for economic development, social inclusion, and environmental sustainability in this era of rapid and sustained digitalisation.



## References

- Alekseev, A. N., Lobova, S. V., & Bogoviz, A. V. (2021). Digitalization and quality of labor: contradictions in developing countries and the prospects of harmonization [Article]. *International Journal for Quality Research*, 15(3), 733-752. <https://doi.org/10.24874/IJQR15.03-04>
- Bihari, A., Tripathi, S., & Deepak, A. (2023). A review on h-index and its alternative indices. *Journal of Information Science*, 49(3), 624-665.
- Boateng, H., Kosiba, J. P. B., & Okoe, A. F. (2019). Determinants of consumers' participation in the sharing economy: A social exchange perspective within an emerging economy context [Article]. *International Journal of Contemporary Hospitality Management*, 31(2), 718-733. <https://doi.org/10.1108/IJCHM-11-2017-0731>
- Bogoviz, A. V. (2021). Limits of the pleasure economy in Industry 4.0: Specifics of developing countries [Article]. *International Journal of Trade and Global Markets*, 14(4-5), 341-348. <https://doi.org/10.1504/IJTGM.2021.116726>
- Bogoviz, A. V., Lobova, S. V., & Alekseev, A. N. (2021). Labour efficiency in the pleasure economy: Perspectives of increase in developing countries in Industry 4.0 [Article]. *International Journal of Trade and Global Markets*, 14(4-5), 370-377. <https://doi.org/10.1504/IJTGM.2021.116730>
- Chiappetta Jabbour, C. J., De Camargo Fiorini, P., Wong, C. W. Y., Jugend, D., Lopes De Sousa Jabbour, A. B., Roman Pais Seles, B. M., Paula Pinheiro, M. A., & Ribeiro da Silva, H. M. (2020). First-mover firms in the transition towards the sharing economy in metallic natural resource-intensive industries: Implications for the circular economy and emerging industry 4.0 technologies [Article]. *Resources Policy*, 66, Article 101596. <https://doi.org/10.1016/j.resourpol.2020.101596>
- Deng, Z., Jean, R. J. B., & Sinkovics, R. R. (2018). Rapid expansion of international new ventures across institutional distance [Article]. *Journal of International Business Studies*, 49(8), 1010-1032. <https://doi.org/10.1057/s41267-017-0108-6>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285-296.
- Dreyer, B., Lüdeke-Freund, F., Hamann, R., & Faccar, K. (2017). Upsides and downsides of the sharing economy: Collaborative consumption business models' stakeholder value impacts and their relationship to context [Article]. *Technological Forecasting and Social Change*, 125, 87-104. <https://doi.org/10.1016/j.techfore.2017.03.036>
- Efendioglu, A. M., & Yip, V. F. (2004). Chinese culture and e-commerce: An exploratory study [Article]. *Interacting with Computers*, 16(1), 45-62. <https://doi.org/10.1016/j.intcom.2003.11.004>
- Kumar, V., Lahiri, A., & Dogan, O. B. (2018). A strategic framework for a profitable business model in the sharing economy [Article]. *Industrial Marketing Management*, 69, 147-160. <https://doi.org/10.1016/j.indmarman.2017.08.021>
- Lehdonvirta, V., Kässi, O., Hjorth, I., Barnard, H., & Graham, M. (2019). The Global Platform Economy: A New Offshoring Institution Enabling Emerging-Economy Microproviders [Article]. *Journal of Management*, 45(2), 567-599. <https://doi.org/10.1177/0149206318786781>
- Ma, D., & Zhu, Q. (2022). Innovation in emerging economies: Research on the digital economy driving high-quality green development [Article]. *Journal of Business Research*, 145, 801-813. <https://doi.org/10.1016/j.jbusres.2022.03.041>
- Popkova, E. G., Bogoviz, A. V., Ekimova, K. V., & Sergi, B. S. (2023). Will Russia become a blueprint for emerging nations' high-tech reforms? evidence from a 26-countries

- dataset [Article]. *International Journal of Innovation Studies*, 7(4), 294-306. <https://doi.org/10.1016/j.ijis.2023.05.001>
- Tirachini, A. (2020). Ride-hailing, travel behaviour and sustainable mobility: an international review [Article]. *Transportation*, 47(4), 2011-2047. <https://doi.org/10.1007/s11116-019-10070-2>
- van Ooijen, C., Ubaldi, B., & Welby, B. (2019). A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance.
- Wu, J., Zeng, M., & Xie, K. L. (2017). Chinese travelers' behavioral intentions toward room-sharing platforms: The influence of motivations, perceived trust, and past experience [Article]. *International Journal of Contemporary Hospitality Management*, 29(10), 2688-2707. <https://doi.org/10.1108/IJCHM-08-2016-0481>

## Appendix

### 1. Bibliometric Analysis Search String (Scopus Database)

( TITLE-ABS-KEY ( "digital economy" ) OR TITLE-ABS-KEY ( "network\* economy" ) OR TITLE-ABS-KEY ( "sharing economy" ) OR TITLE-ABS-KEY ( "platform economy" ) OR TITLE-ABS-KEY ( "gig economy" ) AND TITLE-ABS-KEY ( "developing countr\*" ) OR TITLE-ABS-KEY ( "developing nation\*" ) OR TITLE-ABS-KEY ( "emerging econom\*" ) OR TITLE-ABS-KEY ( "emerging market\*" ) ) AND PUBYEAR > 2002 AND PUBYEAR < 2024 AND ( LIMIT-TO ( SRCTYPE , "j" ) OR LIMIT-TO ( SRCTYPE , "p" ) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) ) )

### 2. A Short Response to Peer Reviewers' Comments on the First Draft

**Response to Peer Reviewer 1:** Thank you very much for taking the time to read, review, and leave comments on the initial draft. Your contribution has provided an opportunity to improve the paper. Specifically, we have now emphasised that *the focus of the study was to consider only scientific papers published in the English language*. This strategy became necessary because the three co-authors only read and write in English, a medium that came in handy when searching for the articles, retrieving them, reading them, analysing them, and presenting the outcome. In other words, downloading papers in other languages from English would have been nonprofitable because we would not have had the language skills to work with them. This clarification is now reflected in the abstract, introduction, justification, research questions, data collection, inclusion and exclusion criteria, discussions, limitation of the study, future research direction, conclusions, and other sections of the paper. In recognition of this limitation, we had already encouraged future researchers to attempt replicating this study in other languages, like Portuguese, Spanish, French, etc.; doing so will be a useful opportunity for comparing or validating our findings – beyond papers published in the English language.

**Response to Peer Reviewer 2:** We appreciate the effort and time you put into reviewing and recommending some improvements on the first draft, by way of your comments. We agree with you that using only Scopus as a source of retrieving data might limit the research breadth; this much we have alluded to in the limitations of the study section in the initial draft. Be that as it may, choosing only one database was strategic, considering the use of VOSviewer for our analysis. In our experience, mixing papers from various databases before analysing them in VOSviewer typically skews the results because the format in which the respective paper metadata are organised varies from database to database. In other words, *VOSviewer works best with downloaded files with the same types of metadata all through*. Moreover, we had already encouraged future researchers to attempt replicating this study by retrieving data from other databases, like Google Scholar, Web of Science, etc. Secondly, we have clarified the issue about using quantitative methods. In actual fact, we only aggregated metadata from the 358 papers analysed and then presented the outcome through visualisation charts, tables, and graphs. In effect, *we did not use any sophisticated statistical method or any other advanced quantitative methods, all we did was to either sum, multiply, divide, or subtract using standard Microsoft Excel prompts*, hence the added phrase: **metadata aggregation**. This clarification can be found in the Data Processing and Analysis section of this updated draft.

# The Role of Augmented Reality in Enhancing Sustainable Decision-Making in the Fashion Industry

Omon Stellamaris Fagbamigbe, Mohammad Awais Shakir, Zeeshan Bhatti  
*University of Portsmouth*

## Abstract

*This paper focuses on the role of Augmented Reality (AR) in enhancing sustainability, specifically in minimising product return in the fashion industry. The fashion industry is the second highest producer of greenhouse gas emissions, leading to climate change. The increased environmental awareness has led to a shift in consumer behaviour. Consumers are now more actively seeking sustainable products and brands. However, consumers' positive environmental sustainability values do not always align with their purchasing decisions due to price or product fit uncertainty, and more. AR allows consumers to "experience" products through virtual-try-on, resulting in higher confidence in purchasing decisions and minimised returns, ultimately enhancing sustainability, in line with global Sustainable Development Goals. Building on insights from extant literature, this research will develop a framework grounded on the Stimulus-Organism-Behaviour-Consequence model using quantitative data, collected through online questionnaires among Millennials and GenZ. This research contributes to information system literature on consumer behaviour and sustainability.*

**Keywords:** Augmented Reality, Consumer Behaviour, Sustainability, Product Return, Retail, Fashion Industry.

## 1. Introduction

Consumers' purchase decisions have been significantly influenced by sustainability (Gazzola et al., 2020; Kapse et al., 2023), driven by advancements in technology (Hoyer et al., 2020; Shankar et al., 2021) such as augmented reality (AR). Consumers are now more aware of the environmental impact of their purchasing decisions (Kapse et al., 2023, p. 236). This increased awareness of sustainability issues is significantly influencing consumer preferences. A global survey on consumer behaviour found that 73% of respondents would unquestionably change their purchasing habits to reduce the environmental effect of their consumption (Nielsen, 2019). As a result, consumers may choose to purchase a product from a brand that promotes sustainability or avoid brands with a poor reputation in terms of environmental sustainability (Joerß et al., 2021). Consequently, retailers are beginning to implement business models that provide consumers with a more sustainable shopping experience (Zhu et al., 2018), in response to the increasing pressure and global attention on environmental sustainability (Hasbullah et al., 2022). In the context of the fashion industry, "sustainability refers to not being wasteful of resources at various stages of fashion design, production, and consumption" (Mesjar et al., 2023, p. 1). This implies, that sustainability is the efficient use of resources in a way that minimizes harm to the environment across the supply chain.

Presently, the fashion industry is deemed unsustainable, being the second highest producer of greenhouse gas (GHG) emissions after the oil and gas industry, accounting for approximately 10% of global GHG emissions (The World Bank, 2019; Zurich, 2023). A significant part of the GHG emission within the retail sector can be attributed to the high volume of online returns, in 2023, 20.8% of total retail purchases worldwide are anticipated to take place online (Baluch, 2023). To put this in context, the total volume of parcels shipped in 2013 was 36 billion (Michels *et al.*, 2022) and is anticipated to reach approximately 225 billion by 2028, reflecting a Compound Annual Growth Rate (CAGR) of 6% from 2023 to 2028 (Pitney Bowes Inc, 2022). This growth in online sales has raised concerns about the environmental impact of online retail (Van Loon *et al.*, 2015; Michels *et al.*, 2022). It is worth noting that logistics and distribution have been identified as the main contributors to environmental degradation in the online retail supply chain (Buldeo Rai, 2021; Michels *et al.*, 2022), particularly in the fashion industry. Every year, £5.8 billion of returned products ends up in landfills, subsequently undergoing degradation into microfibers that contribute to the pollution of oceans over time (Driscoll, 2023) creating problems for human health and the environment (The World Bank, 2019; BoF & McKinsey, 2022). For example, UK online retailer ASOS stated that customer returns accounted for 12% of its carbon emissions (Sword, 2020).

The detrimental impact of product return on the environment is intrinsically linked to consumer behaviour, particularly because the consumer initiates the process of product return. Although consumers may express positive environmental attitudes and values (Mostafa, 2006), they may not necessarily make sustainable purchasing decisions (De Pelsmacker, Driesen and Rayp, 2005; Michels *et al.*, 2022). This situation is known as the ‘attitude-behaviour gap’ (Sadiq, Adil and Paul, 2022; Colombo *et al.*, 2023). Consumers may have positive environmental attitude but struggle to make purchasing decisions that align with their values. These inconsistencies can be attributed to various factors such as lack of awareness of the impact of their purchasing decisions on the environment (De Pelsmacker, Driesen and Rayp, 2005; Nisar *et al.*, 2021), convenience, quality, product price (Bushara *et al.*, 2023), and product fit uncertainty (Mulpuru, 2017; Park and Yoo, 2020). Product fit uncertainty has been recognized as one of the top reasons for the high volume of online return rates (Mulpuru, 2017).

According to Park and Yoo (2020), consumers have difficulty ascertaining the true fit and quality of products before purchase. Similarly, consumers may choose to return a product not only due to the disappointing fit or poor quality of the product but also due to the low quality of their decision-making (Zhu and Zhang, 2010; Wang, Yu and Chen, 2023). These

uncertainties remain a major challenge in the retail industry (Wang *et al.*, 2021). AR has the potential to bridge these gaps by providing consumers with vivid and accurate information about the product (Sihi, 2018), improving decision comfort (Heller *et al.*, 2019); ultimately, enhancing sustainable decision making.

Research shows that AR has the potential to influence consumer purchasing decisions (AL Hilal, 2023) by providing immersive and interactive experiences. AR blends the virtual and real worlds (Huang & Liao, 2015), by using computer-generated images to enhance the user's perception of the real world (Hilken *et al.*, 2022; Yim *et al.*, 2017). These characteristics of AR enabling realistic product visualization (Amorim *et al.*, 2022; Hilken *et al.*, 2022; Hoyer *et al.*, 2020; Lee *et al.*, 2022; Vieira *et al.*, 2022), have been found to profoundly impact consumer's perception of products affecting their purchasing decision (Dwivedi *et al.*, 2021). Thus, AR enhances the mental imagery of products which improves decision comfort (Heller *et al.*, 2019), whilst offering consumers access to fundamental product information, facilitating their purchase decisions (Tanrikulu, 2021). Similarly, AR has been found to increase consumer's confidence when purchasing products via online platforms (Rauschnabel *et al.*, 2022), which could lead to high consumer satisfaction and ultimately minimise product return. Retailers such as Sephora's makeup have deployed AR app that enables consumers to virtually try on products such as lipstick before purchasing, thereby increasing confidence in purchasing decisions (Rauschnabel *et al.*, 2022; Scholz & Duffy, 2018), and minimizing return rate and waste, ultimately promoting environmental sustainability. The intersection of AR, consumer decision making and sustainability, is an emerging area of research.

Previous research has focused on the adoption of AR in sustainable retailing (Caboni & Bruni, 2022), sales of sustainable products (Jäger & Weber, 2020), strategies for creating customer value perception (Hilken *et al.*, 2022), improving customer responses (Rauschnabel *et al.*, 2019), enhancing brand relationships (Scholz & Duffy, 2018), and supply-chain (Wren, 2022). Similarly, extant literature shows that most studies on AR are concentrated on consumer attitude and perception rather than how it reduces product return or changes consumer behaviour in a measurable way (Mesjar *et al.*, 2023). It is therefore important to develop a deeper understanding of how the characteristics of AR can enhance consumer decision-making and ultimately promote environmental sustainability through minimized product return. However, there is limited research on how AR could influence sustainable consumer decision-making from the perspective of product return. Therefore, the focus of this research is on the role of

AR in enhancing consumer decision-making and promoting environmental sustainability in the fashion industry. Therefore, to address these gaps we propose a research question:

RQ: How does consumer decision-making using AR impact sustainability in the fashion industry?

To answer this research question, we propose a theoretical model based on the Stimulus-Organism-Behaviour-Consequence model (SOBC) (Davis and Luthans, 1980); where AR technology is recognised as the stimulus and product return as the consequences of consumer behaviour. This framework provides insight into how features of AR through product visualization (Virtual try-on) and detailed informativeness about the product can enable the consumer to make more confident decisions about the product. As such, virtual try-on (VTO) could lead to a higher tendency for the consumer to retain the product, minimising product return and ultimately bolstering environmental sustainability. AR provides engaging and interactive shopping experience, which may appeal more to Millennials and Gen Z consumers who are known for their love of technology. Research also shows that Millennials and Gen Z have a higher propensity to return products bought online, with 32 % of Gen Z and 36% of Millennials considered high-intensity returners when compared to 40% of those over 41 (The British Fashion Council, 2023).

## **2. Theoretical Framework**

Building on the insight from information systems, marketing and sustainability literature, this research proposes a framework to help consumers and businesses understand how AR can help promote sustainability. The proposed conceptual framework is grounded on the Stimulus-Organism-Behaviour-Consequence (SOBC), which is an extension of the Stimulus-Organism-Response (SOR) (Chopdar et al., 2022). The SOBC model helps to provide insight into the antecedents and consequences of behaviours in environmental psychology (Davis and Luthans, 1980).

The SOBC model stipulates that the stimuli (S) represented as the environment influences the internal state of the consumer denoted as the organism (O), which then triggers a behavioural response (B), leading to consequences (C) which is the behavioural outcomes (Chakraborty *et al.*, 2022; Xuan *et al.*, 2023). This sequential flow of the SOBC model helps to explore the intention behaviour gap and, consequently, the consumer decision-making (Talwar *et al.*, 2021), where (S) interacts with (O) to trigger (B) resulting in C (Chakraborty *et al.*, 2022; Duong, 2023), ultimately providing insight to sustainable retailing.

Although previous research has used the SOBC framework to gain an understanding of consumer behaviour. For example, SOBC was used to identify the drivers of green apparel purchase behaviour that may potentially mitigate the intention-behaviour gap (Dhir et al., 2021), the paradigm to examine buying behaviour towards organic food (Talwar et al., 2021) and to understand consumers' engagement with new media (Yuan et al., 2017). However, to the best of our knowledge and at this point of this research, there is currently no research that has incorporated the intersection of AR, consumer behavioural intention and sustainability in the SOBC model, particularly from the perspective of product return. The SOBC framework is considered suitable for this research because of its adaptability and applications in promoting environmental sustainability. It also accounts for a variety of factors that influences consumer behaviour (Dhir et al., 2021).

### 3. Conceptual Framework and Hypothesis

The SOBC framework highlight the need to consider both environmental stimuli and individual factors when attempting to understand and predict consumer behaviour (Dhir et al., 2021). As a result, this study conceptualizes the following features of AR as stimuli (S) – perceived interactivity (Kim, Park and Kader, 2022; Kumar and Srivastava, 2022) and augmentation (Lee, Xu and Porterfield, 2022). We propose that these features of AR technology acts as a stimulus which interacts with the consumer based on their environmental attitude and environmental knowledge to trigger a buying behaviour leading to a consequence (for the purpose of this research we consider product return as the consequence), shown in figure 1.

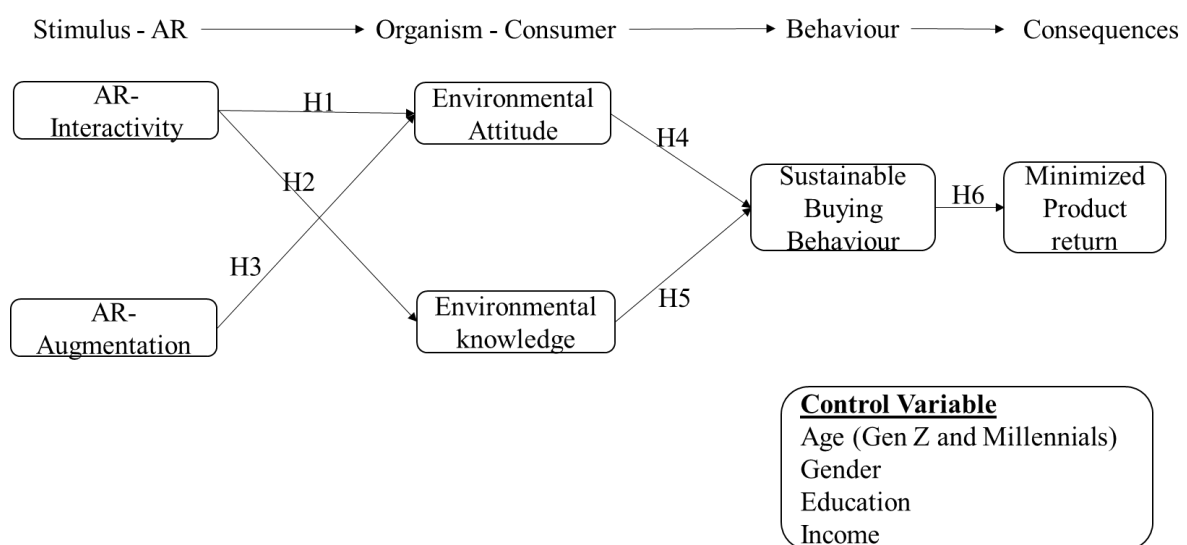


Figure 1: Conceptual Framework



AR technology enhances contextual information and usability of the product (Caboni and Bruni, 2022; Wu, 2023), providing detail information on the eco-friendliness of products (Mesjar et al., 2023), ways to manage and the final disposal of these products. We propose that consumers who interact with AR technology when making purchase, are more likely to develop positive environmental attitudes and greater environmental knowledge, which may influence their purchasing behaviour. This immersive and informative nature of AR experiences provides an opportunity for consumers to connect with environmentally friendly products, leading to more conscious and sustainable purchasing decisions. In consonance, AR informativeness has been found to increase consumers' cognitive consistency which aligns with their attitudes and purchase behaviour (Hilken et al., 2022; Vieira, Rafael and Agnihotri, 2022). Subsequently, AR enhances the consumer's actual buying behaviour of environmentally friendly products (Mesjar et al., 2023) and the post purchase behaviour specifically product return behaviour. However, according to Kumar et.al (2017), "intention has been assumed to be a strong predictor of behaviour but in some cases, it may not act consistently".

In the same vein, we argue that AR VTO provides the consumer with a more realistic assessment of products which will lead to a more convenient and confident sustainable purchase. AR is expected to enhance consumer awareness of environmental problems and promote eco-friendly purchase decision through immersive experiences. We hypothesize that exposure to AR augmentation and interactive features positively influence consumers' environmental attitudes and knowledge, buying behaviour and ultimately leading to minimised product return.

H1: AR interactivity has a positive effect on a consumer's environmental attitude.

H2: AR interactivity has positive effect on perceived environmental knowledge.

H3: AR augmentation has positive effect on environmental attitude.

H4: Environmental attitude has positive association with Buying Behaviour.

H5: Perceived environmental knowledge has positive association with Buying Behaviour.

H6: Sustainable buying behaviour has a positive effect on product return.

By investigating the intricate interplay between these variables, this research seeks to unveil the nuanced dynamics on how AR enhances sustainable consumer decision-making from the perspective of product return in the fashion industry, by adopting quantitative methodology.

#### **4. Methodology**

The study design is analytical and is currently at the data collection stage. Quantitative data is being collected using self-administered questionnaires with structured questions from the Millennials and GenZ population who uses AR to make their purchase via electronic platforms. Data will be analysed using partial least square structural equation modelling (PLS-SEM) (Hair Jr. et al., 2017). This analysis aims to provide insights into the proposed relationships between variables, and a deeper understanding of consumer behaviour and their corresponding decision-making.

#### **5. Study Population**

The study population consists of millennials and GenZ. Millennials and GenZ are typically aged 11 to 42 years and are considered the most fashion-conscious consumer group (Hasbullah et al., 2022; Samala & Katkam, 2020). They are at ease with and enthusiastic about emerging technologies and have an information-age mindset (Samala & Katkam, 2020), thus making them a key market for fashion retailers.

#### **6. Study Area**

England, United Kingdom is purposively selected for convenience and limited funding.

#### **7. Significance and Contributions**

AR provides the consumer with an avenue to experience the product before seeing it in person (Sihi, 2018), through virtual-try-on thereby leading to more confident decision-making, which will invariably have a knock-on effect on minimizing product return rate and its impact on the environment. Environmental sustainability is enshrined in the United Nations 2015 Sustainable Development Goals (SDG) (United Nations, 2022). Sustainability is no longer an option but a top priority and a strategic plan owing to growing pressure and a global focus on socially responsible behaviour (Hasbullah et al., 2022). With the growing awareness of sustainability and the drive for net zero by 2050, consumers are becoming more aware of the impact of their purchasing decisions on the environment (Vlastelica et al., 2023).

Although several scholars have identified the importance of the study of consumer behaviour and their decision-making process (Dimanche & Havitz, 1995; Furaiji & Łatuszyńska, 2012; Jose, 2017; Shahab et al., 2021), it is imperative to understand how AR can enhance consumer intention to make sustainable purchases in the fashion industry.

The results of this study will be an invaluable resource for businesses and managers across the retail industry to formulate sustainable business models based on insights into how AR can be used to enhance consumer decision-making and promote environmental sustainability. This

will inevitably result in improved customer engagement, satisfaction, loyalty, and retention which will invariably lead to an increase in profitability both in the short-term and long term. This research aims to contribute to the information system literature on consumer behaviour in the retail sector as well as the growing need for sustainability, through insight into how AR technology supports sustainable decision-making. Specifically, the study outcomes will contribute significantly to the attainment of SDGs 9 (industry, innovation, and infrastructure), 11 (sustainable cities and communities), 12 (responsible consumption and production) and 13 (Climate action).

## References

- Baluch, A. (2023) eCommerce Statistics of 2023 – Forbes Advisor, Forbes. Available at: <https://www.forbes.com/advisor/business/e-commerce-statistics/> (Accessed: August 24, 2023).
- BoF & McKinsey (2022) The state of Fashion 2022, McKinsey & Company.
- Buldeo Rai, H. (2021) “The net environmental impact of online shopping, beyond the substitution bias,” *Journal of Transport Geography*, 93, p. 103058. Available at: <https://doi.org/10.1016/J.JTRANGE.2021.103058>.
- Bushara, M.A. et al. (2023) “Power of Social Media Marketing: How Perceived Value Mediates the Impact on Restaurant Followers’ Purchase Intention, Willingness to Pay a Premium Price, and E-WoM?,” *Sustainability*, 15(6), p. 5331. Available at: <https://doi.org/10.3390/su15065331>.
- Caboni, F. and Bruni, R. (2022) “Exploring Augmented Reality Applications for Sustainable Retailing,” *International Series in Advanced Management Studies*, pp. 131–146. Available at: [https://doi.org/10.1007/978-3-031-12027-5\\_8/TABLES/3](https://doi.org/10.1007/978-3-031-12027-5_8/TABLES/3).
- Chakraborty, D. et al. (2022) “Exploring consumer purchase intentions and behavior of buying ayurveda products using SOBC framework,” *Journal of Retailing and Consumer Services*, 65, p. 102889. Available at: <https://doi.org/10.1016/J.JRETCONSER.2021.102889>.
- Colombo, S.L. et al. (2023) “Understanding the Environmental Attitude-Behaviour Gap : The Moderating Role of Dispositional Mindfulness,” pp. 1–12.
- Davis, T.R. V and Luthans, F. (1980) “A Social Learning Approach to Organizational Behavior Part of the Business Administration, Management, and Operations Commons, Management Sciences and Quantitative Methods Commons, and the Strategic Management Policy Commons A Social Learning Approach to ,” *Academy of Management Review*, 5(2), pp. 281–290. Available at: <https://digitalcommons.unl.edu/managementfacpubhttps://digitalcommons.unl.edu/managementfacpub/173>.
- Dhir, A. et al. (2021) “Green apparel buying behaviour: A Stimulus–Organism–Behaviour–Consequence (SOBC) perspective on sustainability-oriented consumption in Japan,” *Business Strategy and the Environment*, 30(8), pp. 3589–3605. Available at: <https://doi.org/10.1002/BSE.2821>.

- Driscoll, M. (2023) "Returns Uncovered," (April).
- Duong, C.D. (2023) "Using a unified model of TPB, NAM and SOBC to understand students' energy-saving behaviors: moderation role of group-level factors and media publicity," *International Journal of Energy Sector Management* [Preprint]. Available at: <https://doi.org/10.1108/ijesm-09-2022-0017>.
- Hilken, T. et al. (2022) "How to strategically choose or combine augmented and virtual reality for improved online experiential retailing," *Psychology & Marketing*, 39(3), pp. 495–507. Available at: <https://doi.org/10.1002/MAR.21600>.
- Hoyer, W.D. et al. (2020) "Transforming the Customer Experience Through New Technologies," *Journal of Interactive Marketing*, 51, pp. 57–71. Available at: <https://doi.org/10.1016/j.intmar.2020.04.001>.
- Kim, S., Park, H. and Kader, M.S. (2022) "How augmented reality can improve e-commerce website quality through interactivity and vividness: the moderating role of need for touch," *Journal of Fashion Marketing and Management*, pp. 1–24. Available at: <https://doi.org/10.1108/JFMM-01-2022-0001>.
- Kumar, H. and Srivastava, R. (2022) "Exploring the role of augmented reality in online impulse behaviour," *International Journal of Retail and Distribution Management* [Preprint]. Available at: <https://doi.org/10.1108/IJRDM-11-2021-0535>.
- Lee, H., Xu, Y. and Porterfield, A. (2022) "Antecedents and moderators of consumer adoption toward AR-enhanced virtual try-on technology: A stimulus-organism-response approach," *International Journal of Consumer Studies*, 46(4), pp. 1319–1338. Available at: <https://doi.org/10.1111/IJCS.12760>.
- Van Loon, P. et al. (2015) "A comparative analysis of carbon emissions from online retailing of fast moving consumer goods," *Journal of Cleaner Production*, 106, pp. 478–486. Available at: <https://doi.org/10.1016/J.JCLEPRO.2014.06.060>.
- Mesjar, L. et al. (2023) "The Intersection of Fashion, Immersive Technology, and Sustainability: A Literature Review," *Sustainability (Switzerland)*, 15(4), p. 3761. Available at: <https://doi.org/10.3390/SU15043761>.
- Michels, L. et al. (2022) "Empowering Consumers to Make Environmentally Sustainable Online Shopping Decisions: A Digital Nudging Approach," *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2022-January, pp. 4707–4716. Available at: <https://doi.org/10.24251/HICSS.2022.574>.
- Mostafa, M.M. (2006) "Antecedents of Egyptian consumers' green purchase intentions: A hierarchical multivariate regression model," *Journal of International Consumer Marketing*, 19(2), pp. 97–126. Available at: [https://doi.org/10.1300/J046v19n02\\_06](https://doi.org/10.1300/J046v19n02_06).
- Mulpuru, S. (2017) Think Tank: The Opportunity in Online Returns for Apparel Retailers – WWD, WWD. Available at: <https://wwd.com/feature/sucharita-mulpuru-think-tank-returns-10939339/> (Accessed: August 20, 2023).

Nisar, Q.A. et al. (2021) “Cosmopolitan orientation, consumer consciousness, and green purchase intentions: mediating role of green trust,” *European J. of International Management*, 1(1), p. 1. Available at: <https://doi.org/10.1504/ejim.2021.10034093>.

Park, M. and Yoo, J. (2020) “Effects of perceived interactivity of augmented reality on consumer responses: A mental imagery perspective,” *Journal of Retailing and Consumer Services*, 52, p. 101912. Available at: <https://doi.org/10.1016/j.jretconser.2019.101912>.

De Pelsmacker, P., Driesen, L. and Rayp, G. (2005) “Do consumers Care about ethics? Willingness to pay for fair-trade coffee,” *Journal of Consumer Affairs*, pp. 363–385. Available at: <https://doi.org/10.1111/j.1745-6606.2005.00019.x>.

Pitney Bowes Inc (2022) Pitney Bowes Parcel Shipping Index Confirms Extended Effects of COVID-19 Lockdowns on Chinese Shipping Through 2022 | Business Wire, Business Wire. Available at: <https://www.businesswire.com/news/home/20230809471409/en/Pitney-Bowes-Parcel-Shipping-Index-Confirms-Extended-Effects-of-COVID-19-Lockdowns-on-Chinese-Shipping-Through-2022> (Accessed: August 24, 2023).

Rauschnabel, P.A. et al. (2022) “What is augmented reality marketing? Its definition, complexity, and future,” *Journal of Business Research*, 142, pp. 1140–1150. Available at: <https://doi.org/10.1016/J.JBUSRES.2021.12.084>.

Sadiq, M., Adil, M. and Paul, J. (2022) “Organic food consumption and contextual factors: An attitude–behavior–context perspective,” *Business Strategy and the Environment*, (October), pp. 1–15. Available at: <https://doi.org/10.1002/bse.3306>.

Shankar, V. et al. (2021) “How Technology is Changing Retail,” *Journal of Retailing*, 97(1), pp. 13–27. Available at: <https://doi.org/10.1016/j.jretai.2020.10.006>.

Sword, A. (2020) ASOS achieves 30% reduction in carbon emissions per order, ASOS. Available at: <https://deliveryx.net/asos-achieves-30-reduction-carbon-emissions-per-order/> (Accessed: March 27, 2023).

Talwar, S. et al. (2021) “What drives willingness to purchase and stated buying behavior toward organic food? A Stimulus–Organism–Behavior–Consequence (SOBC) perspective,” *Journal of Cleaner Production*, 293, p. 125882. Available at: <https://doi.org/10.1016/J.JCLEPRO.2021.125882>.

The British Fashion Council (2023) “Solving fashion’s product returns How to keep value in a closed-loop system,” Institute of Positive Fashion [Preprint].

The World Bank (2019) How Much Do Our Wardrobes Cost to the Environment?, World Bank. Available at: <https://www.worldbank.org/en/news/feature/2019/09/23/costo-modamedio-ambiente> (Accessed: June 15, 2023).

Vieira, V.A., Rafael, D.N. and Agnihotri, R. (2022) “Augmented reality generalizations: A meta-analytical review on consumer-related outcomes and the mediating role of hedonic and utilitarian values,” *Journal of Business Research*, 151, pp. 170–184. Available at: <https://doi.org/10.1016/J.JBUSRES.2022.06.030>.

Wang, C. et al. (2021) “Factors influencing consumers’ purchase decision-making in O2O business model: Evidence from consumers’ overall evaluation,” *Journal of Retailing and Consumer Services*, 61. Available at: <https://doi.org/10.1016/j.jretconser.2021.102565>.

Wang, Y., Yu, B. and Chen, J. (2023) “Factors affecting customer intention to return in online shopping: the roles of expectation disconfirmation and post-purchase dissonance,” *Electronic Commerce Research*, pp. 1–35. Available at: <https://doi.org/10.1007/S10660-023-09769-3/TABLES/5>.

Wu, C.-W. (2023) “The Impact of Virtual Reality and Augmented Reality Service Technologies on Consumer Purchase Intention for Fashion Brands.” Available at: <https://doi.org/10.20944/preprints202307.1408.v1>.

Xuan, V.N. et al. (2023) “Using a Unified Model of TPB, NAM, and SOBC to Investigate the Energy-Saving Behaviour of Urban Residents in Vietnam: Moderation Role of Cultural Values,” *Sustainability* 2023, Vol. 15, Page 2225, 15(3), p. 2225. Available at: <https://doi.org/10.3390/SU15032225>.

Zhu, F. and Zhang, X. (Michael) (2010) “Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics,” <https://doi.org/10.1509/jm.74.2.133>, 74(2), pp. 133–148. Available at: <https://doi.org/10.1509/JM.74.2.133>.

Zurich (2023) What is fast fashion and how is it damaging the planet? | Zurich Insurance, Zurich. Available at: <https://www.zurich.com/en/media/magazine/2021/fast-fashion-5-practical-ways-to-cut-the-carbon-from-your-closet> (Accessed: August 27, 2023).

# An Exploration of Contemporary Factors that Influence Adoption of Instant Messaging by Academics in Higher Education

*Research In Progress*

## **Abstract**

*Instant messaging (IM) has been found to support both synchronous and asynchronous learning in Higher Education Institutions (HEIs), but the efficacy of this approach is heavily reliant on staff adoption. Related work has focussed on the effect of IM on learners; frequency of usage, perception of usefulness, and its relationship with academic performance. However, this research will explore perceptions of academic staff towards the use of IM, and identify and classify the barriers that hinder adoption of IM utilising a composite theoretical model and survey intended to capture the impact of IM on stress and academic workload. This Research in Progress paper is informed by the authors' experience of teaching apprentice students within HE, and sets out the background, rationale, and theoretical frameworks for this study. This work intends to contribute insight for Higher Education institutions aiming for optimal adoption of IM for inter-staff and student-staff communication.*

**Keywords:** Instant Messaging, Academic Workload, Technostress, Digitalised Workplace, GenAI, LLMs, Staff Well-Being

## **1.0 Introduction**

Instant messaging (IM) tools emerged initially in the 1990s, with platforms like ICQ, AIM (AOL Instant Messenger), and Yahoo! Messenger allowing users to send text messages and files to one another over the internet. The proliferation of smartphones in the late 2000s brought another significant shift to instant messaging. In 2015, WhatsApp became the most popular IM tool, having registered more than 900 million users (Sun, 2015), and this continued to grow to over 2 billion users in early 2020

(WhatsApp blog, 2020). Later businesses and governmental organisations started to use platforms like Slack, Microsoft Teams, Zoom and Google Chat for internal communication and collaboration. IM tools bring people from different geographical spaces closer together through quasi-synchronous communication (So, 2016) and have proved their worth in facilitating team collaboration, problem solving, coordination and efficient decision making (Hurbean et al, 2022).

While IM tools have become more widespread and accessible within Higher Education (HE) settings, along with the emergence of Generative AI (GenAI) tools such as ChatGPT which are further normalising IM in the form of chatbots, it is unclear how IM is being adopted by academics on a larger scale. While important work has been conducted since the early 2000s to identify the attitudes and engagement of learners with IM technology, there is a need to further understand the perceptions of academics who are important facilitators of effective IM practice within an educational context.

This research aims to explore the adoption of IM tools by teachers in HE institutions and the impact of IM tools on the stress levels and well-being of academic staff. The authors are planning to perform initial explorative research to identify influencing factors in the context of HE. Following this, a wider survey will be used to gain a deeper understanding of the strength of each influencing factor as well as adoption levels of IM for inter-staff and student-staff communication.

To summarise, the authors specify the following research questions:

*RQ1: What are the contemporary barriers, facilitators and stressors that impact the adoption of IM by academic staff within Higher Education contexts?*

*RQ2: What are practical solutions to achieve optimal adoption of IM in Higher Education contexts?*

## **2.0 Related work**

Studies between the early 2000s and 2015, when the use of IM was more select and smartphones were in their infancy, found that learners would feel comfortable or 'very comfortable' with text and instant messaging (Lauricella and Kay, 2013). This sense of



comfort may stem from familiarity with the tools, as De Bakker and colleagues (2007) report that approximately 75% of participants in their study utilise IM tools at least once a day for five out seven days a week. Further studies found that learners used IM to coordinate work and receive answers to queries outside of timetabled sessions (Hrastinski and Aghaee, 2012), and has also provided learners an ability to communicate with academic staff, in contexts where they may have felt inhibited in large class sizes (Lents and Cifuentes, 2010).

The COVID-19 worldwide pandemic prompted a greater uptake in video conferencing software such as Zoom and Microsoft Teams, which also facilitate IM between staff and learners. A study conducted in 2020, sampling over 1700 undergraduates, found a statistically significant interaction between formal and study Academic Instant Messaging Groups (AIMGs) and academic performance, in addition to lower academic stress levels (le Roux and Perry, 2022).

A plethora of research exists that investigates the impact of IM on student-teacher collaboration and learning enhancement (De Bakker et al., 2007; Lauricella and Kay, 2013; Tang and Hew, 2019) but while examples of good practice are evident, it is unclear whether IM is being used intentionally and sustainably on a larger scale.

Like any new technology, the adoption of IM at work can introduce both new opportunities and new threats to employees' work performance and well-being (Hurbean et al, 2022). IM can be a significant cause of distraction that moves attention away from main working goals (Spira and Feintuch, 2005). This flexibility and versatility of IM tools can contribute to employees' technostress.

As of 2024, individuals can interact with GenAI chatbots based on large language models (LLMs) such as ChatGPT, Google Bard, Claude AI, and Microsoft Copilot. For some, this may be normalising the use of IM, even easing academic workload through efficiency gains. For others, it may be a significant stressor and even pose a perceived existential threat (Concannon et al., 2023).

### **3.0 Theoretical Background**

Complexity of technology usage behaviour can be explored from various perspectives and there are numerous theories and frameworks that focus on various aspects of this behaviour. Some models take a usability perspective, others also take technical characteristics into consideration, yet others articulate the social aspect of technology usage and adoption. Two of the most widely used frameworks in the context of technology adoption and usability research are the Information Systems Success Model and Technology Acceptance Model.

The Information Systems Success Model is a widely recognised framework that explains the success of information systems in organisations (DeLone and McLean, 2003). It is frequently used in information systems research and evaluation since it addresses both usability and technical aspects of information systems. The model postulates that components such as system quality, information quality and service quality influence user satisfaction which in turn influences intention to use and actual use of tools and information systems, ultimately leading to benefits for the organisation.

The Technology Acceptance Model (TAM) seeks to explain and predict how users accept and adopt new technology (Davis, 1989). TAM is frequently used in the research related to information systems acceptance and technology adoption due to its transparency and adaptability.

TAM2 (Venkatesh and Davis, 2000) and the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh, et al., 2003) can be viewed as extensions of TAM and incorporate additional factors such as voluntariness, job relevance, social influence, gender, and age. While these models have gained traction in the academic literature (Williams, et al., 2015), the original TAM is still widely used and has now been adapted for contemporary research in education related to emerging technologies such as ChatGPT (Saif, et al., 2024; Abdaljaleel, et al., 2024).

While some additional factors would likely be uncovered by using more recent models such as TAM2, TAM3, and UTAUT, there are many potential context-specific factors that could influence adoption of IM within HE which may not be uncovered by existing models. Examples may include organisational culture, student expectations, digital poverty, safeguarding considerations, pedagogical alignment, cross-generational or cross-cultural communication styles, and technological trends such as GenAI.

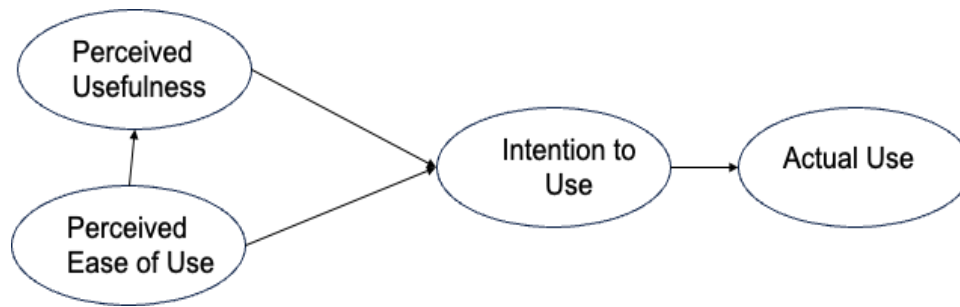
For these reasons, the authors chose TAM as a theoretical framework to support the exploration of barriers and facilitators of IM adoption in academic settings.

While TAM and the Information Systems Success Model mainly address the usability aspects of information systems, the technostress model developed by Ragu-Nathan and colleagues (Ragu-Nathan et al., 2008) takes a distinct perspective on the usage of Information Technology (IT) and focuses on users' well-being and the phenomenon of technostress - stress experienced by end users of IT and its impact on their job satisfaction and commitment to the organisation. This model is less frequently used in the IT usage research, though it can provide useful and insightful perspective on the impact of IT tools, including the experience of using IM.

Other frameworks that articulate the social aspect of communication technology are related to the computer-mediated communication theory, for example Uses and Gratification theory (Weiyan, 2015), but the authors would like to investigate users' well-being along with the usability aspects of IM in Higher Education. For this reason, the TAM and Technostress models are seen as being the suitable frameworks to form a basis for the current research while acknowledging the need for a wider exploration of influencing factors specific to HE. For this reason, the authors are proposing a composite model, an initial version of which is outlined in section 4 of this paper.

### **3.1 Technology Acceptance Model**

The Technology Acceptance Model (TAM), originally proposed by Davis (1989) and represented in Fig. 1, continues to be the prevailing measure of 'usability' with regards to new procedures and processes involving technology. TAM is a widely recognised and validated theoretical framework in the field of information systems and technology management. It aims to understand and predict users' acceptance and adoption of new technologies. It is based on motivational theories and idea that perceived ease of use (PEOU) and perceived usefulness (PU) are key determinants of an individual's willingness to accept and use technology.



**Figure 1. Technology Acceptance Model (TAM).**

### **3.2 Technostress model**

Ragu-Nathan and colleagues developed and validated a conceptual model of technostress (Ragu-Nathan et al., 2008) and is based on the transactional approach to stress that explains stress occurrence as a combination of a challenging stressful condition and the person's response to it (Cooper et al., 2001). Technostress refers to the psychological and physical stress that individuals experience as a result of use of technology (Zielonka and Rothlauf, 2021) and emerged as a concept with the increasing integration of technology into various aspects of our personal and professional lives. The technostress model proposed by Ragu-Nathan et al. (2008) provides a structured framework for understanding the causes and consequences of technostress in the workplace, as represented in Fig. 2.

Differences in age, education, experience, and familiarity with IT define the strength of technostress creators as perceived by an individual. These differences relate to individual beliefs about the usefulness and ease of use of IT. A combination of stressors and technostress inhibitors creates a certain level of Job Satisfaction, Organisational Commitment and Continuance Commitment.

It is appreciated that not all technostress is negative, and some stressors can, in fact, lead to positive outcomes such as improved efficiency and innovation. Depending on an individual's perspective, the stress associated with use of IT could be perceived as challenging in a developmental way (Eustress) or as a threat leading to detrimental outcomes (Distress) as outlined by Tarafdar, Cooper and Stich (2019).

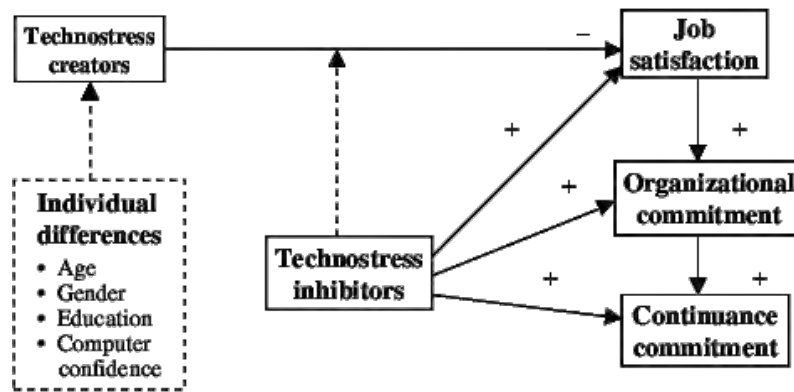


Figure 2. Conceptual Model for Understanding Technostress (Ragu-Nathan et al., 2008).

#### 4.0 Composite Theoretical Model

An initial round of pre-survey exploratory interviews is proposed with the purpose of identifying any influencing factors in the context of HE not revealed by the application of TAM and Technostress.

Based on this initial research, the authors intend to develop and utilise a composite theoretical model to inform the design of a questionnaire to capture the following data regarding academic staff in High Education contexts:

- Current usage levels of IM.
- Stressors (potential and actual) hindering adoption of IM.
- Current perceived ease of use (PEOU) and perceived usefulness (PU) of IM.
- The impact of other influencing factors specific to the HE context.

Based on the work of Ragu-Nathan et al. (2007) the authors aim to assess the strength of a set of technostress creators, categorised as techno-overload, -invasion, -complexity, -insecurity, and -uncertainty. An adapted version of the Technology Acceptance Model questionnaire (Davis, 1989) will be used to capture perceived ease of use and the usefulness of IM. Additional questions will be included, based on other influencing factors identified during pre-survey interviews. The influence of HE-specific factors on PEOU, PU, Technostress creators, and Technostress inhibitors will subsequently be analysed.

## 5.0 Conclusion

This research has explored how a composite theoretical model could be used to identify and classify factors that influence the adoption of IM from an academic staff perspective, which include the technostressors associated with IM and other factors specific to Higher Education, particularly within the contemporary context of emerging GenAI and LLMs.

Interviews will be conducted to inform the development of a survey to be distributed through academic networks to attain a sample which is as representative as possible from among teaching staff within HE. These findings will further deepen our understanding of the research landscape (RQ1) before progression to the next phase of the research.

The next research step would be to propose a set of recommendations for optimal IM adoption in HE and validate these through another round of interviews or focus groups (RQ2).

## References

- Abdaljaleel, M., Barakat, M., Alsanafi, M., Salim, N.A., Abazid, H., Malaeb, D., Mohammed, A.H., Hassan, B.A.R., Wayyes, A.M., Farhan, S.S. and Khatib, S.E. (2024). A multinational study on the factors influencing university students' attitudes and usage of ChatGPT. *Scientific Reports*, 14(1), p.1983.
- Concannon, F., Costello, E., Farrell, O., Farrelly, T. and Wolf, L.G. (2023). There's an AI for that: Rhetoric, reality, and reflections on EdTech in the dawn of GenAI. *Irish Journal of Technology Enhanced Learning*, 7(1).
- Cooper, C. L., Dewe, P., & O'Driscoll, M. P. (2001). Organizational stress: A review and critique of theory, research, and applications.
- Davis, F. D. (1989). *Perceived usefulness, perceived ease of use, and user acceptance of information technology*. *MIS Quarterly*, 13(3), 319–340
- DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. *Journal of management information systems*, 19(4), 9-30.
- De Bakker, G., Sloep, P. and Jochems, W. (2007). Students and instant messaging: a survey of current use and demands for higher education.

- Hrastinski, S. and Aghaee, N.M. (2012). How are campus students using social media to support their studies? An explorative interview study. *Education and Information Technologies*, 17(4), pp.451-464.
- Hurbean, L., Dospinescu, O., Munteanu, V. and Danaiaata, D. (2022). *Effects of Instant Messaging Related Technostress on Work Performance and Well-Being*. *Electronics*, 11(16), 2535.
- Lauricella, S. and Kay, R. (2013). Exploring the use of text and instant messaging in higher education classrooms. *Research in Learning Technology*, 21.
- Lents, N.H. and Cifuentes, O.E. (2010). Increasing student-teacher interactions at an urban commuter campus through instant messaging and online office hours.
- le Roux, D. le Roux, D.B., Parry, D.A. (2022). An exploratory investigation of the use and effects of academic instant messaging groups among university students. *Educational Information Technology* 27, 1055–1080.  
<https://doi.org/10.1007/s10639-021-10631-y>
- Ragu-Nathan, T. S., Tarafdar, M., Ragu-Nathan, B. S., & Tu, Q. (2008). *The consequences of technostress for end users in organizations: Conceptual development and empirical validation*. *Information systems research*, 19(4), 417-433.
- Saif, N., Khan, S.U., Shaheen, I., Alotaibi, A., Alnfai, M.M. and Arif, M. (2024). Chat-GPT; validating Technology Acceptance Model (TAM) in education sector via ubiquitous learning mechanism. *Computers in Human Behavior*, 154, p. 108097.
- So, S. (2016). Mobile instant messaging support for teaching and learning in higher education. *The Internet and Higher Education*, 31, pp.32-42.
- Spira, J. and Feintuch, J. (2005). *The cost of not paying attention: How interruptions impact knowledge worker productivity*. Basex Feintuch 1–21.
- Sun, L. (2015). Facebook Inc.'s WhatsApp hits 900 million users: what now. *The Motley Fool*.
- Tang, Y. and Hew, K.F. (2019). Examining the utility and usability of mobile instant messaging in a graduate-level course: A usefulness theoretical perspective. *Australasian Journal of Educational Technology*, 35(4).
- Tarafdar, M., Cooper, C.L. and Stich, J.F. (2019). The technostress trifecta-techno eustress, techno distress and design: Theoretical directions and an agenda for research. *Information Systems Journal*, 29(1), pp.6-42.
- Venkatesh, V. and Davis, F.D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, 46(2), pp.186-204.
- Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, pp.425-478.

- Weiyang, L. I. U. (2015). A historical overview of uses and gratifications theory. *Cross-Cultural Communication*, 11(9), 71-78.
- WhatsApp (2020). Two billion users—connecting the world privately. WhatsApp Blog.
- Williams, M.D., Rana, N.P. and Dwivedi, Y.K. (2015). The unified theory of acceptance and use of technology (UTAUT): a literature review. *Journal of enterprise information management*, 28(3), pp.443-488.
- Zielonka, J. T., & Rothlauf, F. (2021). Techno-eustress: The impact of perceived usefulness and perceived ease of use on the perception of work-related stressors.



# The Technological Challenges of Blended Learning in Higher Education: A Case Study of Blackboard

## Author Name

Mohamed Daud, Ali Gheitasy, Parisa Saadati and Laden Husamaldin  
[21436149@student.uwl.ac.uk](mailto:21436149@student.uwl.ac.uk), [ali.gheitasy@uwl.ac.uk](mailto:ali.gheitasy@uwl.ac.uk), [parisa.saadati@uwl.ac.uk](mailto:parisa.saadati@uwl.ac.uk),  
[laden.aldin@uwl.ac.uk](mailto:laden.aldin@uwl.ac.uk)

School of Computing and Engineering  
University of West London  
St Mary's Road, Ealing  
London  
W5 5RF

## Abstract

*This research delves into blended learning, the fusion of online and face-to-face education in higher education. Despite the topic's popularity, blended learning faces significant technological challenges. Employing a mixed-methods approach, the study analyses data through literature reviews, surveys, and interviews, exploring these challenges. Data from the University of West London highlights the importance of practical training and support for leveraging platforms like Blackboard effectively. Addressing these challenges can enhance the blended learning experience, fostering engagement and improved outcomes. Thus, this study emphasises the need to recognise and tackle technological challenges in blended learning. Ongoing research is vital to implement best practices and enhance the usability of tools like Blackboard for improved learning experiences. Moreover, the research underscores the potential of blended learning, showing that strategic use of tools can create efficient, engaging education. This study lays a foundation for improving blended learning in higher education by understanding its challenges. It calls for collaborative efforts to optimise this experience, transforming pedagogy.*

**Keywords:** Blended learning, Online learning challenges, User experience

## 1.0 Introduction

Higher education has used blended learning since the late 1990s. Over the following years, it has gained popularity as more and more higher education institutions offer at least some of their programs in a blended format (Mirriahi, et al., 2015). Blended learning refers to a combination of face-to-face interaction between teachers and students with technology-based learning experiences (Venkateswari, 2022). The

significance of blended learning in higher education lies in its capacity to offer more flexibility, accessibility, and personalization of learning experiences. However, implementing blended learning has challenges (Ma'arop & Embi, 2015). Technical challenges include a lack of access to dependable internet, difficulties understanding learning management systems, and problems with online collaboration and communication (Fathema et al., 2015), which can hamper the learning experience for instructors and students. Identifying and resolving these obstacles is crucial for boosting blended learning's overall experience and success.

This research seeks to identify and address the technical issues students and instructors face in blended learning environments, focusing on their experiences with the Blackboard learning management system. This research strives to improve the blended learning experience for all stakeholders by illuminating these obstacles and proposing solutions. A mixed-methods strategy was adopted to achieve these goals, including a literature review, surveys, and interviews with crucial stakeholders at the University of West London. Participants included students and instructors from various academic levels and institutions, resulting in a statistically varied sample.

Blended learning is fraught with several technological challenges, such as problems regarding access, difficulties with technology, a lack of technical understanding, inconsistencies in the use of technology, insufficient degrees of technical assistance, and worries about safety (Johnson et al., 2016). Because of the difficulties presented by these factors, the effectiveness of blended learning may be compromised, which would be detrimental to the performance of online learning integrated with traditional classroom instruction. In addition, it is of the utmost importance to have a deeper understanding of these problems and to develop solutions. Nevertheless, blended learning demands additional research.

The research aims to investigate the technological challenges students and lecturers face in blended learning environments, understand these challenges, and develop solutions to enhance these environments and improve the overall learning experience. The study will focus on one of the leading virtual learning platforms (the Blackboard) as a case study for this research.

The intention is to improve the experiences of both students and instructors and to promote the effectiveness of blended learning in higher education by better understanding these problems and working to address them.

This research structure is as follows: Section 1 includes a literature review that examines previous research on blended learning and its challenges. Section 2 describes the study's methodology, including the data collecting and analytic techniques. Section 3 shows the results of the data analysis and identifies the most significant technological difficulties—section 4 addresses suggestions for addressing these obstacles and enhancing the blended learning experience. Lastly, the conclusion summarises the study's findings and emphasizes the significance of overcoming technological challenges to improve student and instructor experiences in blended learning environments.

## **2.0 Literature Review**

This literature review will investigate the different learning modes and methods, their advantages and disadvantages, and their challenges.

### **2.1 Definition of Blended Learning**

The term 'blended learning' has diverse connotations, and its precise origin remains nebulous, with its first documented usage traceable to a 1999 press release from an Atlanta-based company, EPIC Learning (Friesen, 2012). Despite its ubiquity, the term embodies many pedagogical methods, technologies, and job responsibilities, from conventional mechanisms like film to advanced theories and practices. Blended learning combines face-to-face and online instruction with long-standing histories (Graham et al., 2013).

Driscoll's (2002) research delineates four distinctive interpretations of 'blended learning.' These include integrating different web-based technologies, various pedagogical approaches, instructional technology with in-person instruction, and melding job activities with instructional technology.

Blended learning amalgamates various teaching and learning techniques to foster a dynamic educational experience. These methods include face-to-face instruction, online lectures and videos, online discussions, interactive multimedia, collaborative learning activities, self-paced learning modules, and flipped classrooms (Bishop and Verleger, 2013).

Blended, Face-to-Face, and Online Learning each have unique strengths and weaknesses. Blended Learning leverages the benefits of traditional instruction with the versatility and accessibility of online components, offering a personalized learning

experience, supporting various learning styles, and fostering student engagement (Cents-Boonstra et al., 2021). A meta-analysis by Means et al. (2010) suggests that blended Learning may lead to superior academic outcomes than traditional or entirely online instruction.

Face-to-face learning underscores direct interaction between teachers and students, enabling immediate feedback and fostering real-time discussion (Banna et al., 2015). Conversely, its limitations lie in its lack of flexibility, as learners must conform to fixed schedules and locations.

Online Learning provides maximum flexibility, enabling learners to access content anytime and allowing for asynchronous communication and self-paced Learning (Southard et al., 2015). It can mitigate the need for physical spaces and travel, reducing costs. However, it may not fully replace face-to-face instruction's immediacy and social interaction (Bernard et al., 2009).

## **2.2.0 Advantages and Disadvantages of Blended Learning**

### **2.2.1 Advantages of blended learning**

Blended learning, which combines the advantages of face-to-face and technology-enhanced learning, has emerged as a significant reform in education (Graham et al., 2013). This approach offers numerous advantages for both teachers and students, as outlined below:

- **Flexibility:** Müller and Mildenerger (2021) emphasize that blended learning allows learners to access materials at their own pace and on their schedules, providing greater flexibility in the learning process.
- **Personalization:** Shemshack, A., Kinshuk, & Spector, J. M. (2021), blended learning accommodates different learning styles and allows students to learn at their own pace, fostering a more personalized learning experience.
- **Student engagement:** Chen, p.s.d., Lambert, a.d. and guidrey, k.r. (2010) assert that combining face-to-face interaction with online components promotes active learning and increases student engagement.
- **Access to diverse resources:** Zacharis, N.Z. (2015) notes that blended learning provides students access to various multimedia and online materials, enhancing the learning experience.

- Improved learning outcomes: Al-Samarraie, H., & Saeed, N. (2018) found that blended learning can lead to better academic outcomes than traditional or fully online instruction.
- Collaboration: Al-Samarraie, H., & Saeed, N. (2018) highlight that blended learning fosters student collaboration and communication through group projects and online discussions.
- Cost-effective: Lloyd-Smith, L. (2010) argues that blended learning can reduce costs for institutions and students by minimizing travel, classroom space requirements, and printed materials.

### **2.2.2 Disadvantages of Blended Learning**

While blended learning has gained popularity due to its numerous benefits, researchers have also identified several disadvantages that need to be considered:

**Technological issues:** Namyssova et al. (2019) highlight that limited access to technology, inadequate infrastructure, and technical problems can negatively affect the effectiveness of blended learning. Ensuring reliable technical support is essential for the successful implementation of this approach.

**Digital divide:** Selwyn (2004) addresses the issue of the digital divide, where socioeconomic disparities and unequal access to technology can hinder some students' ability to participate in blended learning. Educational institutions must address these inequalities to ensure that blended learning is accessible to all students.

**Instructor training and support:** Tshabalala et al. (2014) emphasize the importance of providing adequate training and support for instructors in using technology and implementing blended learning. With proper training, instructors may be able to effectively integrate online and face-to-face components, ultimately impacting student learning outcomes.

**Learner motivation and self-discipline:** Gorbunovs et al., (2016) point out that blended learning requires students to exhibit greater self-motivation and discipline to complete online activities and manage their time effectively. Institutions should consider strategies to foster motivation and self-discipline among learners.

**Development and maintenance costs:** Bates (2015) acknowledges that blended learning can be cost-effective in the long run; however, the initial investment in technology, course design, and staff training can be high. Institutions must evaluate

the cost-benefit analysis and determine if blended learning is the most viable approach for their context.

### **2.3 Technological Challenges in Blended Learning**

This section provides previous studies investigated the challenges of blended learning:

- Namyssova et al. (2019) Highlight critical technological and infrastructure problems. The article delves into roadblocks such as a lack of access to technology and appropriate support, providing complete insights into the probable smooth integration of blended learning in educational environments.
- Tshabalala et al. (2014): The researchers identify many challenges to BL, including policy gaps and a lack of faculty support, explicitly emphasising how foundational policy and training might aid in overcoming these barriers in a South African environment.
- Smith et al. (2019) advocate for rigorous governance and leadership to support the effective implementation of BL tactics in academic settings, emphasising the importance of defined BL goals and substantial teacher preparation.
- Mirriahi et al. (2015) emphasises the necessity of clear institutional definitions and robust staff capacity in BL, demonstrating that lacking these factors leads to misconceptions and misapplications of BL approaches.
- Hofmann (2011) divides BL problems into three categories: technological, organisational, and instructional, demonstrating how unthinking technology adoption can be a barrier rather than a facilitator for successful BL deployment.

### **2.4 Previous Studies on Blended Learning**

López-Pérez et al. (2011), affiliated with the University of Granada, conducted a study involving 1,431 students across 17 groups. Their research revealed that blended learning, a combination of online and in-person instruction, was associated with reduced dropout rates, improved exam performance, and the identification of correlations between student perspectives and factors such as age, academic background, and attendance.

Poon (2013) from Nottingham Trent University investigated the advantages of blended learning through a methodology that included interviews and surveys. The

findings indicated that blended learning significantly enhanced higher education experiences, identified critical success factors, and informed recommendations for effective teaching methods.

Singh et al. (2021) delved into the evolution of blended learning during the initial stages of the COVID-19 pandemic. Their study encompassed an analysis of challenges, strengths, weaknesses, opportunities, and threats associated with blended learning. The research offered practical strategies for effectively integrating traditional and online teaching, particularly during crises.

In 2022, Ruiz-Alonso-Bartol et al. studied the transition to online learning during the COVID-19 pandemic. Their findings gathered through questionnaires, teacher journals, and interviews indicated reduced participant stress levels. While some students appreciated the flexibility of online classes, others expressed mixed feelings about the impact on their learning. Notably, teachers recognized the benefits of smaller online sessions.

In conclusion, these studies collectively underscore the significance of blended learning in contemporary education. They provide evidence that blended learning can reduce dropout rates, enhance student experiences, and offer valuable insights into the challenges and opportunities presented by hybrid instruction, especially in times of crisis like the COVID-19 pandemic. While this instructional mode offers flexibility, it also poses challenges that must be addressed to ensure the continued effectiveness of blended learning approaches.

### **3.0 Methodology**

This research adopts a mixed-methods approach that integrates quantitative and qualitative techniques. Creswell et al. (2017) argue that combining questionnaires and interviews as a concurrent “mixed methods research” approach can reduce researcher bias. Surveys include a broader range of participants but may fall short of providing a depth of understanding of their requirements. On the other hand, interviews can offer more in-depth insights but are limited to a smaller group. This approach facilitates the gathering of both quantitative and qualitative data. This triangulation by incorporating both data types, help researchers to reduce bias in the findings and enhance the reliability of the results.

### **3.1 Data Collection**

To discover the problems with technology in BL, we used a combination of a survey and interviews. We targeted both students and lecturers for our survey. Before the main data collection, a pilot study conducted to validate and enhance the survey and interviews questions. Following the pilot study some lesson learned and adjustments were made to some of the questions to improve the comprehensibility of the participants (mainly students). We received 50 responses from students and 7 responses from the lecturers. Of these numbers, 8 students and 3 lecturers agreed to have a follow-up interview over the phone, and we used the transcript of the recorded interviews in the analysis. The survey questionnaire had 20 Likert scale questions for the students and 20 for lecturers to measure how often and how lousy technology problems were in BL. These questions discussed factors like access to technology, problems with online tools, and difficulties with communication and working together. During the interviews, we asked six questions from students, whom we will refer to as S1-S8 and lecturers, whom we will refer to as L1-L3 in this study. These questions helped us to get more information about their experiences and opinions about the technical problems in BL. The questions were meant to discover their experiences with technology, how it affects their learning and teaching, and what could be done to resolve the issues.

### **3.2 Data Analysis**

In this study, we employed a mixed-method approach to data analysis. The primary data collection goal was to assess how technology usage challenges among students and teachers affect blended learning environments. This approach addresses the technological obstacles present in blended learning environments (Creswell and Plano Clark, 2017).

Initially, the researcher conducted a quantitative analysis of survey responses using Microsoft Excel. The survey was designed to produce data that could be analysed using descriptive statistical methods. This has helped calculate the frequencies and percentages based on responses to Likert-scale questions, providing an overall picture of the challenges in the blended learning environment.

In parallel with the quantitative analysis, the researcher also conducted a qualitative analysis of interview data. This process began with the transcription of recorded interviews, followed by thematic analysis to identify recurring themes related to



technological obstacles. After identifying these themes, we developed a coding scheme and applied it to the transcripts using NVivo software. The coded data were then analysed to identify patterns and problems in blended learning.

The combination of quantitative and qualitative analysis gave us a holistic view of the challenges faced in blended learning environments. To improve the reliability of the results, a third-party expert researcher checked the quantitative analysis results as well as the thematic analysis. Then she carried out a sample of coding procedure independently. Consequently, the results from the third-party researcher were compared with the main results. Both results were consistent, the coded data matched well with each other, and the third-party expert validated the analysis of the survey responses. In the following sections, we will offer an in-depth exploration and deliberation of the outcomes derived from this analysis.

## **4.0 Results**

This section will summarise the results of all analyses done in this research by starting the surveys and then the interviews.

### **4.1 Summary of the Students Survey Results**

The survey findings indicate that most students favour the online component of blended learning and Blackboard. Most respondents agreed or strongly agreed with the benefits of the online component, including providing additional resources, the freedom to learn in their own space, and the fact that it is reliable and convenient. However, some students need help grasping the course topics, engaging with their peers, and navigating Blackboard. Refer to the Bar Chart below (Figure 1).

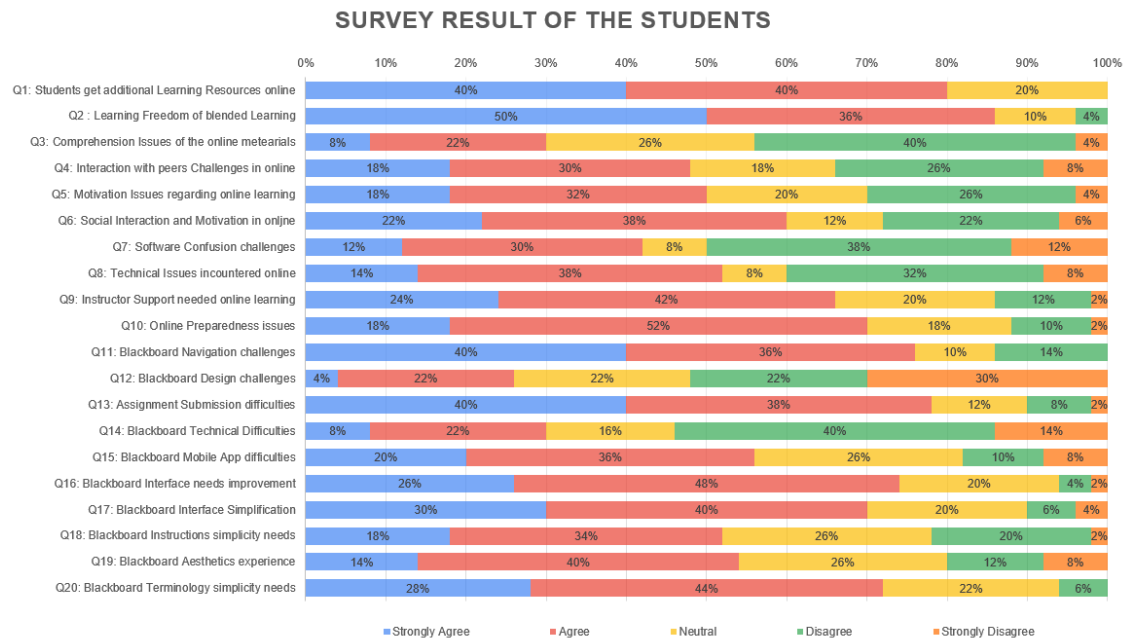


Figure 1. Student’s survey outcome. (Source: screen capture from Excel sheet.)

## 4.2 Summary of the Lecturers’ Survey Results

The analysis of a survey conducted to assess lecturers' perspectives on blended learning technology using the Blackboard as a case study reveals mixed responses from lecturers; while a significant portion struggles with producing high-quality internet videos and is wary of new technologies used in online teaching, many claims to be technically proficient and at ease creating content for Learning Management Systems such as Blackboard. Despite this, a sizable proportion finds the procedure difficult and recognises the need for comprehensive training.

Compared to traditional face-to-face approaches, lecturers have highlighted concerns about the quality of online learning. Despite their belief in the efficiency of blended learning technologies and their opinion of students' appropriate ICT (Information Communication Technology) skills, they emphasise the need for further help to better enlighten students about online resources.

Regarding accessibility, Blackboard is regarded as providing equal access to learning resources for impaired students despite proposals for enhancements to the user interface. The summary emphasises that blended learning technology has its obstacles, emphasising the importance of ongoing assistance and training to improve instructors' capacity to teach online. Refer to the bar chart below (Figure 2) for more details.

## SURVEY RESULTS: LECTURERS BLENDED LEARNING HURDLES

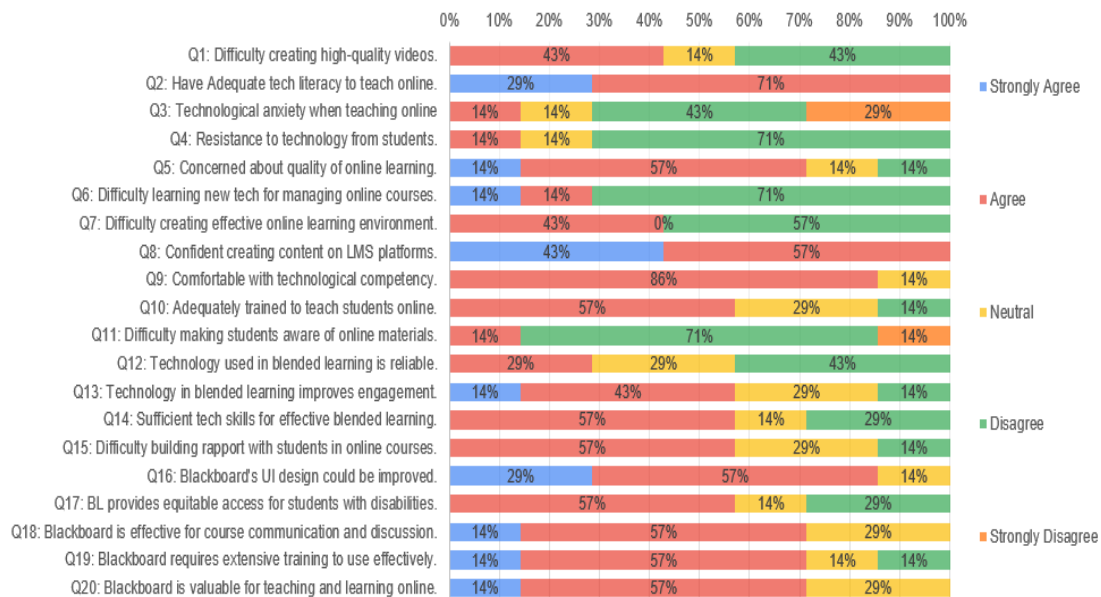


Figure 2. lecturers' survey outcome (Source: screenshot from Excel.)

### 4.3 Summary of the Interview Results of the Students

This research interviewed eight students to gather their experiences and perspectives on blended learning environments, mainly using the Blackboard learning management system (LMS). The students expressed a variety of preferences for their learning environment, with some favouring face-to-face, blended, or online learning approaches. These modalities offer different levels of interaction, engagement, and convenience, depending on the student's learning style and the tools they use. Out of the eight students who expressed a preference for their learning environment, five preferred face-to-face learning (S1,2,3,6,8), highlighting the importance of personal interaction (S1) and the value of social interaction with classmates and educators (S2). Three students favoured blended learning (S4,5,7), citing the flexibility it provides (S4) and the effectiveness of combining online and face-to-face modalities (S5).

The students also discussed the technological challenges they encountered within blended learning environments. These challenges include: 1. Maintaining engagement during online sessions (S1) expressed difficulty staying engaged during online classes compared to face-to-face sessions, which can negatively impact their learning experience. 2. Limitations in online tools for interaction (S2). S2 also noted, "*Some online tools provide a different level of interaction than face-to-face classes, making it harder to collaborate and communicate effectively.*" 3. Difficulties navigating

Blackboard's layout (S4). S4 found the layout of the Blackboard LMS confusing, making it challenging to find the information and resources they need. 4. Lack of customization options (Student 5): Student five expressed the desire for more personalized learning experiences within Blackboard, such as the ability to customize their learning path or access specific resources. 5. Difficulties with notifications (Student 7) Student seven mentioned that they sometimes miss important information or deadlines due to the ineffective notification system in Blackboard.

The students offered suggestions to enhance the blended learning experience. S1 mentioned, *“The importance of smaller group sizes for more effective blended learning.”* S2 also suggested *“Adding a visible progression rate on Blackboard to enhance the learning experience.”* S4 emphasized *“The need for improving notifications and simplifying Blackboard's layout for more straightforward navigation.”* S5 proposed *“better customization options.”*. While S7 recommended *“Improving the notification system.”*. Moreover, that shows that he agrees with S4. Furthermore, S6 discussed their experiences with Blackboard and compared it to another platform they had used, which they could not recall. Moreover, they shared the following: *“Yes, I use one, but I need to remember the name. I don't know if it is blended or packed online, but with this platform, you can log in on your space. You can listen to your course video and have an opening for yourself. For practising notion and also for assessment, you can see yours. Your progression rates, like, yes, I have listened. For the lesson, I am at 40%. For the whole program, this is what the Blackboard does not have. I do not see. I wonder if this is available.”*

These insights from the analyses can be instrumental in developing effective strategies for blended learning tools in higher education. By focusing on the themes and insights gathered from the students' responses and analysing the data using NVivo, it is also possible to gain a deeper understanding of the challenges and potential improvements in blended learning environments for future researchers. Below are some NVivo analyses conducted in this study (Figure 3).

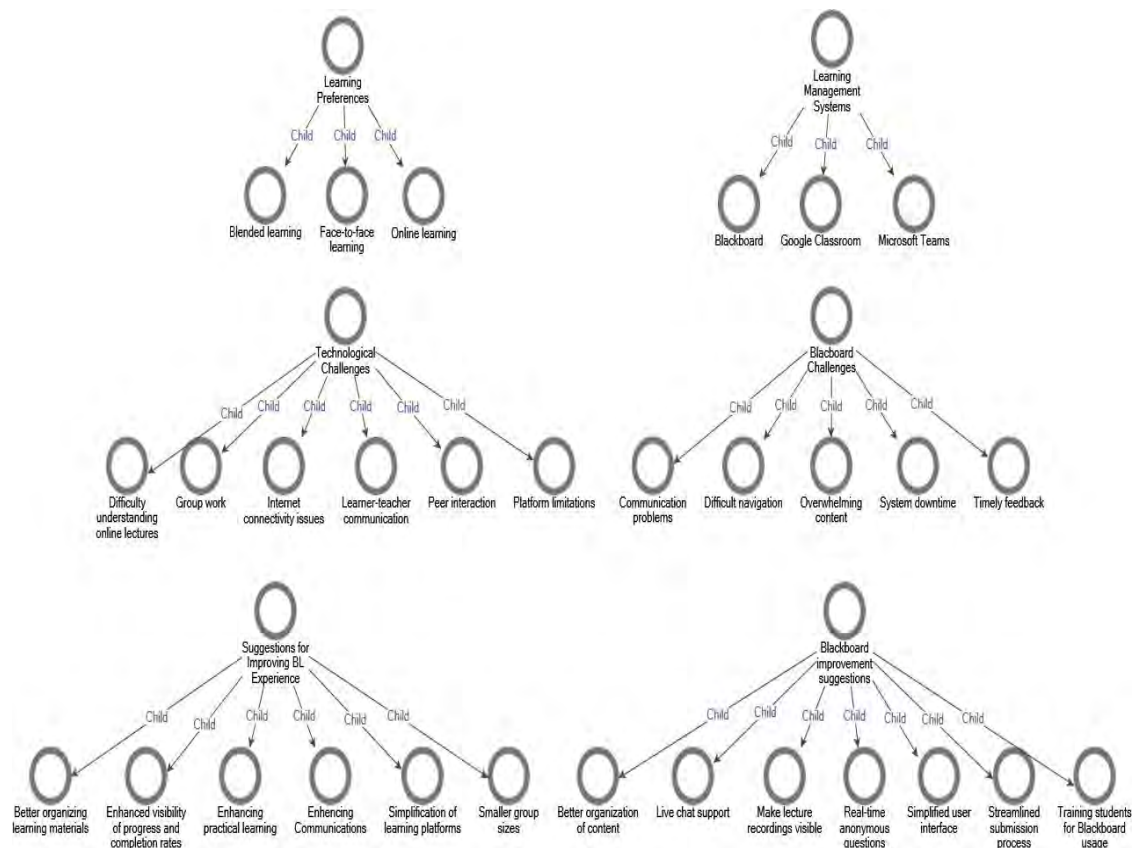


Figure 3. Students Project Map from students' interview results. (Source: Created in NVivo.)

#### 4.4 Summary of the Interview Results of the Lecturers

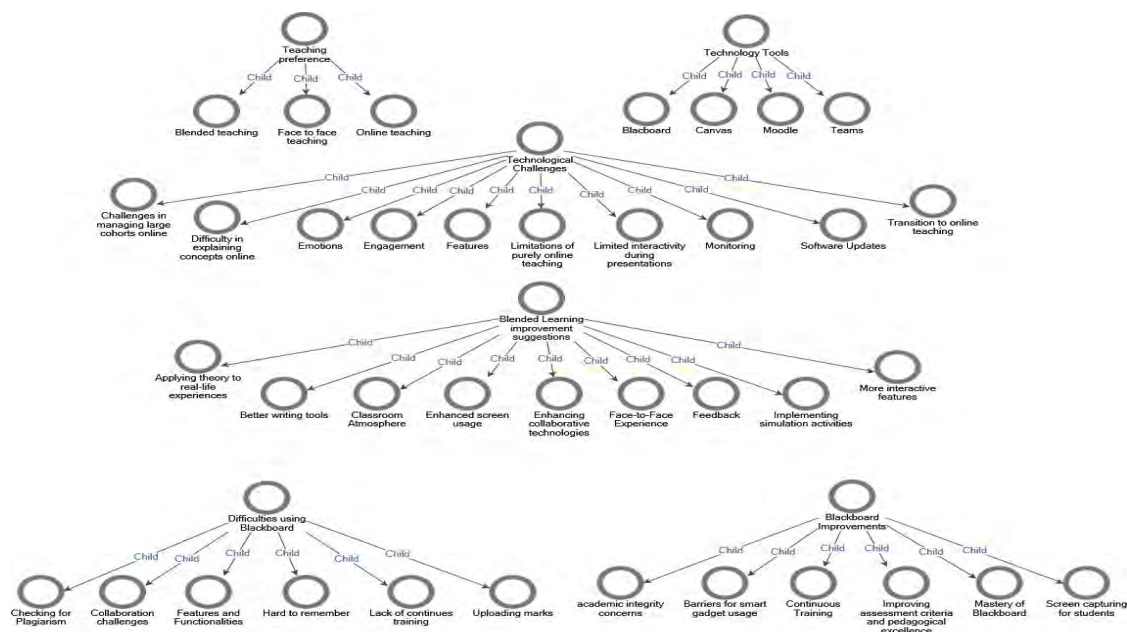
In this research, three lecturers were interviewed to share their experiences and discuss the technological challenges they encountered within blended learning environments, specifically when using the Blackboard learning management system (LMS). There were various preferences among the lecturers, with some favouring face-to-face, blended, or online learning approaches. These modalities offer different levels of interaction, engagement, and convenience, depending on the lecturer's teaching style and the tools they use.

L1, who preferred face-to-face learning, highlighted the challenge of maintaining student engagement in online sessions, stating, "*We do not really have that much engagement.*". L2 favoured face-to-face learning as well and mentioned the limitations in online tools for interaction, saying, "*I prefer to scribble on the whiteboard when in the classroom while teaching, and I feel that there is some sort of restriction in being able to write and explain and go a little bit deeper into whatever it is that you are trying to teach.*". L3 preferred blended learning, citing the effectiveness of synchronizing and asynchronous activities.

The lecturers offered suggestions to enhance the blended learning experience. L1 mentioned “*the importance of simulating an authentic classroom atmosphere.*”. L2 suggested improving writing tools and embedded quizzes, explaining, “*Having better writing tools and embedding quizzes would keep students more engaged during online sessions.*”. L3 emphasized “*the need for continuous training for faculty and addressing concerns related to artificial intelligence and academic integrity.*”

L3 also said, “*Wouldn't it be nice to capture students' screens onto the tutor screen to see how the students are getting on are particularly useful for practical modules. So, if I asked my students to do some coding for a particular task, I wanted to check whether everybody was quickly on the same wavelength. It would be nice to capture individual screens or mine and share good work with others. So, it would be nice to have this two-way system where the students can see my screen, which is standard by default. Nevertheless, also, for me to capture individual students' screens.*”

By focusing on the themes and insights gathered from the lecturers' responses and analysing the data using NVivo, it is possible to gain a deeper understanding of the challenges and potential improvements in blended learning environments. These insights can be instrumental in developing effective strategies and tools for blended learning in higher education. Below are some of the NVivo analyses conducted in this research (Figure 4).



**Figure 4 Project Coding Map form from the lecturers' interview analysis. (Source: Created in NVivo)**

#### **4.5 Analysis and Interpretation of the Results**

The study identified common challenges and opportunities within blended learning environments, mainly using the Blackboard LMS. The insights from the interviews and surveys of students and lecturers provide valuable perspectives that can guide institutions, educators, and LMS developers in refining their blended learning offerings. Key findings and their implications are:

- **Technological Challenges:** Both students and lecturers identified maintaining engagement during online sessions, limitations in online tools for interaction, and difficulties navigating the Blackboard LMS as significant challenges. These may be attributed to the nature of online learning, the design and features of the LMS, or the participants' learning styles and preferences. Addressing these challenges is crucial for enhancing the blended learning experience.
- **Learning Modalities Preferences:** The study found varied preferences among students and lecturers for face-to-face, blended, and online learning approaches. This diversity suggests that institutions should consider offering a mix of learning modalities to cater to different needs and preferences.
- **Suggestions for Improvement:** Participants provided several suggestions to enhance the blended learning experience, such as improving navigation and search systems, implementing a visible progression rate, developing a screen monitoring feature, and creating a customizable dashboard and notification system. These suggestions reflect their desire for a more efficient and engaging learning experience and can guide institutions and LMS developers in refining their platforms and support services.

**Comparison with Previous Studies:** The findings of this study (Rohmani & Andriani, 2021) align with and expand upon the results of previous research in the field of blended learning. This comparison identifies persistent challenges, improvement areas, and potential future research directions.

#### **5.0 Discussion and Conclusion**

This research explored the experiences and perspectives of students and instructors in blended learning environments, focusing on the technological challenges they face and the potential improvements they suggest.

Our findings regarding technological challenges in blended learning echo those of previous research, emphasizing that issues related to engagement, interactivity, and navigation in online learning environments persist (Riel et al., 2016). The varied



preferences for learning modalities also align with earlier findings (Hrastinski, 2019), confirming the importance of tailoring the educational experience to cater to the distinct requirements of each student and instructor.

Our work expands upon previous studies by suggesting specific improvements, such as refining notifications, simplifying layouts, and adding visible progression rates. While these recommendations align with prior research ((Green et al., 2019), they contribute to a deeper understanding of how blended learning environments can be improved to meet the users' needs more effectively.

This study explored the technological barriers to effective blended learning in higher education, employing a mixed-methods approach incorporating surveys and interviews. The results from this investigation offer a comprehensive understanding of the challenges students and instructors face in a blended learning environment.

The research revealed significant issues affecting students and instructors in blended learning environments. These include maintaining engagement during online sessions, limitations in online interaction tools, and difficulties in navigating the Blackboard Learning Management System (LMS).

Like all research, this study has limitations. The small sample size and specific participant selection are limitations of this study, and the participants in the study were chosen from a particular group of people, which might limit the generalizability of our findings. Additionally, our reliance on self-reported data may be subject to recall or social desirability bias. While our study concentrated on the Blackboard LMS, we acknowledge the existence of other LMS platforms that offer different experiences. Comparative studies of these platforms can further comprehend the strengths and weaknesses of each system, informing the choice and design of LMS in future blended learning initiatives.

Future research should continue to explore and potentially integrate more advanced analytics or predictive features to provide deeper insights into student behaviour and progress. One of the limitations of this research was the limited response size of our data collection. In our future study we aim to expand our data collection size and apply the Cronbach's alpha (Agbo, 2010) technique to measure and improve the internal consistency, and to better validate the reliability of our data. Researchers should assess the effectiveness of these tools in a real-world educational setting across different contexts and levels to validate their impact and further refine them based on user feedback and experience.



## 6.0 References

- Agbo, A.A., 2010. Cronbach's alpha: Review of limitations and associated recommendations. *Journal of Psychology in Africa*, 20(2), pp.233-239.
- Al-Azawei, A., Parslow, P. and Lundqvist, K., 2016. Barriers and opportunities of e-learning
- Bates, T. (2015). *Teaching in a digital age: Guidelines for designing teaching and learning*. Tony Bates Associates Ltd.
- Berdanier, C. and Lenart, J. (2020) *So, You Have to Write a Literature Review*. 1st ed. Wiley. Available at: <https://www.perlego.com/book/1881874/--so-you-have-to-write-a-literature-review-a-guided-workbook-for-engineers-pdf> (Accessed: 2nd March 2023).
- Bishop, J. L., & Verleger, M. A. (2013). The flipped classroom: A survey of the research. In *ASEE National Conference Proceedings*, Atlanta, GA (Vol. 30, No. 9, pp. 1-18).
- d.school (2021) *The design thinking process*. Stanford University Hasso Plattner Institute of Design. Available at: <https://dschool.stanford.edu/resources/getting-started-with-design-thinking> (Accessed: 8 May 2023).
- Friesen, N., 2012. Defining blended learning. *Learning Spaces*, p.1. Available at: [https://www.normfriesen.info/papers/Defining\\_Blended\\_Learning\\_NF.pdf](https://www.normfriesen.info/papers/Defining_Blended_Learning_NF.pdf). (Accessed 6 Feb 2023)
- Graham, C., Dziuban, C. and Picciano, A., (2013) *Blended Learning*. 1st edn. Taylor and Francis. Available at: <https://www.perlego.com/book/1557356/blended-learning-research-perspectives-volume-2-pdf> (Accessed: 1st March 2023).
- Hofmann, J. (2014) *Solutions to the Top 10 Challenges of Blended Learning*. InSync Training, LLC. Available at: <https://www.insynctraining.com/> (Accessed: 20<sup>th</sup> February 2023).
- Kintu, M.J., Zhu, C. and Kagambe, E., 2017. Blended learning effectiveness: the relationship between student characteristics, design features and outcomes. *International Journal of Educational Technology in Higher Education*, 14(1), pp.1-20.

López-Pérez, M.V., Pérez-López, M.C. and Rodríguez-Ariza, L., 2011. Blended learning in higher education: Students' perceptions and their relation to outcomes. *Computers & Education*, 56(3), pp.818-826.

Ma'arop, A.H. and Embi, M.A. (2015) Implementation of blended learning in Higher learning institutions: A review of the literature. *International Education Studies*. Canadian Center of Science and Education. Available at: <https://eric.ed.gov/?id=EJ1093338> (Accessed: 19 February 2023).

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluating Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies*. Washington, D.C.: U.S. Department of Education.

Mirriahi, N., Alonzo, D., & Fox, B. (2015). A blended learning framework for curriculum design and professional development. *Research in Learning Technology*, 23(1), 28451. Available at: <https://doi.org/10.3402/rlt.v23.28451> (Accessed: 12 February 2023).

Namyssova, G., Tussupbekova, G., Helmer, J., Malone, K., Afzal, M., & Jonbekova, D. (2019). Challenges and benefits of blended learning in higher education. *International Journal of Technology in Education (IJTE)*, 2(1), 22-31.

Naveh, G., Tubin, D. and Pliskin, N., (2010). Student LMS use and satisfaction in academic institutions: The organizational perspective. *The internet and Higher Education*, 13(3), pp.127-133. Available at: <https://www.sciencedirect.com/science/article/abs/pii/S1096751610000205?via%3DiHub> (Accessed: 25 March 2023).

Rogers, Y., Sharp, H. and Preece, J. (2023) *Interaction Design*. 6th edn. Wiley. Available at: <https://www.perlego.com/book/3856078/interaction-design-beyond-humancomputer-interaction-pdf> (Accessed: 18 April 2023).

Rohmani, N. & Andriani, R. (2021) 'Correlation between academic self-efficacy and burnout originating from distance learning among nursing students in Indonesia during the coronavirus disease 2019 pandemic', *Journal of Educational Evaluation for Health Professions*, 18(9). Available at: <https://doi.org/10.3352/jeehp.2021.18.9> (Accessed: 25 April 2023).

Ruiz-Alonso-Bartol, A., Querrien, D., Dykstra, S., Fernandez-Mira, P. and Sanchez-Gutiérrez, C., 2022. Transitioning to emergency online teaching: The experience of Spanish language learners in a US university. *System*, 104, 102684. Available at:

<https://www.sciencedirect.com/science/article/pii/S0346251X21002384> (Accessed: 12 February 2023).

Singh, J., Steele, K. and Singh, L., 2021. Combining the Best of Online and Face-to-Face Learning: Hybrid and Blended Learning Approach for COVID-19, Post Vaccine, & Post-Pandemic World. *Journal of Educational Technology Systems*, 50(2), pp.140-171. Available at: <http://journals.sagepub.com/doi/pdf/10.1177/00472395211047865> (Accessed: 10<sup>th</sup> February 2023).

Tshabalala, M., Ndeya-Ndereya, C., & Van der Merwe, T. (2014). Implementing blended learning at a developing university: Obstacles in the way. *Electronic Journal of e-Learning*, 12(1), 101-110.

Venkateswari, S.L. (2022) 'Blended Mode of Teaching and Learning in Higher Education', *Language in India*, 22(5), pp. 123–131. Available at: <https://discovery.ebsco.com/linkprocessor/plink?id=776b3a85-190e-34a6-aa25-dcbac0cb4a40> (Accessed: 19 February 2023).

Cents-Boonstra, M., Lichtwarck-Aschoff, A., Denessen, E., Aelterman, N. and Haerens, L., (2021) Fostering student engagement with motivating teaching: an observation study of teacher and student behaviours. *Research Papers in Education*, 36(6), pp.754-779.

Müller, C. and Mildenerger, T., (2021) Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education. *Educational Research Review*, 34, p.100394.

Shemshack, A., Kinshuk and Spector, J.M., (2021) A comprehensive analysis of personalized learning components. *Journal of Computers in Education*, 8(4), pp.485-503.

Chen, P.S.D., Lambert, A.D., & Guidry, K.R. (2010). Engaging online learners: The impact of Web-based learning technology on college student engagement. *Computers & Education*, 54(4), 1222-1232.

Zacharis, N.Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education*, 27, 44-53.

Nortvig, A.M., Petersen, A.K., & Balle, S.H. (2018). A literature review of the factors influencing e-learning and blended learning in relation to learning outcome, student satisfaction and engagement. *Electronic Journal of E-learning*, 16(1), 46-55.

- Al-Samarraie, H., & Saeed, N. (2018). A systematic review of cloud computing tools for collaborative learning: Opportunities and challenges to the blended-learning environment. *Computers & Education*, 124, 77-91.
- Lloyd-Smith, L. (2010). Exploring the advantages of blended instruction at community colleges and technical schools. *MERLOT Journal of Online Learning and Teaching*, 6(2), 508-515.
- Riel, J., Lawless, K.A. and Brown, S.W., 2016. Listening to the teachers: Using weekly online teacher logs for ROPD to identify teachers' persistent challenges when implementing a blended learning curriculum. *Journal of Online Learning Research*, 2(2), pp.169-200.
- Hrastinski, S. (2019). What do we mean by blended learning? *TechTrends*, 63(5), 564-569.
- Green, D., Allan, C.N. and Crough, J., (2019) What Is the Purpose? Using Blended Learning Designs to Purposefully Focus on Student Engagement, Support and Learning. *Blended Learning Designs in STEM Higher Education: Putting Learning First*, pp.35-58.
- Southard, S., Meddaugh, J. and France-Harris, A., (2015) Can SPOC (self-paced online course) live long and prosper? A comparison study of a new species of online course delivery. *Online Journal of Distance Learning Administration*, 18(2), p.8.
- Gorbunovs, A., Kapenieks, A. and Cakula, S., (2016) Self-discipline as a key indicator to improve learning outcomes in e-learning environment. *Procedia-Social and Behavioral Sciences*, 231, pp.256-262.
- Banna, J., Lin, M.F.G., Stewart, M. and Fialkowski, M.K., (2015) Interaction matters: Strategies to promote engaged learning in an online introductory nutrition course. *Journal of online learning and teaching/MERLOT*, 11(2), p.249.
- Creswell, J.W. and Clark, V.L.P., 2017. *Designing and conducting mixed methods research*. Sage publications.

# Digital Ethics: Resolving “Wicked” Problems and Dilemmas

*Completed Research*

## **Abstract (around 150 words)**

*Advances in generative Artificial Intelligence (AI), hardware, the Internet of Things (IoT), smart phones, and software, especially Machine Learning (ML) and social media, are creating new ethical challenges globally. Digital ethics challenges impact Information Technology (IT) professionals. IT professionals and managers play a crucial role in ensuring that digital data are captured, stored, processed, and used in an ethical manner. Digital ethics questions are often “wicked” problems for multiple stakeholders. Recently, political actors have brought greater attention to the ethical implications of capturing, and using digital data, including monetising data. Guidance for resolving problems, including laws, e.g., the European Union’s General Data Protection Regulation (GDPR) and ACM and IEEE Professional Ethics codes, are evolving. This article defines what is meant by digital ethics and digital trust and explores five “wicked” problems and dilemmas associated with digital ethics. By evaluating the dilemmas associated with making ethical decisions related to data from a normative ethics perspective, principles and strategies are proposed to help managers and IT professionals mitigate digital ethics challenges and support ethical decision making within their organisations. Embracing ethical practices contributes to the well-being of individuals and society but also helps build trust in the technology industry. Digital ethics is an important, neglected topic in academic degree programmes and organisations. Future discussions of standards and professional codes of practice should emphasise win-win solutions. Potential research should define the scope of the problem and assess the harms that are occurring or that may result from the misuse of digital data.*

**Keywords:** Ethics, digital ethics, digital data, code of conduct, normative ethics, IT professionals.

## **1.0 Introduction**

The context of business ethics has changed dramatically with the digital transformation of business. Rapid adoption and innovation with digital technologies, expanding data collection, increasing competition, and an interconnected global economy are creating new ethical dilemmas and issues for IT professionals and managers. The problems differ based upon political and social differences across countries. For example, the monetisation of data is of concern in market-driven,

capitalist economies. Expedient actions to adopt and use IT are convenient and practical, but there may be improper or immoral uses.

Every aspect of an organisation disrupted and altered by IT impacts stakeholder trust. In addition, there is a global crisis of trust that is gaining momentum. The trust issue extends to government, large technology organisations, news/journalism organisations, and more recently public healthcare professionals and related international bodies like the World Health Organisation (WHO). With an increase in the number of devices (especially the Internet of Things), there are many new ways by which personal data is accessed, collected, processed, and communicated (Côte-Real, Ruivo, & Oliveira, 2020). This digitisation is changing the way businesses create value, with data being considered a core valuable asset (Fleckenstein & Fellows, 2018). Further, the use of ChatGPT and other generative Artificial Intelligences (AI) could affect new technology, work, and employment in the future (George and George, 2023). ChatGPT is a subset of generative AI that can be applied to create diverse types of data (e.g., text, code, audio, images, and videos) which is designed to follow human conversational instructions and provide a detailed response/predictions based on that input (Korzynski et al. 2023). Using IT has blurred many boundaries between private and public life, resulting in privacy concerns for individuals (Bélanger & Crossler, 2011; Price, 2020; Stahl, 2013; Yun, Lee, & Kim, 2019). As evidenced by the passage of the General Data Protection Regulations (GDPR), people are increasingly concerned about protection of their personal data and what happens to it once it is shared with a third party (Consumers International, 2018). As technology becomes more powerful and misuse of digital data is a possibility, IT professionals and managers should assume primary responsibility. Part of assuming responsibility is developing an understanding of digital ethical principles and the “*wicked*” problems that must be resolved in the context of ethical dilemmas.

As ethical issues are identified and resolved, managers commonly develop policies and frameworks related to the issues, develop compliance approaches, and reduce risks associated with non-compliance. The broad issues of trust, privacy, and ethics are not new for managers. IT professionals are generally not prepared to resolve the “*wicked*” digital ethical problems and dilemmas. There is a need for more discussion and research associated with digital ethics. Google trends indicate searches for

"digital ethics" phrase have declined since 2004. A Google Scholar search for "digital ethics" returns about 2,610,000 results, but no articles are frequently cited. For example, the seventeenth most relevant citation is a book on Digital Media Ethics. It is tangential to digital ethics broadly defined and with 447 citations has many more citations than any of the prior 16 articles. The first 16 are each cited an average of 18.4 times. For these reasons, it is important to examine emerging ethical issues related to digital data that are increasingly confronting IT professionals and managers.

The aim of this paper is to consider digital ethics as a "wicked problem" for IT professionals, especially IT managers. A wicked problem is any decision issue or problem that is difficult or seemingly impossible to resolve because of incomplete, contradictory, and changing requirements (Graves & Garrett Jr, 2014). Wicked problems are often difficult to identify and recognise and as a result have not been widely adopted in management (Gharehgozli, Mileski, Adams, & von Zharen, 2017). Drawing on recent examples, we explore some of the common ethical dilemmas encountered in digital business and we briefly consider these through three normative ethics perspectives.

## **2.0 Defining Ethics, Digital Ethics, and Digital Trust**

Ethics refers to applying "moral rules, codes, or principles which provide guidelines for right and truthful behavior in specific situations" (Lewis, 1985, p. 382). There are several ethical schools of thought, three prominent normative ethics lenses include deontological, consequentialist, and virtue ethics. Deontological ethics, often associated with Immanuel Kant, focuses on the inherent nature of actions rather than their consequences (Alexander & Moore, 2021). It asserts that certain actions are inherently right or wrong, regardless of the context or outcomes. The central principle is the "categorical imperative", which promotes action only according to principles that could be applied universally (Railton, 1986). Keeping promises, telling the truth, and respecting the autonomy of individuals are considered intrinsically moral, regardless of the consequences. Consequentialist ethics, represented by utilitarianism, evaluates the morality of actions based on their outcomes or consequences with the overarching aim to maximise happiness, pleasure, or well-being and minimise suffering or pain (Bentham, 1789; Mill, 1861; Sidgwick, 1907). If an action leads to

the greatest overall happiness for the greatest number of people, it is considered morally right under consequentialist ethics. Virtue ethics focuses on the development of virtuous character traits rather than rules or consequences (Hursthouse & Pettigrove, 2023). It focuses on the importance of cultivating virtues to lead a morally good life. Virtues include characteristics such as honesty, courage, compassion, and justice. For example, a virtuous person, guided by traits like honesty and compassion, is more likely to make morally sound decision in a given situation. These three schools of normative ethics provide different perspectives on how to determine what is morally right or wrong. Deontological ethics looks at the inherent nature of actions, consequentialist ethics evaluates outcomes, and virtue ethics emphasises a virtuous character. Many ethical frameworks and theories are variations or combinations of these fundamental approaches.

Digital ethics involves applying ethical principles and considerations to the specific context of digital technologies, their development, deployment, and use. Data misuse and privacy abuses have motivated the development of digital ethics focusing on the way technology and digital data are altering, shaping, and disrupting our political, business, social, and moral existence. In their Gartner report, Buytendijk, Vashisth, Duncan, and Moran (2016) define digital ethics as *“the system of values and principles a company may embrace when conducting digital interactions between businesses, people, and things. Digital ethics sits at the nexus of what is legally required; what can be made possible by digital technology; and what is morally desirable...”*. Also, increasingly people mistrust those who collect and use data. Managers and IT professional must understand what is possible, viable, and ethical in the collection and use of data.

Compromising an individual’s privacy, vis-à-vis IT, has fundamental ethical implications (Baldini, Botterman, Neisse, & Tallacchini, 2018; Bowie & Jamal, 2006; Sarathy & Robertson, 2003; Spinello, 1998). As business managers continue to grapple with the challenges and complexities of digital transformation (Mohamed & Mohamed, 2020), they are also responsible for assessing digital products and services where ethical concerns should guide their professional decisions (Martin, 2020; Ritter & Pedersen, 2020).



Managers and IT professionals in private, profit-making organisations especially need to create and maintain the perception and reality of digital trust by clients, customers, employees, suppliers, government regulators, and other stakeholders. Meeting this goal requires regularly measuring and quantifying stakeholder expectations to ensure that reality and perception are accurate. To encourage and promote good practices and ethical decision making, managers need to adopt organisational strategies that ensure that the organisation is first compliant, and more importantly *"doing the right thing"*, when it comes to leveraging big data analytics and data mining techniques to generate and possibly monetise and share insights from data.

Digital ethics questions are often *"wicked"* problems for managers, IT professionals, organisations, academics, and government decision makers. According to Mitroff and Mason (1973), an unstructured or *"wicked"* decision problem is one where all the variables, terms, or sets are totally unknown or not known with any high degree of *"confidence."* To tackle a *"wicked problem, the decision-maker needs to form a new appreciation of the situation"* (Mitroff & Mason, 1973). The Austin Center for Design (AC4D), a group focused on targeting social and technological challenges, characterise a wicked problem as *"a social or cultural problem that is difficult or impossible to solve for as many as four reasons: there is incomplete or contradictory knowledge, the number of people and opinions involved, the large economic burden, and the interconnected nature of these problems with other problems."* Wicked problems are messy and challenging to solve (Austin Center for Design (AC4D), 2018).

Central to interpersonal and commercial relationships is the concept of trust (Golembiewski & McConkie, 1975; Soule, 1998), which becomes more important in conditions of risk, uncertainty, and vulnerability (Albrecht, 2002). Hardin (2001) defines trust as a dyadic concept focused on three specific and related actions: *"A trusts B to do X."* This reflects an expectancy that the promise of an organisation can be relied upon (Rotter, 1971). When applied to the digital economy, digital trust has been defined in numerous ways (Akram & Ko, 2014) and the measure of consumer, partner, and employee confidence in an organisation's ability to protect and secure data and the privacy of individuals. It is the expectancy that organisations will collect,

store, process, and use personal information of digital citizens in a manner that benefits and protects those to whom the information pertains.

The 'conditions' that require digital trust involve risk, uncertainty, and vulnerability, as distinct from digital trust itself, defined as digital trustworthiness (adapted from Gefen, 2002). Digital trustworthiness refers to the perceptions about the motives, intentions, and qualities of an influential actor within the digital economy. Building on the work of Mayer, Davis, and Schoorman (1995) and Mayer and Davis (1999), digital trustworthiness comprises three core characteristics, namely (1) Digital Competence, (2) Digital Benevolence, and (3) Digital Integrity. In the digital economy landscape, digital competence reflects a manager's digital expertness, knowledge, influence, impact, and ability in an organisation, digital benevolence reflects a manager's goodwill to an organisation's stakeholders (including consumers), and digital integrity refers to a manager's adherence to principles and values that an organisation's stakeholders find acceptable. While all three components of trustworthiness are important, it is argued (Butler Jr, 1991) that [digital] integrity is the component most important to trust. Using digital data inappropriately jeopardises digital integrity, and hence trust in the organisation.

Managers need to create and maintain the perception and reality of digital ethics and hence the trust by clients, customers, employees, suppliers, government regulators, and other stakeholders (Soule, 1998). Digital trust may be created or destroyed by every digital interaction. Managers should regularly measure and quantify stakeholder expectations to ensure that reality and perceptions of trust are accurate. An organisation may be perceived like a "real" person and stakeholders need to be reassured about the claims, image, and the anticipated and expected behaviour of the organisation. The monetary damage to an organisation from losing digital trust can be enormous. With digital ethics comes the added variable of investigating the ethical implications of things which may not yet exist, or things which may have unintended or unexpected impacts that we have yet to predict. Developing guidelines and codes of conduct and embedding ethical practices for decision making is a precursor and prerequisite to building trust. The literature suggests that making ethical decisions about digital issues leads to perceptions of digital ethical behaviour by customers and stakeholders and hence digital trust and organisation success.

### **3.0 Examples of Digital Ethics Scenarios**

Big data analytics and data mining are relatively new capabilities, and these IT trends have attracted attention from an ethics and a privacy perspective. Managers of businesses in a capitalist market system need to especially address ethical issues associated with these emerging technologies (Steinberg, 2020). Even the appearance of unethical use or abuse of data can permanently harm a business. Managers and IT professionals must respect people's privacy and maintain their trust.

Advances in Big Data Analytics, Data Mining, AI, and ML have created new ethical dilemmas that no one had previously conceived. The use of these emerging technologies has potential ethical implications in a range of areas including healthcare, financial services, national elections, and consumer applications. The following are some examples of recent trends with digital ethics implications:

- The introduction of ChatGPT has fast-tracked the use of generative AI within education and businesses globally. Yet, reports indicate that some information produced by ChatGPT is incorrect, albeit while sounding credible (a term commonly referred to as 'ChatGPT Hallucination,' Den & Lin, 2023), and presents data based on the libraries they were trained upon and hence can become outdated quickly and do not have original new thoughts (Ali & Djalilian, 2023). This is ethically problematic as there exists the risk of bias based on training datasets and plagiarism, and the lack of transparency regarding response/content generation.
- In recent years, direct-to-consumer DNA analysis services have attracted significant attention. Their core service traditionally involves providing customers information about their genetic information, specifically their ancestry, using highly sophisticated genomic data processing and analysis enabled by big data analytics. Some reports indicate that these organisations are undertaking some tests without scientific validity. Subsequently, the information they provide directly to customers may be erroneous. Some of these businesses have started to promote customised precision healthcare products and services to customers e.g., diets to promote youth and lifespan. This is highly problematic as a number of these organisations are not

endorsed to offer this service. In addition, they do not have the adequate professional support services in place to provide customers with potentially life altering advice that may directly impact their personal health.

- Cambridge Analytica is a recent high-profile example of unethical data mining of social media data to target and manipulate individual citizens with tailored election campaign advertisements and messages. This scandal has required marketers, businesses, and governments to rethink how new sources of data, i.e., social media data, are accessed, processed, and used.

- In July 2019, a Russian owned company FaceApp launched and within a short period 150 million people used the app to realistically age their face to understand what they would look like in a couple of decades' time (Cashin, 2019). The US Government have prompted the FBI to launch an investigation of the company who claim they own the images uploaded to their app.

- By summer 2020, more than one million people had downloaded the Irish public health service Covid Tracker app (<https://covidtracker.gov.ie/>). This app requests users to voluntarily provide information such as age, sex, and locality. The app does not ask for specific user data. Users are asked if they want to supply a phone number so the HSE can contact them, this is optional. The majority of the information processed through the app requires user consent and these settings can be changed.

Some other important ethics challenges are associated with behavioural biometrics, datafication of children, and Toronto Sidewalk Labs to create a hyper-efficient city, cf., Baron (2018). One of the fundamental issues with these examples is understanding the boundary of what is both legal and what is also ethical. We closely examine these examples; we are likely to find that what is occurring is legal under current laws and not uncommon. However, the question is whether it is ethical to offer a service that claims to full, unrestricted ownership of personal user data. For some it has become the norm, for users to “*just click agree*” and to trust that organisations will protect their digital identity.

#### **4.0 Identifying Ethics Dilemmas in the Digital Economy**

A tension exists between technological progress and power, and the ethical implications of that progress. Wicked problems and ethical dilemmas are often

difficult to identify and recognise. We briefly outline some current digital ethics dilemmas and “*wicked*” problems facing IT professional and managers:

### **Dilemma #1: Data Ownership**

While policies, legislation, and regulations have been adopted, e.g., data protection by design and default (previously known as Privacy by Design), as part of the European Union GDPR, the question of data ownership remains grey. Personal data ownership is a key consideration of the GDPR, in terms of who owns the sensitive personal data, e.g., financial and health data, which is captured and collected (Romanou, 2018). The troubling question for an IT manager is who owns the personal data once a company has invested in its collection, analysis, and interrogation? Data ownership is typically considered in terms of the data subject - the person to whom the personal data relates and the data controller – the person/entity that captures and processes the personal data.

While GDPR does not explicitly define who owns the data, the regulation empowers the data subject to exercise more control over their personal data. To be compliant, it is essential that business managers understand the legal, statutory, and ethical frameworks in place within the markets they operate (Luckner, 2015; Steinberg, 2020). This understanding should inform actionable and reasonable policies and practices within the organisation. With increasing, changing legislation, it is important that business managers develop capabilities enabling them to be flexible enough to adapt within the changing regulatory data landscape.

**Wicked Problem #1:** The more data an organisation owns the more data is available for sale. Managers are financially motivated to claim ownership of as much data as possible. Exploring this dilemma through an ethical lens, deontological ethics would emphasise the value of respecting individual rights and obtaining appropriate informed consent. Consequentialist ethics would focus on evaluating the long-term consequences of widespread data ownership, considering potential harm to individuals and society. Virtue ethics would encourage cultivating virtues such as honesty and fairness in the decision-making process, questioning the moral character of individuals driven solely by financial motivations.

### **Dilemma #2: Security of Personal Data**

The decisions made by humans related to IT, data capture, and data analyses may involve ethical dilemmas. Personal data is collected, stored, processed, and shared in a variety of ways. Personal data may be shared voluntarily through the completion of online purchases or by creating a personal profile on a social network. It may also be observed based on the data that is available about the person e.g., location-based data, click-stream or browsing preferences or it may be inferred by analysing existing accessible data available about the person (Al-Khouri, 2012). Inferential data includes personal credit ratings, the prediction of future consumption etc. (Al-Khouri, 2012). With many questions being asked of data mining techniques, big data analytics, and AI, there is an increasing responsibility for business managers to develop operational ethical frameworks to ensure permissible collection, use, and storage of customer data (Rieger, 2018; Steinberg, 2020). This approach may involve providing customers/clients/patients with more information about the algorithms and technologies used, as well as creating greater transparency about how their data is managed by organisations in the future.

**Wicked problem #2:** Providing security for data has a cost. Managers are motivated to minimise data security costs even though a data breach may hurt digital trust. Considering this dilemma from an ethics perspective, deontological ethics would stress the obligation to protect individual rights and privacy, suggesting that senior decision makers have a moral duty to invest in data security. Consequentialist ethics would highlight the potential negative consequences of a data breach on digital trust and overall well-being, indicating that prioritising security could lead to better long-term positive outcomes for individuals and organisations. Virtue ethics would encourage managers to embody virtues such as responsibility and integrity, recognising that ethical decisions involve finding a virtuous balance between financial considerations and the well-being of stakeholders.

### **Dilemma #3: Transparency around Personal Data Management**

*“Transparency implies that any information and communication concerning the processing of personal data must be easily accessible and easy to understand”* (“Six data protection principles - FutureLearn”) Terms and Conditions (T&Cs) and Privacy Policies (PPs) documents outline how users’ data is collected, stored, used, and shared. Mostly, these documents are long and technical. Inaccessibility of this valuable information means that users often *“Just Tick Agree”* to register (Rowan,

O'Connor, Lynch, & Heavin, 2017). This information is often perceived by users as a barrier to accessing “an app.” Research has found that users just respond agree because they have a low level of awareness and limited interest in the detail of T&C’s and PP statements (Rowan et al., 2017). Users are motivated to choose current consumption (with benefits ‘free of charge’) over future consumption (based on the long-term value of personal data).

Together with the known problem of “*time preference*”, the omission of ownership results in asymmetry in the digital market economy (Udo, 2018). In delivering new digital products and services to clients, business managers are responsible for ensuring that clear and accessible information, e.g., Terms and Conditions documents, about how client data is shared is available to all users. While users may be happy to “*just tick agree*”, organisations should provide “*the right type of information to the right type of customer at the right time*” to promote user awareness.

Using data for AI development is not ethically neutral. Developing the software is contingent on the training dataset and the bias of the AI developer(s). As the decisions being made by AIs are becoming more important e.g., access to healthcare treatment, determinant of incarceration, access to financial resources the lack of transparency and visibility of the AIs predictions or recommendation is becoming increasingly important for legal, health, financial reasons. Many AIs are perceived as a “*black box*”, yet they can be highly influential in terms of the data they provide to support decision making.

**Wicked problem #3:** Transparency is an ideal, but it creates bureaucracy and legal situations that slow data collection and use. Also, a disclaimer will often be ignored. Ethics provides a lens to further interrogate this dilemma, where deontological ethics would stress the moral duty to be transparent, even in the face of bureaucratic challenges and potential disregard of disclaimers. Consequentialist ethics would focus on weighing the potential positive outcomes of transparency (such as trust-building) against the negative consequences of bureaucracy and legal issues. Virtue ethics would encourage managers to embody virtues such as honesty and responsibility, striving to find a balance that upholds transparency while addressing practical challenges in a virtuous manner.

#### **Dilemma #4: Explicit Data Monetisation Business Model**

Many business managers see their enterprise's data as an asset and a source of genuine business value. Due to the increase in technology use by both businesses and consumers, Udo (2018) argues that the value creation of businesses has "*reversed*" in recent years, with customers being offered services "*free of charge*" albeit paid for by advertisements of businesses to gain access to aggregated data output from users. In his Forbes article, Goodson (2012) explores Google's data collection strategy and use of personal user data. He coined the phrase "*If you are not paying for it, you are the product*" (Goodson, 2012).

More recently in the Cambridge Analytica case, personal data of the friends of consenting users were collected. It is estimated that Cambridge Analytica accessed the personal profiles of approximately 50 million Facebook users while only 270,000 users provided consent for their data (Cadwalladr & Graham-Harrison, 2018; Granville, 2018). This data was collected, analysed and used to personalise and directly target messages at specific types of users. This international scandal has highlighted the capabilities of data mining large volumes of unstructured data to generate insights and target users. The legality and the ethics of these data monetization practices have been called to question. This case has reinforced the notion "*If there is no product, you are the product.*" Personal data is being collected, analyzed, and shared with third party organisations who pay to access and analyse valuable data which may be used to create competitive advantage. This dilemma is closely linked to Dilemma 3, many organisations are not upfront and explicit about their business model i.e., the monetisation of the data they collect on their digital platforms.

**Wicked problem #4:** In a market economy there is an explicit profit motive. Managers do not want competitors to understand the business model. Greed is also sometimes a factor. When we explore this dilemma through an ethical lens, deontological ethics would highlight the importance of conducting business in an honest and fair manner, cautioning against unethical practices even in the pursuit of financial gain. Consequentialist ethics would evaluate the overall consequences of actions, considering the potential negative impacts on competition, stakeholders, and the market as a whole. Virtue ethics would support managers to embody virtues such as integrity and responsibility, advocating for a balance between the profit motive and ethical business practices.



### **Dilemma #5: Identification of personal data**

The nature of much of the data being collected and stored for analysis is sensitive and valuable. Many organisations are interested in user's personal preferences, their financial data, and data about their personal health and well-being. There is data protection and privacy legislation in place across jurisdictions, i.e., in the United States (data privacy laws like COPPA and HIPPA) and Europe (GDPR), but it is fragmentary. The implementation of the GDPR has prompted much discussion and debate around protecting the personal identify of the data subject as part of the use of big data with algorithms and analytics. Experience suggests however that it is difficult to completely anonymise data and protect the privacy of an individual. Also, some organisations engage in the practice of matching pseudonymised data with publicly available information as means on identifying individuals and households. This activity is known as data re-identification or de-anonymisation.

A number of strategies are proposed to protect identities including a data de-identification procedure called pseudonymisation, that is “[t]he process of distinguishing individuals in a dataset by using a unique identifier which does not reveal their ‘real world’ identity” and “anonymous information - data where no information relating to or identifying any individual is shown” (Stalla-Bourdillon & Knight, 2016). However, the interpretation of the legislation is ambiguous and there is little agreement on the optimum approach to de-identifying personal data.

**Wicked problem #5:** Identified data often has a higher value than de-identified. The overriding goal for most business managers is profit maximisation. Deontological ethics would stress the importance of respecting individuals' rights and privacy, even in the face of the higher value of identified data. Consequentialist ethics would focus on evaluating the overall consequences of prioritising identified data, considering potential negative impacts on individuals and society. Virtue ethics would encourage managers to embody virtues such as integrity and responsibility, advocating for a balanced approach that aligns profit maximisation with ethical principles promoting the opportunity to achieve on both fronts.

## **5.0 Code of Conduct for IT Professionals**

By the very nature of their jobs and roles, IT professionals have access to sensitive valuable data about customers, patients, and citizens. They are custodians of powerful

tools and data. IT professionals may not even be aware of when they are acting unethically. In some situations, it is plausible that ethics, values, and morals may be a secondary concern and that employees feel compelled or required to discharge their roles. To stay relevant as IT changes, ethical codes and issues of ethical practice must be reviewed periodically.

Digital ethics and “*Data protection by design and default*” (previously referred to as “Privacy by Design”) should be a guide and focus for data scientists, data analysts, decision support analysts, managers, and IS/T professionals. Those charged with designing technological frameworks and solutions to counter the threats of undesirable, unethical attacks on privacy and trust violations must be diligent and vigilant. Digital ethics is about managing oneself and acting ethically and professionally using online and digital communications and digital storage mediums. Veliz (2019) notes digital ethics is much more political than areas (such as medicine) as it is largely dominated by private forces. For these reasons, she asserts the area of digital ethics requires its own ethical practices. She recommends three steps: 1) the development of ethical codes, 2) the use of ethics committees, and 3) respect for personal autonomy. Veliz asserts two fundamental principles, “*Our data should never be used against us, and a person’s welfare should take precedence over economic interests.*”

Ethical oversight and thoughtful policies must maintain an appropriate balance of interests. Data and analytics have the potential to impact people both positively and negatively. The dilemmas and trade-offs must be resolved for the overall benefits of society in general without harming individual rights. Data analysis of merged and combined data sets create both opportunities and serious ethical challenges. Managers should strive to build a culture that values ethical behaviour. One hopes that in such a culture that win-win solutions can be found like removing identifying particulars, and hence anonymizing data, that will promote the knowledge discovery opportunities of big data analytics and data mining technologies. Building ethical considerations into analytics, data mining, and digital business is much better and easier than trying to add new policies, processes, and system updates later and as an afterthought.

Five strategies should be considered when evaluating the dilemmas and wicked problems associated with making ethical decisions associated with all aspects of digital data capture, storage, and use.

1. *Awareness and Education*: Promote an actionable understanding of digital ethics among business managers in organisations. With changing regulation and heightened tension around GDPR, organisations need to prioritize digital ethics to ensure “*that business interests do not take precedence over the public’s well-being*” (Dimitriadis, 2019). This will involve developing and implementing educational material on digital ethics for information systems, computer science, data science, and business professionals. Business managers should take responsibility to promote a new professional code of conduct including digital ethics.

2. *Understandable, actionable regulation (e.g., GDPR) sanctions need to exist and be implemented*: interpreting existing legislation is a challenge for everyone. Supported by the organisation’s legal expertise, business managers need to understand the legislation and develop and implement policies and processes that adhere to existing legislation.

3. *Regional Autonomy*: context and culture impact on ethics. Business managers need to consider that “*what is considered ethical can vary across individuals, groups, religions and cultures, and in a global and fast-moving digital society, these leave considerable room for interpretation*” (Yardley, 2018). There is no “*one size fits all*” approach to digital ethics. Different legislative frameworks are in operation across geographic jurisdictions which poses significant challenges for businesses. Managers should promote the design and development of flexible and adaptable solutions that promote trust among customers, support data management, and meet the local legislative requirements.

4. *Regulating the profit motive to collect and sell data*: Large technology organisations are integral to the recent commoditisation of data and the global market for data continues to grow. In the US, there is a move towards breaking down these technology giants to reduce their power and influence in international markets and reduce their capacity to collect, store, and analyse data in such a linked manner (some of these organisations own many tech platforms that allow for data linkage across individual profiles). To date, some of the large tech firms have paid significant

financial penalties for GDPR breaches. There have been some recent discussions about implementing GDPR in the US.

5. *Promote Innovation on digital ethics using emerging technologies:* Emerging technologies such as AI, Big Data Analytics, and data mining are used to capture, store, analysis, and share new data insights. These technologies could be leveraged to be part of the solution in several ways. Organisations should use these technologies to support users to better understand what is happening to their data, how their data is processed, where their data is stored, and how their data is being used. Specifically, AI should be used to present a more tailored approach to T&Cs, privacy policies, and electronic consent. Thus, providing users with richer information about their data and better support them to make more informed decisions about the provision of personal data. Further, innovative technology solutions could be developed with ethical principles embedded in their design and development. This approach mandates an “ethics first” approach early in the analysis and design process.

## **6.0 Conclusion**

In their recent paper Gal, Hansen, & Lee (2022) assert that the IS field has a long history of research on ethics, however much of this work either applies no formal ethical theory or it is frequently hidden. This research begins to tackle the intersection of digital ethics and dilemmas in IS through the lens of normative ethics perspectives, presenting a rich area for exploration and development. Addressing these challenges and proposing strategies supported by theoretical and empirical grounding can significantly contribute to IS.

Digital data is a valuable business asset. Managers buy and sell data. Decisions that governments make, customer trends in response to privacy concerns, and the new ethical norms will create an environment in which firms will either flourish or go out of business (Hopkins, 2019). Digital transformation has resulted in a decision environment for all managers that involves making choices about using, storing, and collecting data. Each manager must learn to recognize and resolve multiple, ethically wicked, novel data-related problems. The problems may occur only sporadically, but the choices will be important. Assessing the ethical implications of current and potential data related issues is now part of the job of every manager. Organisational

decision-making environments have fundamentally change. Digital ethics must be part of an organisation's digital transformation strategy to be embedded in a meaningful way in terms decision making around new technologies, new organisational strategies, and new products. Davenport (2020) and Veliz (2019) suggest that organisations should establish a "*data ethics board*" to discuss the ethical implications of new data and analytics-related initiatives. Ethical dilemmas and issues arise during the entire analytics and decision-making process, from what data to use, to how to represent the extracted knowledge and exploit the insights to create economic and social value. Ethical concerns like illegitimate or inappropriate surveillance, invasion of privacy, job elimination and unemployment, malicious use of data and other problems associated with AI and analytics create a real threat to human rights and perhaps even a danger to civilisation as we have known it.

Digital data is proliferating, and the ethical implications remain hard to resolve. Digital ethics is a complex philosophical topic. There is no panacea for answering digital ethics questions. Diverse people, data, and technology operating within the digital space make it difficult to define the boundary of the problem. By combining strong theoretical foundations and empirical evidence, the IS field can foster a culture of digital ethics, where ethical considerations are woven into the fabric of technology design, development, and decision-making. This approach not only mitigates risks associated with ethical dilemmas but also contributes to the creation of more responsible, sustainable, and socially beneficial systems. Case studies offer an opportunity to gather rich qualitative data from real-world applications. Examining in-depth case studies illustrating how various organisations across diverse contexts and industries have navigated specific digital ethics dilemmas, such as contrasting strategies employed by technology companies, healthcare institutions, and educational establishments, would reveal industry-specific challenges and possible solutions.

This review has raised unresolved issues. We have tried to summarise opinions about complex ethical dilemmas. In our opinion, due to the complexity of the ethical issues, future commentary and research should focus upon seeking win-win solutions for all parties, defining the scope of these problems, and assessing the potential harms that are occurring or may result from digital monetisation and broad data sharing. For managers, it is not only imperative, but mandatory, to ensure that the privacy of

individuals is protected, while using data to create value and improve operations. Also, political leaders need to implement actionable legislation that is well understood by governments, organisations, and individuals.

## ACKNOWLEDGEMENT

Removed for review purposes.

## References

Akram, R., & Ko, R. 2014. Digital Trust - Trusted Computing and Beyond a Position Paper, Retrieved from <https://www.jamovi.org/https://pure.royalholloway.ac.uk/portal/files/23295482/Paper.pdf>.

Alexander, L. & Moore, M. 2021. "Deontological Ethics", *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2021/entries/ethics-deontological/>>.

Albrecht, S. L. 2002. Perceptions of integrity, competence, and trust in senior management as determinants of cynicism toward change. *Public Administration and Management: An Interactive Journal*, 7(4), pp. 320-343.

Al-Khouri, A. M. 2012. Data ownership: who owns “my data.” *International Journal of Management & Information Technology*, 2(1), pp. 1-8.

Ali, M.J. and Djalilian, A., 2023, March. Readership awareness series—paper 4: Chatbots and chatgpt-ethical considerations in scientific publications. In *Seminars in ophthalmology* (pp. 1-2). Taylor & Francis.

Austin Center for Design (AC4D). 2018. Wicked Problems, Retrieved from <http://www.ac4d.com/our-work-philosophy-and-approach-to-education/understanding-wicked-problems/>.

Baldini, G., Botterman, M., Neisse, R., & Tallacchini, M. 2018. Ethical design in the internet of things. *Science and engineering ethics*, 24(3), pp. 905-925.

Bentham, J., 1843. *Rationale of Reward*, Book 3, Chapter 1, in *The Works of Jeremy Bentham*, J. Bowring (ed.), Edinburgh: William Tait.

Bélanger, F., & Crossler, R. E. 2011. Privacy in the digital age: a review of information privacy research in information systems. *MIS Quarterly*, 35(4), pp. 1017-1042.

- Bowie, N. E., & Jamal, K. 2006. Privacy rights on the internet: self-regulation or government regulation? *Business Ethics Quarterly*, 16(3), pp. 323-342.
- Butler Jr, J. K. 1991. Toward understanding and measuring conditions of trust: Evolution of a conditions of trust inventory. *Journal of Management*, 17(3), pp. 643-663.
- Buytendijk, F., Vashisth, S., Duncan, A., & Moran, M. 2016. Kick-Start the Conversation on Digital Ethics, Retrieved from <https://www.gartner.com/en/documents/3296619/kick-start-the-conversation-on-digital-ethics-2016>.
- Cadwalladr, C., & Graham-Harrison, E. 2018. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*, Retrieved from <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
- Cashin, R. 2019. FBI requested to investigate FaceApp over fears of data ownership, Retrieved from <https://www.joe.ie/tech/fbi-requested-investigate-faceapp-fears-data-ownership-675670>.
- Consumers International. 2018. The state of data protection rules around the world: A briefing for consumer organizations, Retrieved from <https://www.consumersinternational.org/media/155133/gdpr-briefing.pdf>.
- Côrte-Real, N., Ruivo, P., & Oliveira, T. 2020. Leveraging internet of things and big data analytics initiatives in European and American firms: Is data quality a way to extract business value? *Information & Management*, 57(1), p. 103141.
- Deng, J. and Lin, Y., 2022. The Benefits and Challenges of ChatGPT: An Overview. *Frontiers in Computing and Intelligent Systems*, 2(2), pp.81-83.
- Dimitriadis, C. 2019. Digital ethics rising in importance, Retrieved from <https://www.csoonline.com/article/3396037/digital-ethics-rising-in-importance.html>.
- Fleckenstein, M., & Fellows, L. 2018. Valuing Data as an Asset, *Modern Data Strategy*: 11-14, Springer.
- Gal, U., Hansen, S., & Lee, A. S. 2022. Research Perspectives: Toward Theoretical Rigor in Ethical Analysis: The Case of Algorithmic Decision-Making Systems. *Journal of the Association for Information Systems*, 23(6), 1634-1661.
- Gefen, D. 2002. Reflections on the dimensions of trust and trustworthiness among online consumers. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 33(3), pp. 38-53.

George, A.S. and George, A.H., 2023. A review of ChatGPT AI's impact on several business sectors. *Partners Universal International Innovation Journal*, 1(1), pp.9-23.

Gharehgozli, A. H., Mileski, J., Adams, A., & von Zharen, W. 2017. Evaluating a “wicked problem”: A conceptual framework on seaport resiliency in the event of weather disruptions. *Technological Forecasting and Social Change*, 121, pp. 65-75.

Golembiewski, R. T., & McConkie, M. 1975. The centrality of interpersonal trust in group processes. *Theories of Group Processes*, pp. 131-185.

Goodson, S. 2012. If You're Not Paying for It, You Become the Product, Retrieved from <https://www.forbes.com/sites/marketshare/2012/03/05/if-youre-not-paying-for-it-you-become-the-product/#2f702eac5d6e>.

Granville, K. 2018. Facebook and Cambridge Analytica: What you need to know as the fallout widens, Retrieved from <https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html>.

Graves, H., & Garrett Jr, R. K. 2014. A template for solving wicked problems. Paper presented at the INCOSE International Symposium.

Hardin, R. 2001. Conceptions and explanations of trust. In K. S. Cook (Ed.), *Russell Sage foundation series on trust*, Vol. 2. *Trust in society*, 3-39. Russell Sage Foundation.

Hursthouse, R. & Pettigrove, G. 2023. "Virtue Ethics", *The Stanford Encyclopedia of Philosophy* (Fall 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/fall2023/entries/ethics-virtue/>.

Hopkins, B. 2019. Facebook's recent moves highlight the grand challenge of digital ethics, Retrieved from <https://www.zdnet.com/article/facebooks-recent-moves-highlight-the-grand-challenge-of-digital-ethics/>.

Korzynski, P., Mazurek, G., Altmann, A., Ejdys, J., Kazlauskaite, R., Paliszkiwicz, J., Wach, K. and Ziemba, E., 2023. Generative artificial intelligence as a new context for management theories: analysis of ChatGPT. *Central European Management Journal*.

Lewis, P. V. (1985). Defining ‘business ethics’: Like nailing jello to a wall. *Journal of Business ethics*, 4(5), 377–383.

Lucker, J. 2015. The Dangers of Monetizing Data, Retrieved from <https://deloitte.wsj.com/cio/2015/06/30/the-dangers-of-monetizing-data/>.



- Martin, K. 2020. Breaking the Privacy Paradox: The Value of Privacy and Associated Duty of Firms. *Business Ethics Quarterly*, 30(1), pp. 65-96. DOI: <https://doi.org/10.1017/beq.2019.24>
- Mayer, R. C., & Davis, J. H. 1999. The effect of the performance appraisal system on trust for management: A field quasi-experiment. *Journal of Applied Psychology*, 84(1), pp. 123.
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. 1995. An integrative model of organizational trust. *Academy of Management Review*, 20(3), pp. 709-734.
- Mill, J. S., 1861. *Utilitarianism*, edited with an introduction by Roger Crisp. New York: Oxford University Press, 1998.
- Mitroff, R., & Mason, I. 1973. A program for research on management information systems. *Management Science*, 19(5), pp. 475-591.
- Mohamed, S., & Mohamed, D. 2020. Data in the Internet of Things Era: The Propertization of Data in Light of Contemporary Business Practices. *International Journal of Business and Society*, 21(S1), pp. 81-94.
- Price, M. S. 2020. Internet privacy, technology, and personal information. *Ethics and Information Technology*, pp. 1-11.
- Railton, P. (1986). Moral realism. *The philosophical review*, 95(2), 163-207.
- Rieger, J. 2018. Data Monetization Step Three: Ethical use of data and building trust. In L. A. J. f. U. <https://liveramp.com/blog/data-monetization-step-3-building-trust/>. (Ed.).
- Ritter, T., & Pedersen, C. L. 2020. Digitization capability and the digitalization of business models in business-to-business firms: Past, present, and future. *Industrial Marketing Management*, 86, pp. 180-190.
- Romanou, A. 2018. The necessity of the implementation of Privacy by Design in sectors where data protection concerns arise. *Computer law & security review*, 34(1), pp. 99-110. DOI: <https://doi.org/10.1016/j.clsr.2017.05.021>
- Rotter, J. B. 1971. Generalized expectancies for interpersonal trust. *American Psychologist*, 26(5), pp. 443-452.
- Rowan, W., O'Connor, Y., Lynch, L., & Heavin, C. 2017. Exploring user behaviours when providing electronic consent on health social networks: A 'just tick agree' approach. *Procedia Computer Science*, 121, pp. 968-975.
- Sarathy, R., & Robertson, C. J. 2003. Strategic and ethical considerations in managing digital privacy. *Journal of Business Ethics*, 46(2), pp. 111-126.

- Sidgwick, H., 1907. *The Methods of Ethics*, seventh edition, London: Macmillan; first edition, 1874.
- Soule, E. 1998. Trust and managerial responsibility. *Business Ethics Quarterly*, 8(2), pp. 249-272.
- Spinello, R. A. 1998. Privacy rights in the information economy. *Business Ethics Quarterly*, 8(4), pp. 723-742.
- Stahl, B. C. 2013. Responsible research and innovation: The role of privacy in an emerging framework. *Science and Public Policy*, 40(6), pp. 708-716. DOI: <https://doi.org/10.1093/scipol/sct067>
- Stalla-Bourdillon, S., & Knight, A. 2016. Anonymous Data v. Personal Data-False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data. *Wisconsin International Law Journal*, 34(2), pp. 284-322.
- Steinberg, E. 2020. Big Data and Personalized Pricing. *Business Ethics Quarterly*, 30(1), pp. 97-117.
- Swinhoe, D. 2018. What is digital trust? How CSOs can help drive business, Retrieved from <https://www.csoonline.com/article/3297037/what-is-digital-trust-how-csos-can-help-drive-business.html>.
- Udo, M. 2018. The GDPR: Halfway between consumer protection and data ownership rights, Retrieved from <https://www.ingentaconnect.com/content/hsp/jdb001/2018/00000003/00000001/art00002>.
- Yardley, D. 2018. Are You Making Ethical Decisions During the Digital Transformation Process? Retrieved from <https://www.koganpage.com/article/essential-ethics-for-digital-transformation>
- Yun, H., Lee, G., & Kim, D. J. 2019. A chronological review of empirical research on personal information privacy concerns: An analysis of contexts and research constructs. *Information & Management*, 56(4), pp. 570-601.

# Online Health Communities for Parents of Children with Autism Spectrum Disorder

*Research In progress*

## **Abstract**

*Online health communities (OHCs) play a crucial role as support hubs for patients and caregivers grappling with various diseases. Despite widespread research attention on OHCs, those designed specifically for parents of children with Autism Spectrum Disorder (ASD) have surprisingly been overlooked. This research aims to address this gap by delving into the emotional expressions and information-seeking behaviours exhibited by parents within these OHCs. Employing a Netnography approach complemented by Natural Language Processing (NLP) techniques, we undertake a case study focused on an OHC related to ASD. Preliminary results indicate that predominant emotional expressions revolve around stress, anxiety, and depression. Upon completion, this research seeks to illuminate critical aspects of OHCs catering to parents of children with autism, underscoring the significance of specialised social support systems.*

**Keywords:** Online Health Community, Autism Spectrum Disorder, Information Needs, Emotional Support, Netnography

## **1 Introduction**

Online health communities (OHCs) have emerged as indispensable platforms facilitating the acquisition of health information. These communities foster interactions between healthcare professionals, medical experts, patients, and caregivers, encompassing various topics ranging from physical ailments to mental health concerns. Characterised by their open membership and anonymous environment, OHCs have assumed a crucial role in providing timely information and emotional support, positioning themselves as supplementary services complementing formal health treatments and practices (White & Dorman, 2001).

Recognising the potentialities offered by OHCs, previous studies have explored their utilisation among individuals afflicted with diseases such as cancers and diabetes. However, there exists a notable dearth of research focusing on parents and caregivers of children diagnosed with Autism Spectrum Disorder (ASD) (Mohd Roffeei et al., 2015). ASD, a neurodevelopmental disorder, exerts profound effects on an individual's social and communicative capacities from early childhood onward. Medical reports posit that early intervention significantly benefits the developmental trajectory of children, aiding in overcoming social disabilities. This hence necessitates the undivided attention, substantial effort, and continual care from parents and caregivers. The distinctive nature of ASD and the considerable stress experienced by parents underscore the imperative for heightened ASD-specific support mechanisms. OHCs that integrate emotional and informational support emerge as of paramount importance for parents navigating the challenges associated with raising children with ASD.

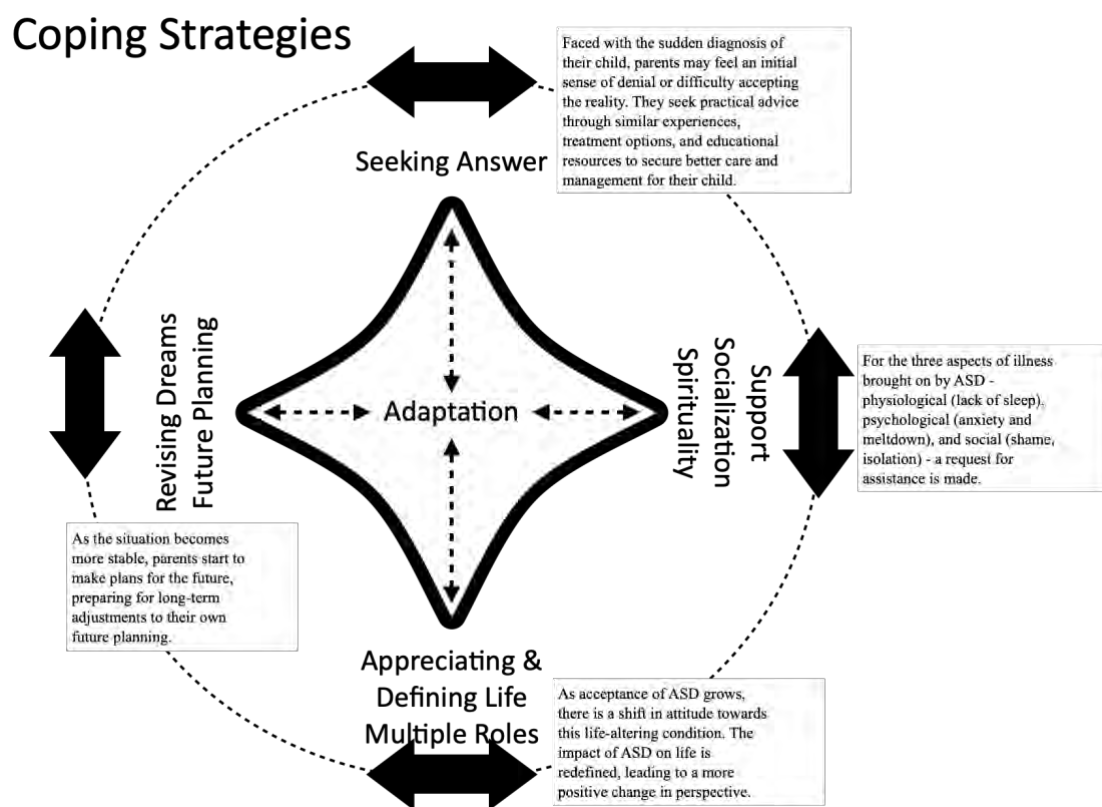
Recognising the evident potential of OHCs in meeting the unique needs of children with ASD and their parents, we believe that it is important to delve deeper into the Research Question (RQ): How do parents of children with ASD seek information and emotion support within OHC? We will employ the Adaptation Journey (Lutz et al., 2012) as the theoretical framing to observe the long-term emotional experiences and transitions over stages. By exploring this RQ, we aim to illuminate how these communities fulfil the specific emotional and informational needs of the parents and illuminate the emotional journey and challenges under different stages. In this research, we take netnography approach (Kozinets, 1998) to investigate user-centric interpretation and interaction among ASD parents in dealing with the disease and challenging events. We contend that this exploration will reveal insights into optimising these communities for more targeted and effective support.

## **2 Literature Review**

OHCs have emerged as an important channel for patients and caregivers to seek health information. Facilitated by their reciprocal, anonymous and networked nature, interactions within the communities are conducted in a more equal footing basis, devoid of prejudgments based on socio-economic status (White & Dorman, 2001). The literature also indicates that the emergence of OHC has created significant social value in certain special diseases (Goh, 2016). Existing literature has extensively documented the prevalent use of OHCs by individuals grappling with health concerns such as, diabetes (Hilliard et al., 2015; Litchman et al., 2019), lung cancer (Lobchuk et al., 2015; Walsh & Al Achkar, 2021), breast cancer (Bender et al., 2013; Zhang et al., 2017), Parkinson's disease (Visser et al., 2016), and HIV/AIDS patients (Maestre et al., 2018). In contrast to the considerable attentions afforded to the aforementioned conditions, research focusing on OHCs as they pertain to parents to parents and caregivers of children with ASD is notably limited (Mohd Roffeei et al., 2015).

ASD, a neurodevelopmental disorder, affecting an individual's social, communicative, and expressive abilities, typically manifesting in early childhood and persist throughout one's lifespan (World Health Organization, 2019). Recent studies indicate an increase of ASD, suggesting that individuals with ASD may no longer be considered a minority group (Cardinal et al., 2021; McConkey, 2020). Given the distinctive symptoms associated with ASD, ongoing support from parents and family members becomes imperative, introducing daily challenges for these caregivers. Recent research underscores the significance of providing caregivers with comprehensive ASD-related information and sustained emotional supports as integral components of persisting early interventions and treatments for affected children, as well as essential social supports

for caregivers (Hyman et al., 2020) However, without a clear understanding of the online support, the efficacy of ASD-related OHCs in delivering substantial assistance to those in need remains questionable. A nuanced understanding of caregivers' emotional expression and information-seeking behaviours is crucial for the value of ASD-related OHCs, thereby offering valuable resources for individuals navigating the challenges associated with ASD. Parenting a child with ASD is a continuous journey, marked by various challenges and emotional difficulties at each stage. Using the Adaptation Journey (Lutz et al., 2012), we categorize parenting of children with ASD into four stages based on coping strategies, enhancing our observation of emotional expression characteristics and transitions.



**Figure 1. Four stages referencing the Adaptation Journey**

The recent research also shows that emotional expression and emotional support on OHCs have become a focus in recent years (Zhou et al., 2023), indicating the increasing

importance of emotions in OHCs. In pursuit of our objective, our investigation delves into the ways through which parents of children with ASD express their emotions within OHCs, and how they seek support or information to fortify their emotional well-being. Our theoretical focus is especially directed towards the exploration and categorisation of emotions commonly experienced by these parents: stress, anxiety, depression, loneliness, and shame. These emotions are explained in the following.

- **Stress** is defined as a distinct interaction between an individual and their environment, where the individual perceives the demands as exceeding their capacity to manage, consequently posing a threat to their well-being (Lazarus & Folkman, 1984).
- **Anxiety** is recognised a future-oriented mood state associated with preparation for possible, upcoming negative events (Barlow, 2004).
- **Depression** is extreme sadness or despair that lasts more than days. It interferes with the activities of daily life and can cause physical symptoms such as pain, weight loss or gain, or lack of energy.
- **Loneliness** refers to the emotional state of feeling isolated or lacking companionship, often accompanied by a sense of social disconnection or the perception that one's social needs are not being met (de Jong Gierveld, 1998).
- **Shame** involves a deep sense of humiliation, disgrace, or regret about one's actions, thoughts, or identity (Brown, 2006). The emotion of "shame" is particularly highlighted when discussing parents of children with ASD. Literature reveals that many parents of children with ASD often feel a sense of shame in certain situations (Farrugia, 2009; Gray, 1993, 2002).

Categorising the emotions offers a profound insight of the intricate emotional states of parents, providing an opportunity for us to reveal the mechanisms underlying emotional support within OHCs.

Furthermore, the information shared, sought, and exchanged within OHCs is diverse and complex. Ingwersen (1986) identified three primary types of information needs for

users: Verificative Information Need (VIN), Conscious Topical Information Need (CIN), and Muddled Topical Information Need (MIN).

- **VIN** is that user wants to *verify* information objects with known *non-topical* (structured) data (Ingwersen, 2000, p.164), such as information source, journal name, and facts.
- **CIN** is that the user wants to *clarify, review, or pursue* information in a known subject matter and domain, where known subject matter signifies *topical* (unstructured) data about contents, such as terms, concepts, and image representation (Ingwersen, 2000, p.164).
- **MIN** is engaged in the *exploration* of new concepts and relations outside known subject matter or domain (Ingwersen, 2000, p.164-165).

By identifying VIN, we can determine whether parents are *verifying* certain facts they are already known, such as verifying the source of information or validating non-topical facts. Conversely, recognizing CIN helps us discern if parents are *deepening their understanding* of known topics, while MIN indicates whether they are *exploring* new concepts beyond their current knowledge. The categorised information needs provide a conceptual foundation for our investigation into the information-seeking behaviour of parents in the context of OHC.

## **4 Methodology**

To fulfil our research objective, we selected the *Autism Parenting* OHC on Reddit as our research setting. This particular OHC is notably active, witnessing an average of 20 posts daily on the site, which is one of the richest, openly available, and spans a long-term period, making it one of the few datasets of its kind. The dataset's period starts from April 10, 2023, and extends up to November 11, 2023. We employed the research methodology of Netnography (Kozinets, 1998), involving a thorough investigation of conversations through the reading of posts and comments. Netnography is effective in studying OHCs' culture, values, language, behaviors, and interactions. It captures the



significant self-disclosure of struggles and challenges among members, essential for understanding them, while ensuring non-intrusive observation (Addeo et al., 2019). Given the substantial volume of textual data available, we utilised Python Reddit API Wrapper (PRAW) to retrieve online information. Through PRAW, we successfully collect more than 4,000 posts post from the OHC over the past six months. This approach ensures that our dataset is extensive and recent, allowing us to capture the latest and most relevant information pertaining to discussions and trends related to autism parenting.

Furthermore, to enhance our comprehension of the sentiments conveyed in textual data, we employed Natural Language Processing (NLP) techniques to facilitate our data analysis. The data analysis comprises the following procedures:

- **Data pre-processing:** Initial steps involve the removal of irrelevant posts, such as advertisements and bot-generated content. Subsequently, NLP tools are used for tokenisation, stop word removal, and part-of-speech tagging to emphasise meaningful content.
- **Content analysis:** This phase begins with defining analysis units, such as words, sentences, and paragraphs. For this study, theoretical concepts introduced earlier—emotions and information needs—constitute the central focus of our analysis. Both manual and automatic coding are performed using NLP tools, and a coding guide is developed to facilitate systematic analysis.
- **Statistical evaluation:** to measure the relationships between specific discussion topics and emotions, as well as information-seeking behaviour, a statistical evaluation is undertaken. This approach enables a quantitative assessment of the associations within the dataset.

The research design is primarily qualitative, incorporating analysis techniques such as PRAW and NLP. While it is not uncommon for prior studies to employ such mixed methods, given the alterations in participants and the intricate nature of conversations in this study, the current research design may not be fully matured. To enhance the robustness of this research, we eagerly seek additional suggestions and valuable

feedback from the conference audience. Your insights will greatly contribute to refining the quality of our study.

## **5 Preliminary Results**

The research is currently in progress, and while ongoing, preliminary outcomes incorporating descriptive statistical information have unveiled noteworthy phenomena. Initially, concerning the expression of emotions, stress emerges as the most prevalent emotion, accounting for 25% of emotional expressions, followed by anxiety (15%) and depression (10%). Notably, within the realm of emotional expressions, 17% of posts lack specific information needs, serving solely as a platform for emotions to vent. Moving on to information needs, the VIN category encompasses 7% of messages, Compulsory Information-Seeking CIN constitutes about 42%, and the Miscellaneous Information-Seeking MIN represents approximately 33%.

Moreover, we have observed that numerous posts reflect stress stemming from societal prejudices and judgments. Within this specific OHC, parents frequently seek support in addressing and navigating societal perceptions of ASD, including those from relatives, friends, and strangers. Parents often pose questions about social judgments, such as “how can we help others understand autism?” or inquiries about methods to enhance their children's social skills and societal inclusion. This emerging theme appears to be a potentially significant issue deserving further exploration.

## **6 Conclusions**

The current research is ongoing, yet the progress has been promising. We are optimistic that by the time of the conference, we will have uncovered even more intriguing

findings. We eagerly look forward to sharing these insights with fellow researchers and contributing to the collective body of knowledge.

## References

- Addeo, F., Paoli, A. D., Esposito, M., & Bolcato, M. Y. (2019). Doing social research on online communities: The benefits of netnography. *Athens journal of social sciences*, 7(1), 9-38.
- Barlow, D. H. (2004). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. Guilford Press, New York
- Bender, J. L., Jimenez-Marroquin, M. C., Ferris, L. E., Katz, J., & Jadad, A. R. (2013). Online communities for breast cancer survivors: a review and analysis of their characteristics and levels of use. *Supportive Care in Cancer*, 21, 1253-1263.
- Brown, B. (2006). Shame resilience theory: A grounded theory study on women and shame. *Families in Society*, 87(1), 43-52.
- Cardinal, D. N., Griffiths, A. J., Maupin, Z. D., & Fraumeni-McBride, J. (2021). An investigation of increased rates of autism in US public schools. *Psychology in the Schools*, 58(1), 124-140.
- Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas, A., & Holyst, J. A. (2011). Collective emotions online and their influence on community life. *PLOS ONE*, 6(7), e22207.
- de Jong Gierveld, J. (1998). A review of loneliness: concept and definitions, determinants and consequences. *Reviews in clinical gerontology*, 8(1), 73-80.
- Hilliard, M. E., Sparling, K. M., Hitchcock, J., Oser, T. K., & Hood, K. K. (2015). The emerging diabetes online community. *Current Diabetes Reviews*, 11(4), 261-272.
- Farrugia, D. (2009). Exploring stigma: medical knowledge and the stigmatisation of parents of children diagnosed with autism spectrum disorder. *Sociology of Health & Illness*, 31(7), 1011-1027.
- Goh, J. M., Gao, G. (Gordon), & Agarwal, R. (2016). The Creation of Social Value: Can an Online Health Community Reduce Rural–Urban Health Disparities? *MIS Quarterly*, 40(1), 247-264.
- Gray, D. E. (1993). Perceptions of stigma: The parents of autistic children. *Sociology of Health & Illness*, 15(1), 102-120.
- Gray, D. E. (2002). ‘Everybody just freezes. Everybody is just embarrassed’: Felt and enacted stigma among parents of children with high functioning autism. *Sociology of Health & Illness*, 24(6), 734-749.

- Hyman, S. L., Levy, S. E., & Myers, S. M.; American Academy of Pediatrics, Council on Children With Disabilities and Section on Developmental and Behavioral Pediatrics. (2020). Clinical report: Identification, evaluation, and management of children with autism spectrum disorder. *Pediatrics*, 145(1), e20193447.
- Ingwersen, P. (1986). Cognitive analysis and the role of the intermediary in information retrieval. In R. Davies (Ed.), *Intelligent information systems* (pp. 206–237). Chichester, UK: Horwood.
- Ingwersen, P. (2000). Users in Context. In: Agosti, M., Crestani, F., Pasi, G. (eds) *Lectures on Information Retrieval. ESSIR 2000. Lecture Notes in Computer Science*, vol 1980. Springer, Berlin, Heidelberg.
- Kozinets, R. V. (1998). On netnography: initial reflections on consumer research investigations of cyberculture. *Advances in Consumer Research*, 25(1), 366–71.
- Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer Publishing Company, New York.
- Litchman, M. L., Walker, H. R., Ng, A. H., Wawrzynski, S. E., Oser, S. M., Greenwood, D. A., Gee, P. M., Lackey, M., & Oser, T. K. (2019). State of the science: a scoping review and gap analysis of diabetes online communities. *Journal of diabetes science and technology*, 13(3), 466-492.
- Lobchuk, M., McClement, S., Rigney, M., Copeland, A., & Bayrampour, H. (2015). A qualitative analysis of “naturalistic” conversations in a peer-led online support community for lung cancer. *Cancer nursing*, 38(5), E21-E31.
- Maestre, J. F., Herring, S. C., Min, A., Connelly, C. L., & Shih, P. C. (2018). Where and how to look for help matters: analysis of support exchange in online health communities for people living with HIV. *Information*, 9(10), 259.
- McConkey, R. (2020). The rise in the numbers of pupils identified by schools with autism spectrum disorder (ASD): a comparison of the four countries in the United Kingdom. *Support for learning*, 35(2), 132-143.
- Mohd Roffeei, S. H., Abdullah, N., & Basar, S. K. R. (2015). Seeking social support on Facebook for children with Autism Spectrum Disorders (ASDs). *International Journal of Medical Informatics*, 84(5), 375-385.
- World Health Organization. (2019). *Autism spectrum disorders* (No. WHO-EM/MNH/215/E). World Health Organization. Regional Office for the Eastern Mediterranean.
- Visser, L. M., Bleijenbergh, I. L., Benschop, Y. W., Van Riel, A. C., & Bloem, B. R. (2016). Do online communities change power processes in healthcare? Using case studies to examine the use of online health communities by patients with Parkinson's disease. *BMJ open*, 6(11), e012110.

- Walsh, C. A., & Al Achkar, M. (2021). A qualitative study of online support communities for lung cancer survivors on targeted therapies. *Supportive Care in Cancer*, 29, 4493-4500.
- White, M., & Dorman, S. M. (2001). Receiving social support online: implications for health education. *Health education research*, 16(6), 693-707.
- Zhang, S., Grave, E., Sklar, E., & Elhadad, N. (2017). Longitudinal analysis of discussion topics in an online breast cancer community using convolutional neural networks. *Journal of biomedical informatics*, 69, 1-9.
- Zhou, J., Zhang, Q., Zhou, S., Li, X., & Zhang, X. M. (2023). Unintended emotional effects of online health communities: A text mining-supported empirical study. *MIS Quarterly*.

# Sharing of personal health information for secondary use: A scoping review from the perspective of trust

*Completed Research*

## **Abstract (around 150 words)**

*The adoption and use of health information systems depend on the gathering of personal health information (PHI), and trust is considered a prerequisite for sharing PHI for secondary use. However, the operationalisation of trust in empirical research is often not explicit about the underlying constructs. This article aims to evaluate and enhance the operationalisation of trust in research involving PHI sharing. A scoping literature review was conducted, encompassing 76 articles. The articles' analysis was based on seminal models of trust, particularly the model presented in (Mayer et al., 1995), and data were extracted and synthesised with a concept-centric approach. Constructs such as ability, benevolence, and integrity from the Mayer et al. model are frequently found in PHI contexts. Additionally, constructs specific to PHI research, such as confidentiality, are identified. Based on the review, recommendations for changes to existing conceptualisations and operationalisations of trust are proposed to improve PHI research.*

**Keywords:** Trust, Health Information Systems, Information Sharing, Scoping Literature Review, Confidentiality, Privacy

## **1. Introduction**

Health information systems (HIS) offer the potential to address the challenges facing healthcare. Recent advances in information technologies, such as big data analytics, artificial intelligence, and mHealth offer a plethora of new possibilities for managing healthcare (Cavallone & Palumbo, 2020) and improving the diagnosis and treatment of diseases (Wolf et al., 2019). However, new HIS solutions require the accumulation of personal health information (PHI), a necessity which is thought to be accompanied by a “dark side”, raising concerns about transparency, security, fairness (Mikalef et al., 2022) and privacy (Pool et al., 2020). While those concerns can hinder PHI sharing, trust has been thought to be a major enabler of PHI sharing, as it has been shown to increase users' intention to share their information for Covid-19 contact tracing (Hong & Cho, 2023; Lin et al., 2021), or for participating in a sustainable health system (Hillebrand et al., 2023). Perhaps more tellingly, big data initiatives in healthcare, such as the care.data project in the United Kingdom, are believed to have collapsed under the weight of widespread public distrust (Carter et al., 2015), showing that trust is a necessary foundation of initiatives that involve the collection of PHI. Despite the significance of trust, its operationalisation as a construct has been problematic,

especially in health services research (Taylor et al., 2023). Furthermore, past literature reviews on PHI sharing, although rigorous, have not focused on the ontology and methodological operationalisation of trust.

Understanding trust as a construct, its antecedents and its impact on information disclosure is paramount for successful HIS adoption and use. This paper presents a current assessment of trust's operationalisation in PHI-sharing research, and provides a theoretical framework to support future HIS research around trust. Specifically, we seek to answer the following research questions:

*How has the construct of trust been operationalised in extant, empirical studies involving PHI sharing?*

*How can the conceptualisation and operationalisation of trust be improved when studying PHI sharing?*

To answer these questions, we form an analytical lens, by identifying the relevant antecedents and constructs directly associated with trust, drawing on the work of (Mayer et al., 1995) and relevant PHI-sharing research. Thereafter, we report on the methods, research instruments and results of empirical PHI-sharing studies, by comparing them to our analytical lens. Finally, we synthesise our findings and arrive at better-suited theoretical pathways for measuring trust in the context of PHI-sharing. As a result, this review expands upon the model of (Mayer et al., 1995) by including constructs that were not originally present, while considering adjustments of existing constructs that the context of PHI sharing necessitates. Consequently, our study delivers both theoretical and research insights, by elaborating and expanding on an established ontology of trust, studying its application in the specific context of PHI sharing, and proposing a context-relevant trust model. Our findings can be utilised in HIS research, to improve the conceptualisation and operationalisation of trust in studies concerning PHI sharing.

## **2. Theoretical Background**

### **2.1 Definition of Trust/Original Model**

Trust has been thoroughly studied by scholars specialising in various disciplines (Kim et al., 2008) and is usually conceived as a multidimensional construct. (e.g., McAllister 1995; McKnight et al., 2002). As McEvily and Tortoriello (2011) summarise, trust is thought to operate in three distinct ways: perceptually, as an expectation about another party (trustworthiness beliefs); volitionally, as a willingness to put oneself into a vulnerable position

(trusting intentions); and behaviourally, as a risk-taking act (trusting behaviours). In 1995, Mayer and colleagues presented a model of trust (hereafter the MDS model) that described trust's antecedents and downstream relations with risk-taking actions. Considered seminal in trust literature (Dirks & De Jong, 2022; McEvily & Tortoriello, 2011), the MDS model provides a definition of trust which comprises those perceptual, volitional and behavioural manifestations:

*[trust is] the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party* (Mayer et al., 1995, p. 712).

The MDS model defines trust as one's willingness to be vulnerable, and is shaped by expectations for the trustee. These expectations are derived from a cognitive assessment of the trustee's characteristics and form his perceived trustworthiness (hereafter trustworthiness), which comprises three dimensions: ability, integrity, and benevolence. Most often synonymous with competence (McEvily & Tortoriello, 2011), ability refers to the trustor's notion that the trustee possesses the necessary skills and competencies within a specific domain. Integrity refers to the trustee's morality, promise-keeping and honesty (Mayer et al., 1995). Benevolence is "the extent to which a trustee is believed to want to do good to the trustor, aside from an egocentric profit motive" (Mayer et al., 1995, p. 718). The MDS model contains another factor, propensity to trust (hereafter propensity) which is represented both as a moderator of the relationship between trustworthiness and trust, and as a direct predictor of trust. Propensity is described as a person's general inclination to trust others and it is based on the personality, experiences and culture of the trustor (Mayer et al., 1995; Schoorman et al., 2007).

## **2.2 Additional Antecedents and Underlying Constructs**

Over the years, literature reviews have studied trust both theoretically and through its use in empirical research. Notably, investigating interpersonal trust in organisational settings, Lyu and Ferrin (2018) based their literature review on four categories of determinants: trustor-related factors (incl. propensity to trust), trustee-related factors (incl. benevolence, ability, and integrity), contextual factors (incl. network and culture) and relationship factors (incl. relationship length, communication, and similarities). They found that previous empirical research supports the significant effect of trust's determinants, while pointing out trust's role as a mediator between perceived trustworthiness and risk-taking, as proposed by (Mayer et al.,



1995), and later supported by (Colquitt et al., 2007). The remainder of this sub-section elaborates on constructs that fit into the latter two categories of factors described by (Lyu & Ferrin, 2018), contextualising them in PHI sharing.

#### Privacy and confidentiality

Two of the most widely used antecedents of trust not originally included in the MDS model are confidentiality and privacy concerns, which are thought to have a significant influence on trust, especially in information-sharing scenarios (Wu et al., 2012). This influence is particularly pronounced in healthcare (Sterckx et al., 2016; Platt & Kardia, 2015) and a correlation between perceptions of privacy risks and trust in confidential handling of data has been confirmed empirically (Belfrage et. al., 2022). One of the most influential scales of measuring trust in healthcare is the Wake Forest scale (Platt et al., 2018). Developed by Hall et al. (2001), it introduces the construct of confidentiality, in addition to the three dimensions of trustworthiness. According to its developers, confidentiality refers to a patient's assurance that their private information will be protected, used properly, and will not be used against them.

#### Temporal effects on trust

Although trust can be seen as a state or an attitude, it is not static (Korsgaard, et.al., 2018). Indeed, the MDS model does not present trust as static, but instead proposes an iterative and dynamic development of trust through a feedback loop. A subsequent meta-analysis has pointed out the positive relationship between relationship length and trust (Vanneste et al., 2014). Similar discoveries suggest an influence of past interactions with a party on perceived trustworthiness, as summarised by (Kramer, 1999). Recent theorising regarding the trust-building processes at different stages of a relationship is of particular interest. Ongoing discourse in trust research has proposed heuristics, cognition and affect as bases of trust, as the length of a relationship increases (Baer & Colquitt, 2018; Lind, 2018; van Knippenberg, 2018).

#### Distrust

While literature traditionally defines distrust as the absence of trust and places both on a single spectrum (McKnight & Chervany, 2001; Schoorman et al., 2007), other streams delve deeper into the subtleties of the construct of distrust and point out details that might separate the two in a more meaningful way (Lewicki et al., 1998). Bijlsma-Frankema and colleagues (2015) have indicated two main characteristics of distrust: the presence of negative expectations about a party and a self-amplifying cycle of distrust. Moreover, it has been proposed that distrust can

have a dampening effect on the antecedents of trust (Sitkin & Bijlsma-Frankema, 2018). Still, (Schoorman et al., 2007) hypothesise that studying only one of the two constructs at a time might be the reason why they appear to be different and to exist simultaneously and independently of each other. As (Saunders et al., 2014) find in their empirical, mixed-methods study, the simultaneous coexistence of both trust and distrust towards the same trustee is exceedingly rare. In sub-section 3.2, we explain how we handled distrust in our literature review in light hereof.

### Culture and third parties

As mentioned above, propensity is believed to have roots in both one's personality and experiences, and in their culture (Schoorman et al., 2007). Trust literature has studied the impact of culture, finding direct or indirect effects on trust (Wu et al., 2012) and on ability and benevolence (Schoorman et al., 2007). Apart from being an individual factor, culture can also influence relationships. Studying business executives, Jiang et al. (2011) showed that a common ethnical background between the two trusting parties influences affect-based trust.

Within the MDS model, trust is studied between two parties, the trustor and the trustee. However, the role of third parties has attracted the attention of trust scholars, pushing trust further from a narrow, dyadic perspective (Dirks & De Jong, 2022). Embeddedness theory and over- and under-socialised communities sit at the centre of several studies concerned with a field of interaction outside the isolated dyadic relationship (Ferrin et al., 2006; McEvily et al., 2003). Concerning the present literature review, such an influence appears relevant, as it is common for a trusted third party (e.g., the family doctor) to request the PHI on behalf of another entity (e.g., the country's healthcare system).

## **3. Review Methodology**

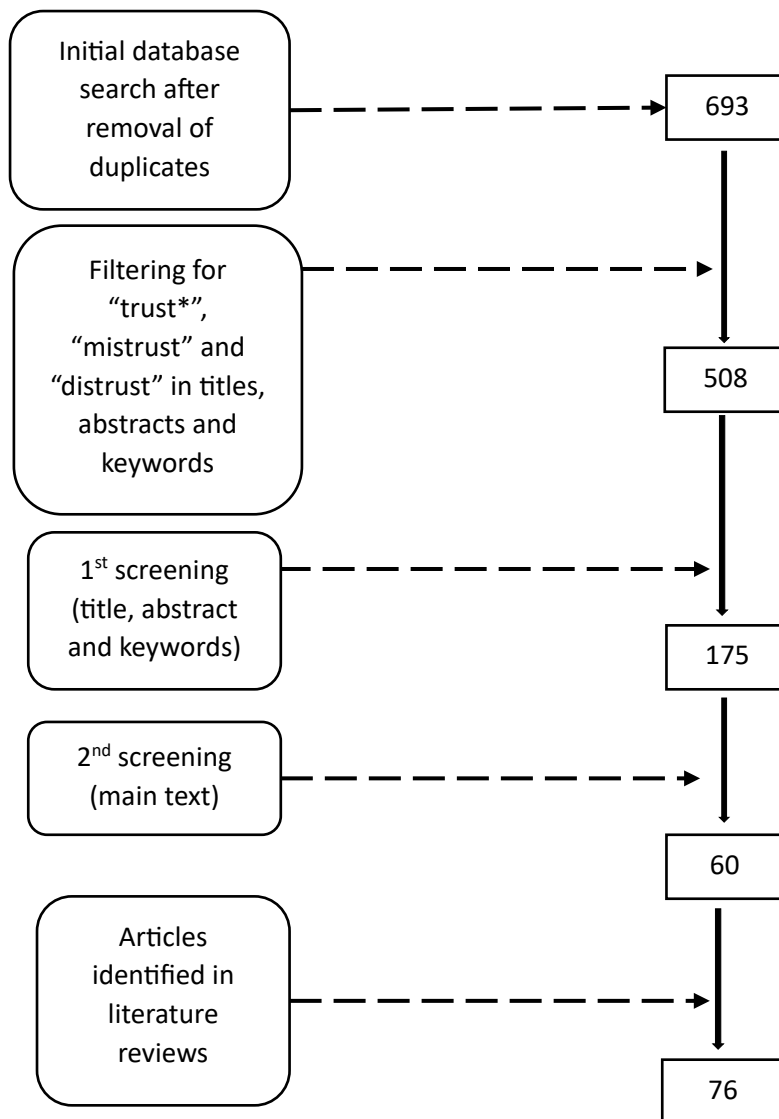
### **3.1 Searching the Literature**

The Web of Science, Scopus and PubMed databases were searched by one reviewer in March of 2021. The search fields included keywords pertaining to data sharing, health information, attitudes, empirical studies and trust (for the complete set of keywords see appendix A). This search provided a total of 981 articles, of which 693 were retained after the removal of duplicates. A second filtering through the Mendeley desktop citation manager (v.1.18) was conducted, by searching the titles, abstracts and keywords of the articles for "trust" (truncated), "mistrust", and "distrust", using the built-in search function. This yielded a total of 508 articles.

In the next phase, the articles were screened for fit by reading the abstract and noting the publication type for each. Here, particular attention was paid to inclusiveness, by erring on the side of caution, as advised by (Okoli & Schabram, 2012). The articles had to satisfy all of the following criteria:

- They reported empirical findings using primary data. Any literature reviews identified were excluded from this scoping review, although they aided the subsequent article search.
- They concerned perspectives, opinions, or decisions regarding PHI aggregation, reuse, disclosure, sharing, linking, or other similar action. This was done to include studies regardless of the purpose of request to share PHI.
- Trust was used to inform the research instrument or was mentioned in the empirical findings. References which were merely superficial were not considered.
- They studied patients or members of the general population as trustors (i.e., articles measuring opinions of only physicians or other stakeholders were excluded).
- The information requested or shared was for secondary uses (i.e., for inclusion in Electronic Health Records, research participation or similar reasons - not for immediate treatment of the responder), as this is the focus of our scoping review.
- The sharing of information concerned living adults and affected the responders only (subjects must be self-consentees; e.g., studies on the sharing of information of sick adults by their families were excluded in order to assure a direct link between trusting beliefs and personal risk of sharing). Studies on people with intellectual disabilities were also excluded.
- The responses in the reported studies were collected in the years 2001-2021 (both years inclusive), to ensure recency.
- They were published in journals or books. Conference proceedings and grey literature were excluded.
- They had a full-text version in English.

After applying the above criteria to the article abstracts, 174 articles were deemed fit for main-text screening. The main-text screening applied the same criteria used for the abstract screening, reducing the article pool to 60 articles. The references of literature reviews identified in the initial search were examined (manual backwards search) and led to the inclusion of a further 16 articles, based on the same criteria and a main-text examination, bringing the final number of articles to 76 (Figure 1).



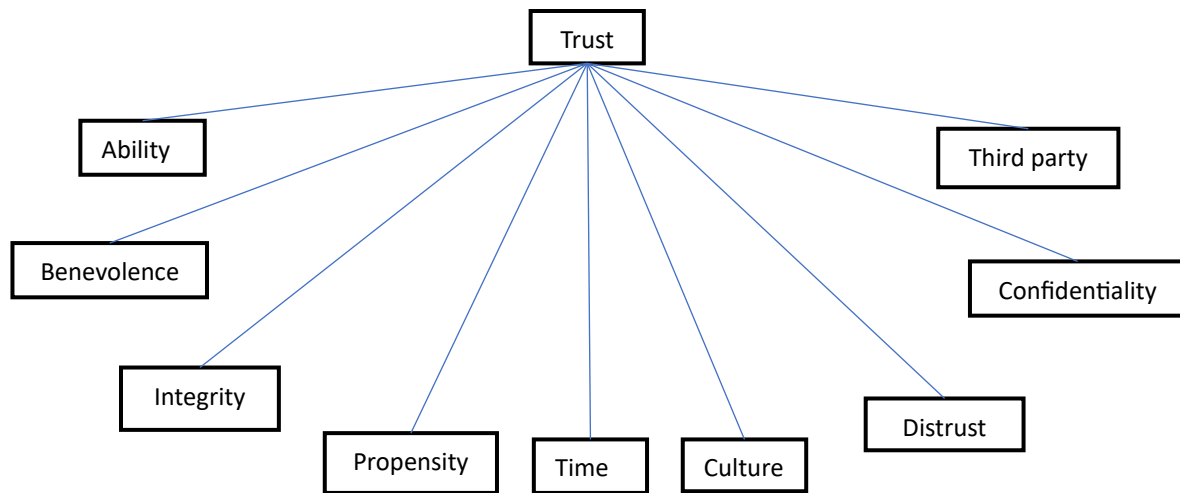
**Figure 1. The literature searching process**

### 3.2 Analysing the Literature

This scoping review used the MDS model as the core component of its analytical lens. Constructs not present in the original model were also included, and the already included ones were further developed (see 2.2), in order to provide an investigation that is both updated and more suitable to the studied field of PHI sharing. The constructs found in the reviewed empirical papers were deductively categorised based on the analytical lens. These constructs are ability, benevolence, integrity, propensity, confidentiality, time, culture, distrust, and third-party influence (Figure 2).

To clarify the constitution of some constructs described in section 2, time-related constructs included experience, relationship length, familiarity, and frequency of interaction. Following

the remark of (Schoorman et al., 2007) regarding the distinction between trust and distrust, we distinguished between these two constructs only if both had been reported or operationalised in the same empirical study. In any other case, distrust was synthesised together with trust. Finally, confidentiality comprised concepts relating to the protection of private information. This categorisation was based on (Hall et al., 2001; see 2.2 above), with the exception of the “will not be used against them” dimension, which was included in the benevolence construct.



**Figure 2. The analytical lens (influencers of trust)**

Each construct was matched to a code, with each code comprising several sub-codes (see Appendix B). To avoid situations where two identified constructs were named differently but referred to the same construct (and vice versa), the coding was based on the essential definition of the constructs, instead of how they were labelled in the articles.

The evaluation of past operationalisations of trust in empirical PHI-sharing research was informed by the research instruments employed by the researchers. To that end, the collected articles’ methods sections were studied. Only studies employing one or more elements of the analytical lens, either as separate constructs or as dimensions of a composite one, were considered. To assess whether the constructs included in our analytical lens can be used for improving the operationalisation of trust in PHI-sharing research, we looked into the results of the collected literature. The results sections and, occasionally, the discussion and methods sections were consulted. In the case of the qualitative studies, in order for a study to be included in the analysis, an element of the analytical lens had to be mentioned explicitly, or be discernible from the available quotes or the researchers’ coding. Moreover, the construct of

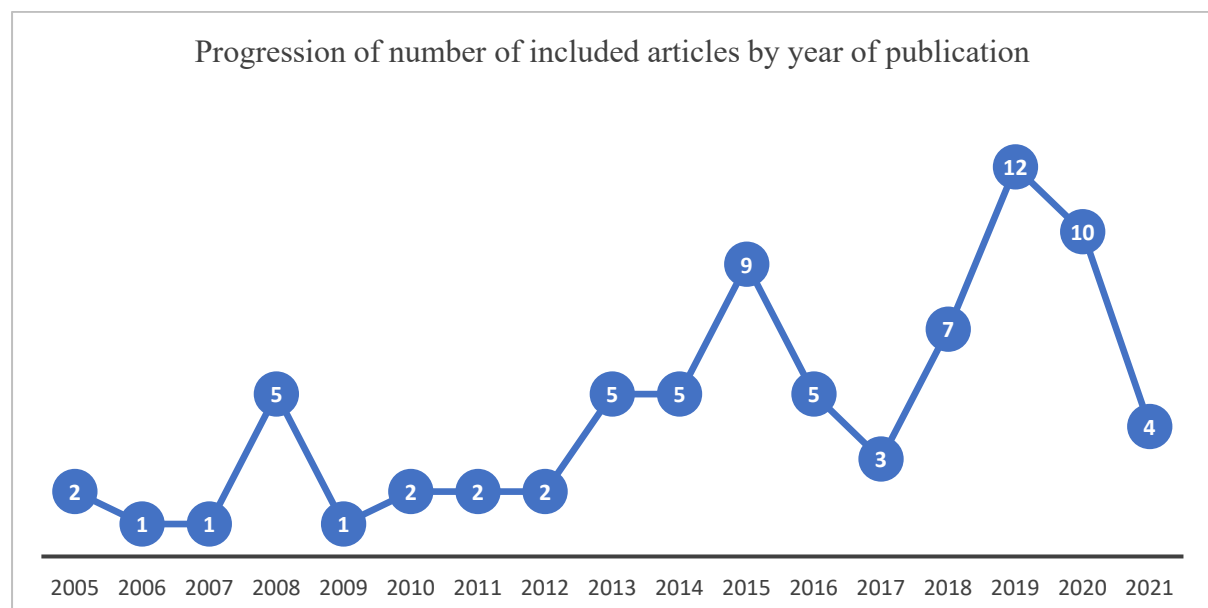
trust had to be mentioned explicitly by either the researcher or the respondent, to assure relevance. In quantitative studies, the elements of the analytical lens had to either have a significant effect on trust or a sharing intention (if they were utilised as separate constructs), or to show high reliability, if they were used as a dimension of a trust construct.

All data was extracted by one reviewer, by manual coding, and each article was marked for the inclusion of a construct/code. Multiple identifications in the same article did not affect the reported results.

## 4. Results

### 4.1 General Findings

The 76 collected articles were published between 2005 and 2021 (see Figure 3). 36 studies employed quantitative, 29 qualitative and 11 mixed methods. The most common data collection method was the survey (39 studies), followed by interviews (21 studies), focus groups (18 studies), and deliberations (3 studies). Out of the 75 studies using unique samples, 38 studies took place in the USA, 7 in England, 5 in Australia, 4 in Canada, 3 in Switzerland, 2 in Scotland, 2 worldwide, and 14 in other countries or across one or more countries. The participants were asked their opinions on a wide range of uses of their information, including use for or participation in medical research, biobanking, genomic and DNA research, Electronic Health Record systems, collection of administrative data, and access to medical records. The respondents were either sampled because they belonged to the general population (45 studies), or they were approached through their patient status or their proximity to other patients or health institutes, hospitals, or other similar places (33 studies).



**Figure 3. Frequency of included articles by year of publication**

#### **4.2 Operationalisation of Trust - Research Instruments**

In total, 37 studies informed their measuring instruments by using at least one of the elements of the analytical lens. The findings relating to the main and additional constructs of the analytical lens were primarily gathered from quantitative studies, and are presented in tables 1 and 2.

14 studies informed their research instruments based on at least one antecedent of the MDS model, either utilizing it as a dimension of trust or trustworthiness to create a construct, or operationalising it as a separate variable. Amongst them, seven took all three trustworthiness dimensions into consideration, and six used them to develop the studies' research instruments (Li et al., 2014; Platt et al., 2018; Platt & Kardia, 2015; Platt, Raj, Büyüktür, et al., 2019; Raj et al., 2020; Shen et al., 2019). It is noteworthy that three of those six studies were published by the same first author, who also contributed in a fourth one. Finally, only four studies included all four antecedents of trust in the MDS model (ability, benevolence, integrity, propensity), with three of them published by the same first author. This indicates that the already small number of references is inflated by the persistent work of a small group of researchers. The remaining identified constructs show a similar relative frequency to the ones noted in the article results' synthesis (see 4.3). Specifically, confidentiality is represented most often, followed by culture and time, while third-party influence was identified only once.

From the above, it becomes evident that trust is operationalised with limited regard for the MDS model. Additionally, in a significant portion of the studies, none of the analytical lens' constructs were identified (Figure 4). While it is not to say that those studies used an uninformed or generic trust construct, as a review with a different analytical lens could have identified important patterns in their measuring instruments, the low number is still surprising. Occasionally, such an absence of a theoretically rich, multidimensional trust construct was reported by the authors as a limitation of their research (e.g., Abdelhamid et al., 2017; Serrano et al., 2016). The findings of this literature review converge towards a widespread inconsistency between generally accepted trust models and the operationalisation of trust in PHI-sharing research.

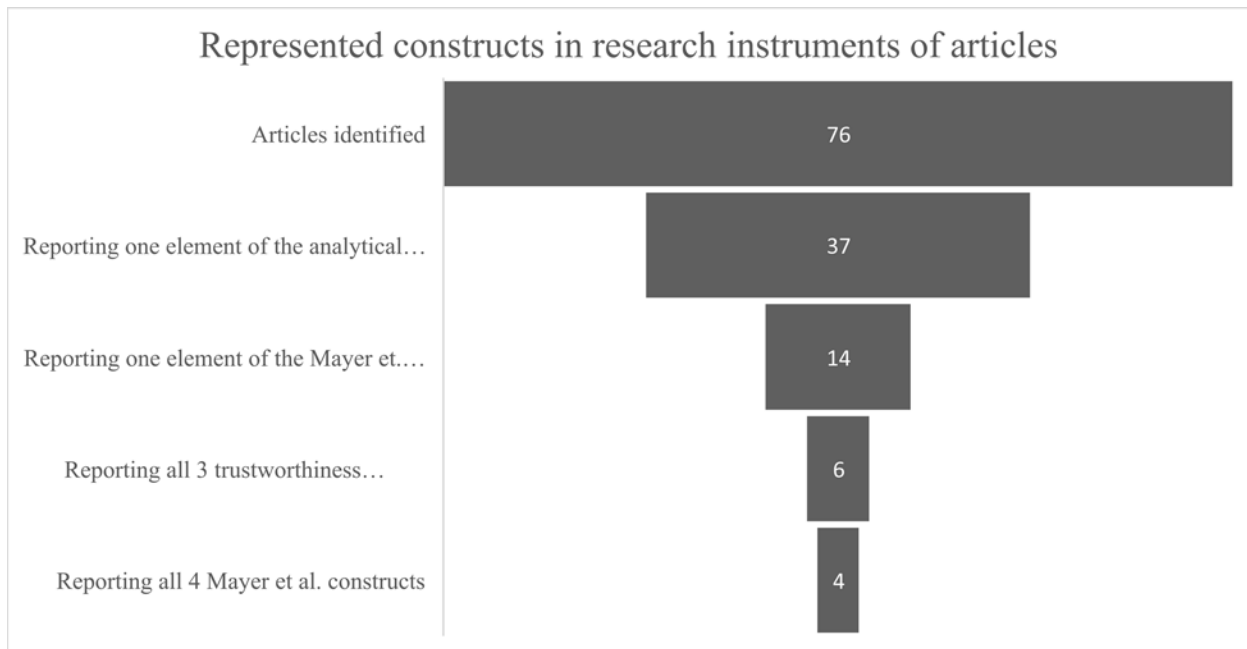
Construct	Count	References
Ability	9	Esmailzadeh (2019b); Kettis-Lindblad et al. (2006); Li et al. (2014); Platt et al. (2018); Platt and Kardia (2015); Platt, Raj, Büyüktür, et al. (2019); Raj et al. (2020); ***Shen et al. (2019); Teixeira et al. (2011)
Benevolence	12	Beskow and Dean (2008); Brewer et al. (2014); Esmailzadeh (2019a); Li et al. (2014); Overby et al. (2015); Platt et al. (2018); Platt and Kardia (2015); Platt, Raj, Büyüktür, et al. (2019); Raj et al. (2020); Serrano et al. (2016); ***Shen et al. (2019); Teixeira et al. (2011)
Integrity	11	Abdelhamid (2018); Beskow and Dean (2008); Esmailzadeh (2019a, 2019b); Li et al. (2014); Overby et al. (2015); Platt et al. (2018); Platt and Kardia (2015); Platt, Raj, Büyüktür, et al. (2019); Raj et al. (2020); ***Shen et al. (2019)
Propensity	5	Mello et al. (2018); Platt et al. (2018); Platt and Kardia (2015); Platt, Raj, Büyüktür, et al. (2019); ***Shen et al. (2019)

**Table 1. Main constructs from the MDS model (synthesis of research instruments) \*\*\*qualitative study**

Construct	Count	References
Confidentiality	22	Abdelhamid (2018); Abdelhamid et al. (2017); Andrews et al. (2014); Bearth and Siegrist (2020); Buckley et al. (2011); Damschroder et al. (2007); Dhopeswarkar et al. (2012); Dinev et al. (2016); Esmailzadeh (2019b); ***Greenhalgh et al. (2008); Holm et al. (2020); Juga et al. (2021); Morin et al. (2005); Papoutsi et al. (2015); Platt and Kardia (2015); Pratap et al. (2019); Rho et al. (2015); ***Shen et al. (2019); ***Stone et al. (2005); Weng et al. (2019); Willison et al. (2009); Willison et al. (2008)
Time	5	Esmailzadeh (2019b); Herian et al. (2014); Holm et al. (2020); Platt and Kardia (2015); Weidman et al. (2019)
Culture	7	Lor and Bowers (2018); Middleton et al. (2020); Papoutsi et al. (2015); Platt et al. (2018); ***Shen et al. (2019); ***Slegers et al. (2015); Weidman et al. (2019)
Distrust**	1	***Shen et al. (2019)
Third party influence	1	Weidman et al. (2019)

**Table 2. Additional constructs not part of the MDS model (synthesis of research instruments) \*\*searched only in qualitative studies \*\*\* qualitative study**





**Figure 4. Represented constructs in research instruments of articles**

### 4.3 Relevance of the Studied Constructs

To evaluate the relevance of the studied constructs in the context of PHI sharing, we analysed the results of the collected literature. The results of the qualitative studies were generally more suitable for that. 52 articles included at least one construct of the analytical lens and are presented in tables 3 and 4. The findings concerning the MDS model and those concerning the other related constructs are presented separately for clarity.

The three trustworthiness dimensions (ability, benevolence and integrity) were widely associated with positive trusting attitudes. Each of them was present in nearly half of the collected articles that included at least one element of the analytical lens. Considering that they were mostly mentioned spontaneously by the responders, their influence on trust appears to be fundamental. Propensity was also present, but much less frequently than trustworthiness. In total, of the 53 articles, only 13 have no mention of any of the above four constructs in their results. Although it is not always clear if the four elements of the MDS model influence PHI sharing directly or through the mediation of trust, it is nevertheless evident that they play a significant role in the PHI-sharing process.

Construct	Count	References
Ability	21	Alaqla et al. (2018); Bosisio et al. (2021); Bussone et al. (2020); Carson et al. (2019); Damschroder et al. (2007); Darquy et al.

		(2016); *Esmailzadeh (2019a); *Esmailzadeh (2019b); Harle et al. (2018); Jones et al. (2020); Jones et al. (2017); Kerns et al. (2013); *Kettis-Lindblad et al. (2006); *Li et al. (2014); Mozersky et al. (2020); Papoutsi et al. (2015); *Platt et al. (2018); Shen et al. (2019); Spencer et al. (2016); Stablein et al. (2015); *Teixeira et al. (2011)
Benevolence	25	*Brewer et al. (2014); Bussone et al. (2020); Carson et al. (2019); Damschroder et al. (2007); Darquy et al. (2016); De Vries et al. (2019); *Esmailzadeh (2019a); Greenhalgh et al. (2008); Harle et al. (2018); Jones et al. (2020); Jones et al. (2017); Kerns et al. (2013); Lemke et al. (2010); *Li et al. (2014); Lysaght et al. (2020); Mählmann et al. (2018); McGuire et al. (2008); *Platt et al. (2018); Pratap et al. (2019); Shen et al. (2019); Slegers et al. (2015); Spencer et al. (2016); *Teixeira et al. (2011); Trinidad et al. (2010); Weng et al. (2019)
Integrity	22	*Abdelhamid (2018); Alaqra et al. (2018); Bosisio et al. (2021); Bussone et al. (2020); De Vries et al. (2019); *Esmailzadeh (2019a); *Esmailzadeh (2019b); Greenhalgh et al. (2008); Harle et al. (2018); Jamal et al. (2014); Lee et al. (2016); Lemke et al. (2010); *Li et al. (2014); Lor and Bowers (2018); Mählmann et al. (2018); Merson et al. (2015); Mozersky et al. (2020); *Platt and Kardia (2015); Pratap et al. (2019); Shen et al. (2019); Slegers et al. (2015); Trinidad et al. (2010)
Propensity	8	Bosisio et al. (2021); Jones et al. (2017); McGuire et al. (2008); *Mello et al. (2018); *Platt et al. (2018); *Platt and Kardia (2015); Shen et al. (2019); Spencer et al. (2016)

**Table 3. Main constructs from the MDS model (synthesis of results)**

**\*quantitative study**

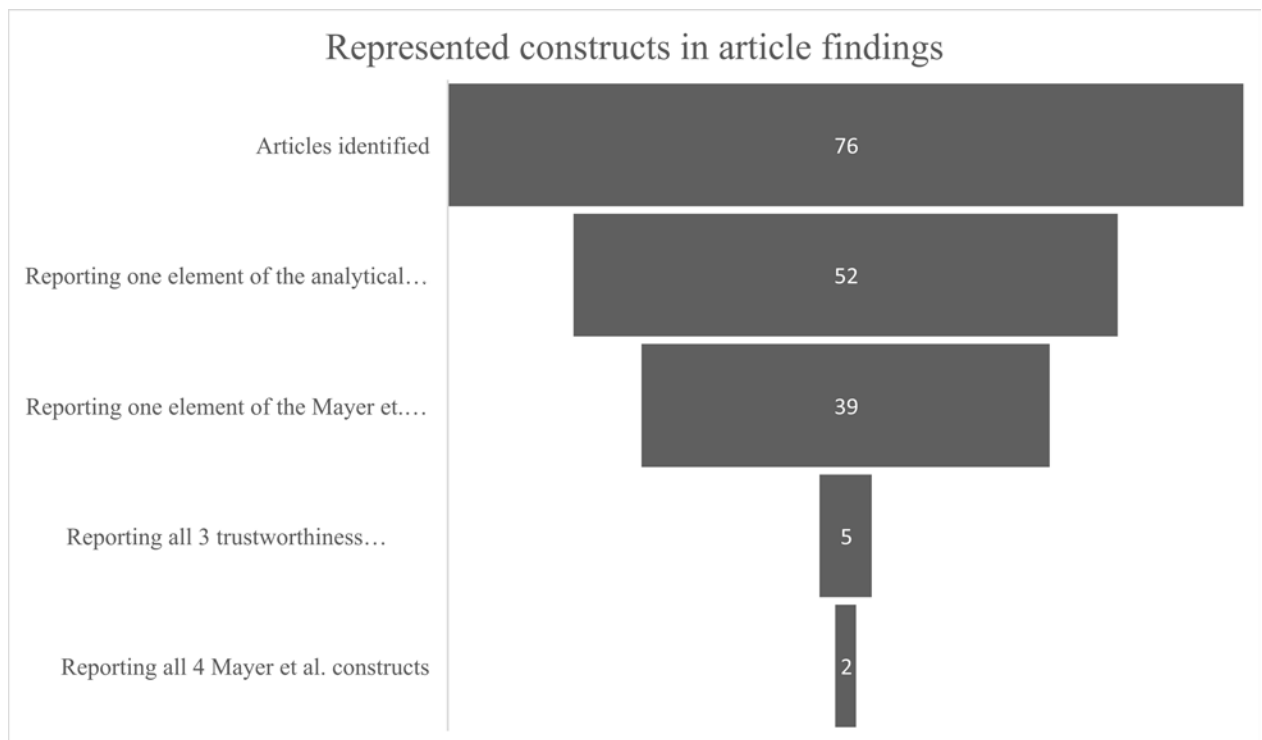
Apart from the MDS model constructs, the other constructs of the analytical lens were also present in the collected literature results. Confidentiality was the most frequently mentioned construct by the respondents, with more than two thirds of the articles including references to it. Distrust and mistrust were also mentioned by the researchers and the respondents, but in almost all cases they were used to denote the opposite of trust, without being placed on separate spectra. Time-related constructs were also identified, in the form of familiarity, relationship length, and frequency of contact. In most cases their influence on trust was significant and positive, but it rarely was of no significance (Weidman et al., 2019). Cultural influence on trust was also present in a significant portion of the collected literature, usually stemming from the norms or values of the respondent's culture, or the minority status of the responder. Lastly, constructs pertaining to third-party influence were found mainly in the form of trust transfer and trust by proxy.

<b>Construct</b>	<b>Count</b>	<b>References</b>
Confidentiality	37	Alaqra et al. (2018); *Bearth and Siegrist (2020); Beskow and Dean (2008); Broes et al. (2020); Bussone et al. (2020); Carson et al. (2019); Damschroder et al. (2007); Darquy et al. (2016); *Dinev et al. (2016); Esmaeilzadeh (2019b); Grant et al. (2013); Harle et al. (2018); Hill et al. (2013); Jamal et al. (2014); Jones et al. (2020); Jones et al. (2017); Kerns et al. (2013); Lee et al. (2016); Lemke et al. (2010); Lucero et al. (2015); Lysaght et al. (2020); Mählmann et al. (2018); McGuire et al. (2008); Merson et al. (2015); Morin et al. (2005); Mozersky et al. (2020); Papoutsi et al. (2015); *Platt et al. (2018); *Platt and Kardia (2015); Pratap et al. (2019); Shen et al. (2019); Skatova et al. (2019); Spencer et al. (2016); Stablein et al. (2015); Trinidad et al. (2010); Weng et al. (2019); Willison et al. (2009)
Time	19	Broes et al. (2020); Bussone et al. (2020); Carson et al. (2019); De Vries et al. (2019); *Esmaeilzadeh (2019b); Grant et al. (2013); Greenhalgh et al. (2008); *Herian et al. (2014); Jamal et al. (2014); Jones et al. (2020); Kerns et al. (2013); Lee et al. (2016); Lor and Bowers (2018); Mozersky et al. (2020); *Platt and Kardia (2015); Shen et al. (2019); Skatova et al. (2019); Stone et al. (2005); Willison et al. (2009)
Culture	7	Alaqra et al. (2018); *Dinev et al. (2016); Lee et al. (2016); Lor and Bowers (2018); *Middleton et al. (2020); *Platt et al. (2018); *Weidman et al. (2019)
Distrust**	18	Alaqra et al. (2018); Broes et al. (2020); De Vries et al. (2019); Grant et al. (2013); Harle et al. (2018); Jones et al. (2020); Jones et al. (2017); Lee et al. (2016); Lemke et al. (2010); Lucero et al. (2015); Lysaght et al. (2020); Mählmann et al. (2018); McGuire et al. (2008); Shen et al. (2019); Skatova et al. (2019); Slegers et al. (2015); Stablein et al. (2015); Trinidad et al. (2010)
Third party influence	7	Alaqra et al. (2018); Bussone et al. (2020); Kerns et al. (2013); Lee et al. (2016); Lor and Bowers (2018); Skatova et al. (2019); Willison et al. (2009)

**Table 4. Additional constructs not part of the MDS model (synthesis of results)**

**\*quantitative study \*\*searched in qualitative articles only**

From the synthesis of the literature results (Figure 5) it is concluded that the general trust conceptualisations, especially the MDS model, are applicable in the context of PHI sharing, as most of the constructs of the model were independently and frequently identified in the results of the literature. Furthermore, confidentiality and time-related constructs to a large extent, and culture and the influence of third parties to a lesser extent, were frequently found to affect trust.



**Figure 5. Represented constructs in article results**

## 5. Discussion

The successful adoption and use of HIS are contingent on trust. Although numerous studies have conceptualised and operationalised trust in health research (for a review see Taylor et al., 2023), most empirical investigations do not clarify their perspective and, consequently, state-of-the-art knowledge is fragmented. Therefore, this scoping review looked into this broad field research in order to understand what influences trust when people are asked to share their personal health information.

A number of literature reviews have synthesised similar empirical findings, in the context of PHI sharing. Hutchings et al. (2020) extensively and systematically reviewed healthcare consumer attitudes, pointing out the importance of confidentiality, privacy and familiarity in establishing trust, while Aitken et al. (2016) found interesting links between trust, familiarity and PHI sharing for research. Reviewing five studies on the attitudes of stakeholders towards research participation, (Bull et al., 2015) confirmed the need for integrity and benevolence, and pronounced the role of trusted third parties in building trust. Concentrating on the UK and Irish public, (Stockdale et al., 2018) confirmed the importance of perceived ability, benevolence and privacy assurance in building trust. Shen et al. (2019) emphasised on the bidirectional

relationship of privacy concerns and trust, while reporting on several of the trust-influencing constructs that were identified in this review as well.

Nevertheless, none of the above-mentioned reviews sought to construct an ontological map of trust's antecedents and related constructs based on a deductive analysis of empirical research. Moreover, they have not focused on the operationalisation of trust in PHI-sharing research, or encompassed a broader, more inclusive range of secondary uses of PHI. Therefore, this review offers valuable insight into the conceptualisation and operationalisation of trust in the context of PHI sharing. The collected literature results confirm the appropriateness of the MDS model for the purpose of researching PHI sharing. The constructs of confidentiality, time, third-party influence, and culture are also relevant when measuring trust. At the same time, it appears that the operationalisations of trust in empirical research have not fully integrated these constructs into their applied research instruments.

### **5.1 Discussing the Results**

With regard to the MDS model, the responders of the reviewed studies underlined the significance of ability, benevolence and integrity in building trust. The fact that these three constructs usually emerged in an unforced, natural manner during the interviews or the focus groups reaffirms their trust-measuring relevance in the PHI-sharing context. Propensity to trust was not identified as frequently as the three dimensions of trustworthiness. A possible explanation for this may lie in the nature of the construct. In contrast to trustworthiness, propensity focuses on the trustor instead of the trustee and, since the trustee was the main point of focus in most studies, it could have overshadowed the importance of trustor characteristics. Indeed, this focus on trustees seems to be a common trend in trust research (Möllering, 2019).

Of the constructs not included in the MDS model, confidentiality was the most commonly occurring one, followed by time, culture and third-party influence. Often, confidentiality and its related constructs were not easily distinguishable from trustworthiness and trust. An example of this can be found in (Weng et al., 2019), where ability, a trust antecedent, is mentioned as a factor that fosters confidentiality, hinting that confidentiality and the dimensions of trustworthiness overlap. This raises the question whether the construct labelled as confidentiality in this review is an independent construct, or a context within which trust operates. Even in the case of confidentiality being an independent construct, it is still unclear whether it is an antecedent of trust, or a result of it. Scholars have studied confidentiality as a precursor of trust (Esmailzadeh, 2019a), while others have studied the opposite direction

(Dinev et al., 2016), with both finding significant relationships. Shen, Bernier et al. (2019), explored a bidirectional relationship between the two constructs, and found evidence for both directions of influence. Given the prominent role of confidentiality and privacy in PHI-sharing behaviour, research capable of causal inference would be particularly valuable.

Regarding the effect of time on trust, long-lasting relationships, repeated interaction and familiarity were, in most cases, positively associated to trusting attitudes. This finding is in line with the dynamic development of trust proposed in the MDS model. Being aware of the temporal effects on trust is essential to our understanding of trust and its related constructs. Propensity, for example, has been theorised to be more important in the beginning of a relationship, when there has not been enough time for the other trust antecedents to develop (initial trust; McKnight et al., 1998). This can be relevant when researching trust in HIS, where the users' familiarity can be low, such as the with the use of AI and chatbots in healthcare.

The influence of culture on trust constituted a significant part of the literature results. The most characteristic example was found in (Middleton et al., 2020) where the researchers accentuated the varying trust attitudes towards genomic data sharing on a worldwide scale with the replication of an identical survey instrument. Similar findings were reported by (Dinev et al., 2016), regarding privacy concerns and trust towards Electronic Health Records between citizens in Italy and the USA. For explaining such between-cultures differences, one should consider the familiarity of the studied population with the technologies involved with PHI-sharing, or their propensity, which is partly based on culture (Schoorman et al., 2007).

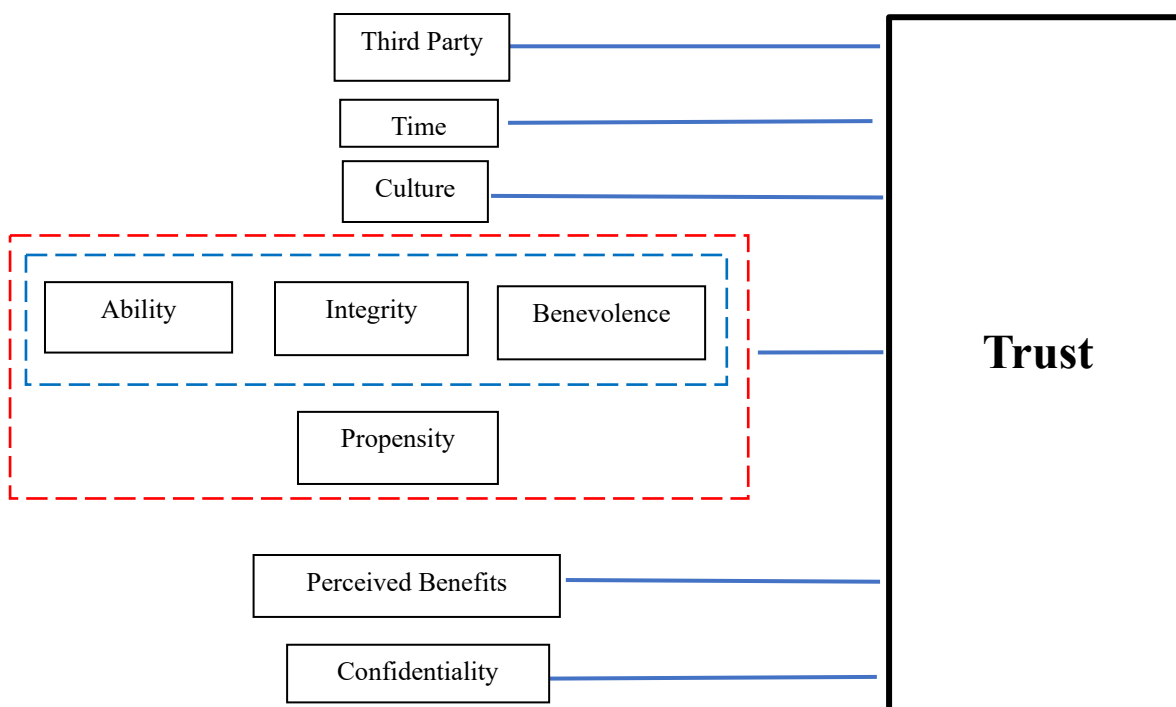
The results for distrust point towards a single-spectrum trust construct, as distrust was generally found to be used as merely a synonym for lack of trust. In one study distrust was briefly mentioned in the discussion to be different from the lack of trust, without further elaboration (Grant et al., 2013). Indeed, it appears that the difference between low trust and distrust is small, if it exists at all, as noted by (Schoorman et al., 2007).

Our review also identified research showing perceived benefits to be a predictor of trust. In a survey-based empirical study, (Esmailzadeh, 2019b) found evidence supporting the hypothesis that the perceived benefits of a Health Information Exchange system increase people's cognitive trust in that system. Similar findings have been reported in (Platt, Raj, & Kardia, 2019) and in (Jenkner et al., 2022; not included in our review). These serendipitous observations could offer yet another angle into the understanding of trust in PHI sharing. For example, perceived benefits of PHI disclosure could signal trust by the trustee, enabling

cooperation (or making the possibility of cooperation salient), which, in turn, might increase trust (Korsgaard et al., 2018). Alternatively, one could perceive them as a gesture of care and goodwill, thus as a sign of benevolence.

## 5.2 Improved Conceptualisation and Operationalisation of Trust

Trust's operationalisation has been reported to be fragmented even within the same field (McEvily & Tortoriello, 2011; organisational research). Our review finds a similar fragmentation in the case of PHI-sharing research. The MDS model has not been found to be widely replicated, and any replications of it were usually conducted by the same authors. Moreover, when its constructs are used, it is done so sparingly and incompletely. Looking at all elements of our analytical lens, only confidentiality was broadly present in the measuring instruments used by the identified empirical studies. This points out a significant inconsistency between the appropriateness of the constructs included in our analytical lens, and their use in the operationalisation of trust in extant PHI-sharing research. Building on our review's findings, we suggest PHI-relevant directions for conceptualising and operationalising trust, based on the MDS model, and enhancing it with context-relevant constructs (Figure 6). Apart from an improvement of the measuring of trust in a PHI context, the summation of our suggested constructs into a trust model can streamline future studies through the standardisation of the research instruments. This could not only make comparisons between studies more efficient, but more meaningful as well (McEvily & Tortoriello, 2011).



**Figure 6. Proposed constructs to consider when measuring trust in PHI-sharing contexts (the MDS model antecedents enclosed in red, trustworthiness dimensions enclosed in blue).**

This study offers contributions to theory and research. Firstly, it assesses the applicability of an organisational trust model in the context of PHI sharing. Based on the literature's results' synthesis, it can be argued that the MDS model remains relevant in the studied context of PHI sharing, thus its application and further study are justified. Secondly, the synthesis of other trust-related constructs is a step towards a more cohesive model of trust in the context of PHI, with increased content validity. The frequent and usually unprompted mentioning of confidentiality and privacy suggests that they are naturally relevant, and measuring trust without considering them might result in incomplete and context-blind models. Similarly, time, familiarity, and relationship length show a clear effect on trust. Finally, taking into account the effect of trusted third parties can help with the overcoming of the narrow, dyadic measuring of trust, while the consideration of cultural influences appears to be relevant and is highly suggested.

Understanding the importance of trust in the context of healthcare, especially when it comes to the adoption and use of new technologies, and the sharing of PHI that they necessitate, is vital. By analysing trust's structure at a deeper level, the public's concerns can be addressed more clearly, and data collection can be carried out more effectively, when the responsible entities navigate in ways that are aligned with the public's volitions, insecurities, or objections. We believe that the above recommendations can progress that understanding, as a necessary step towards realising the benefits of HIS in addressing the challenges facing healthcare systems now and in the future.

### **5.3 Limitations**

The main limitation of this study lies in the selection of literature that formed the core of the analytical lens. Although the additional constructs that formed the analytical lens provided some necessary recency and contextual relevance, the MDS model brought a dated perspective, which originated in a different literature discipline than that of the reviewed articles. Moreover, there is an absence of emotional factors in the MDS model, as its authors note in their revisiting the original paper (Schoorman et al., 2007). Indeed, the importance of emotions in trust (also see affect-based trust; van Knippenberg, 2018) has been noted in medical (Hall et al., 2001) and information-sharing contexts (Tomlinson et.al., 2020). Thus, our cognitive-centred lens may not have been able to capture some important antecedents of trust. Articles studying



information sharing for primary use (i.e., for treating the patient sharing the information), or the opinions of minors, were excluded from this literature review, as were papers not written in English. This further limits the generalisability of our results.

#### **5.4 Future Research**

As with any conceptual suggestion, research that collects the highlighted constructs into a model and empirically validates it is of great interest. At the same time, those constructs could benefit from further exploration. As elaborated above, confidentiality is highly relevant in the field of PHI-sharing research. Nevertheless, its nature remains nebulous. To that end, supplementary research is required to understand its relationship with trust in a PHI-sharing context, especially when combined with recent conceptualisations of perceived privacy risks (Karwatzki et al., 2022) and privacy concerns (Bansal & Nah, 2022) in IS research. Regarding multiple trustees and third-party influence, there is a lack of empirical, PHI-sharing research combining the two, despite the rich theoretical background. Future research can study the effect of trust transferability, especially when a trusted party mediates between the data holder and a distrusted party. This is especially relevant in HIS research, when the trustee is a technology new and unfamiliar to the patient (e.g., trust in AI-aided diagnosis), and a familiar and trusted person, such as the family doctor, acts as the mediator.

A perspective which includes both cognitive and emotional dimensions of trust presents a promising research avenue. Suggested research areas include the effect of temporal factors, such as relationship length, on trust, especially since affect-based trust is thought to be very closely linked to the relationship between the trusting parties (van Knippenberg, 2018). Our review's serendipitous findings regarding the role of the benefits in increasing trust provide an incentive for further research. Connecting this with the influence of culture (in the scope of individualism and collectivism; Hofstede, 1984), we find that studying the impact of perceived altruistic or egoistic benefits (White et al., 2009) on trust to be a particularly appealing direction, especially when the trustor is asked to share PHI for both primary and secondary (e.g., research) use.

#### **6. Conclusion**

Trust is necessary for the adoption and use of HIS. While the MDS conceptualisation is widely represented in the empirical findings of the published PHI-sharing research, trust's operationalisation in the methods of the empirical studies is not sufficiently founded on it, with the measurement of the construct of trust being fragmented and inconsistent. Based on the

findings of this scoping review, PHI-relevant approaches for measuring of trust are proposed, incorporating confidentiality, trust transfer, as well as time-related and cultural factors into the MDS model. These suggested pathways can offer a better operationalisation of trust when conducting studies that concern the sharing of PHI.

### **Declaration of interests**

The authors have no conflict of interest to declare.

### **Availability of data and materials**

The datasets used and analysed for the review are available from the corresponding author upon reasonable request.

## References

- Abdelhamid, M. (2018). Greater patient health information control to improve the sustainability of health information exchanges. *Journal of Biomedical Informatics*, 83, 150-158. <https://doi.org/10.1016/j.jbi.2018.06.002>
- Abdelhamid, M., Gaia, J., & Sanders, G. L. (2017). Putting the focus back on the patient: How privacy concerns affect personal health information sharing intentions. *Journal of Medical Internet Research*, 19, 1-12. <https://doi.org/10.2196/jmir.6877>
- Aitken, M., De St Jorre, J., Pagliari, C., Jepson, R., & Cunningham-Burley, S. (2016). Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. *BMC Medical Ethics*, 17(1), 1-24. <https://doi.org/10.1186/s12910-016-0153-x>
- Alaqra, A. S., Fischer-Hübner, S., & Framner, E. (2018). Enhancing privacy controls for patients via a selective authentic electronic health record exchange service: Qualitative study of perspectives by medical professionals and patients. *Journal of Medical Internet Research*, 20. <https://doi.org/10.2196/10954>
- Andrews, L., Gajanayake, R., & Sahama, T. (2014). The Australian general public's perceptions of having a personally controlled electronic health record (PCEHR). *International Journal of Medical Informatics*, 83, 889-900. <https://doi.org/10.1016/j.ijmedinf.2014.08.002>
- Baer, M. D., & Colquitt, J. A. (2018). Why do people trust?: Moving toward a more comprehensive consideration of the antecedents of trust. *The Routledge companion to trust*, 163-182.
- Bansal, G., & Nah, F. F. H. (2022). Internet privacy concerns revisited: oversight from surveillance and right to be forgotten as new dimensions. *Information & Management*, 59(3), 103618.
- Bearth, A., & Siegrist, M. (2020). Psychological factors that determine people's willingness-to-share genetic data for research. *Clinical Genetics*, 97, 483-491. <https://doi.org/10.1111/cge.13686>
- Belfrage, S., Helgesson, G., & Lynøe, N. (2022). Trust and digital privacy in healthcare: a cross-sectional descriptive study of trust and attitudes towards uses of electronic health data among the general public in Sweden. *BMC medical ethics*, 23(1), 1-8.
- Beskow, L. M., & Dean, E. (2008). Informed consent for biorepositories: assessing prospective participants' understanding and opinions. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 17, 1440-1451. <https://doi.org/10.1158/1055-9965.EPI-08-0086>
- Bijlsma-Frankema, K., Sitkin, S. B., & Weibel, A. (2015). Distrust in the balance: The emergence and development of intergroup distrust in a court of law. *Organization Science*, 26(4), 1018-1039. <https://doi.org/10.1287/orsc.2015.0977>
- Bosisio, F., Barazzetti, G., Koutaissoff, D., & Spencer, B. (2021). Patients' decision to contribute to a biobank in the light of the patient-recruiter relationship-a qualitative study of broad consent in a hospital setting. *Journal of Community Genetics*, 12, 15-25. <https://doi.org/10.1007/s12687-020-00479-z>
- Brewer, L. C., Hayes, S. N., Parker, M. W., Balls-Berry, J. E., Halyard, M. Y., Pinn, V. W., & Breitkopf, C. R. (2014). African American women's perceptions and attitudes regarding participation in medical research: The mayo clinic/the links, incorporated partnership. *Journal of Women's Health*, 23, 681-687. <https://doi.org/10.1089/jwh.2014.4751>
- Broes, S., Verbaanderd, C., Casteels, M., Lacombe, D., & Huys, I. (2020). Sharing of Clinical Trial Data and Samples: The Cancer Patient Perspective. *Frontiers in Medicine*, 7. <https://doi.org/10.3389/fmed.2020.00033>

- Buckley, B. S., Murphy, A. W., & MacFarlane, A. E. (2011). Public attitudes to the use in research of personal health information from general practitioners' records: A survey of the Irish general public. *Journal of Medical Ethics*, 37, 50-55. <https://doi.org/10.1136/jme.2010.037903>
- Bull, S., Cheah, P. Y., Denny, S., Jao, I., Marsh, V., Merson, L., ... & Parker, M. (2015). Best practices for ethical sharing of individual-level health research data from low-and middle-income settings. *Journal of Empirical Research on Human Research Ethics*, 10(3), 302-313.
- Bussone, A., Kasadha, B., Stumpf, S., Durrant, A. C., Tariq, S., Gibbs, J., Lloyd, K. C., & Bird, J. (2020). Trust, Identity, Privacy, and Security Considerations for Designing a Peer Data Sharing Platform between People Living with HIV. *Proceedings of the ACM on Human-Computer Interaction*, 4. <https://doi.org/10.1145/3415244>
- Carson, C., Hinton, L., Kurinczuk, J., & Quigley, M. (2019). 'I haven't met them, I don't have any trust in them. It just feels like a big unknown': A qualitative study exploring the determinants of consent to use Human Fertilisation and Embryology Authority registry data in research. *BMJ Open*, 9. <https://doi.org/10.1136/bmjopen-2018-026469>
- Carter, P., Laurie, G. T., & Dixon-Woods, M. (2015). The social licence for research: Why care.data ran into trouble. *Journal of Medical Ethics*, 41, 404-409. <https://doi.org/10.1136/medethics-2014-102374>
- Cavallone, M., & Palumbo, R. (2020). Debunking the myth of industry 4.0 in health care: insights from a systematic literature review. *The TQM Journal*, 32(4), 849-868.
- Colquitt, J. A., Scott, B. A., & Lepine, J. A. (2007). Trust, Trustworthiness, and Trust Propensity: A Meta-Analytic Test of Their Unique Relationships With Risk Taking and Job Performance. <https://doi.org/10.1037/0021-9010.92.4.909>
- Damschroder, L. J., Pritts, J. L., Neblo, M. A., Kalarickal, R. J., Creswell, J. W., & Hayward, R. A. (2007). Patients, privacy and trust: Patients' willingness to allow researchers to access their medical records. *Social Science and Medicine*, 64, 223-235. <https://doi.org/10.1016/j.socscimed.2006.08.045>
- Darquy, S., Moutel, G., Lapointe, A.-S., D'Audiffret, D., Champagnat, J., Guerroui, S., Vendeville, M.-L., Boespflug-Tanguy, O., & Duchange, N. (2016). Patient/family views on data sharing in rare diseases: Study in the European LeukoTreat project. *European Journal of Human Genetics*, 24, 338-343. <https://doi.org/10.1038/ejhg.2015.115>
- De Vries, R. G., Ryan, K. A., Gordon, L., Krenz, C. D., Tomlinson, T., Jewell, S., & Kim, S. Y. H. (2019). Biobanks and the Moral Concerns of Donors: A Democratic Deliberation. *Qualitative Health Research*, 29, 1942-1953. <https://doi.org/10.1177/1049732318791826>
- Dhopeswarkar, R. V., Kern, L. M., O'Donnell, H. C., Edwards, A. M., & Kaushal, R. (2012). Health care consumers' preferences around health information exchange. *Annals of Family Medicine*, 10, 428-434. <https://doi.org/10.1370/afm.1396>
- Dinev, T., Albano, V., Xu, H., D'Atri, A., & Hart, P. (2016). Individuals' Attitudes Towards Electronic Health Records: A Privacy Calculus Perspective. 19-50. [https://doi.org/10.1007/978-3-319-23294-2\\_2](https://doi.org/10.1007/978-3-319-23294-2_2)
- Dirks, K. T., & de Jong, B. (2022). Trust within the workplace: A review of two waves of research and a glimpse of the third. *Annual Review of Organizational Psychology and Organizational Behavior*, 9, 247-276. <https://doi.org/10.1146/annurev-orgpsych-012420-083025>
- Esmailzadeh, P. (2019a). The Impacts of the Perceived Transparency of Privacy Policies and Trust in Providers for Building Trust in Health Information Exchange: Empirical Study. *JMIR Medical Informatics*, 7(4), e14050-e14050. <https://doi.org/10.2196/14050>

- Esmaeilzadeh, P. (2019b). The process of building patient trust in health information exchange (HIE): The impacts of perceived benefits, perceived transparency of privacy policy, and familiarity. *Communications of the Association for Information Systems*, 45(1), 364-396. <https://doi.org/10.17705/1CAIS.04521>
- Ferrin, D. L., Dirks, K. T., & Shah, P. P. (2006). Direct and indirect effects of third-party relationships on interpersonal trust. *Journal of Applied Psychology*, 91(4), 870-883. <https://doi.org/10.1037/0021-9010.91.4.870>
- Grant, A., Ure, J., Nicolson, D. J., Hanley, J., Sheikh, A., McKinstry, B., & Sullivan, F. (2013). Acceptability and perceived barriers and facilitators to creating a national research register to enable 'direct to patient' enrolment into research: The Scottish Health Research Register (SHARE). *BMC Health Services Research*, 13. <https://doi.org/10.1186/1472-6963-13-422>
- Greenhalgh, T., Wood, G. W., Bratan, T., Stramer, K., & Hinder, S. (2008). Patients' attitudes to the summary care record and HealthSpace: Qualitative study. *BMJ*, 336, 1290-1295. <https://doi.org/10.1136/bmj.a114>
- Hall, M. A., Dugan, E., Zheng, B., & Mishra, A. K. (2001). Trust in Physicians and Medical Institutions: What Is It, Can It Be Measured, and Does It Matter? *Milbank Quarterly*, 79(4), 613-639. <https://doi.org/10.1111/1468-0009.00223>
- Harle, C. A., Golembiewski, E. H., Rahmanian, K. P., Krieger, J. L., Haggmayer, D., Mainous III, A. G., & Moseley, R. E. (2018). Patient preferences toward an interactive e-consent application for research using electronic health records. *Journal of the American Medical Informatics Association*, 25, 360-368. <https://doi.org/10.1093/jamia/ocx145>
- Herian, M. N., Shank, N. C., & Abdel-Monem, T. L. (2014). Trust in government and support for governmental regulation: The case of electronic health records. *Health Expectations*, 17, 784-794. <https://doi.org/10.1111/j.1369-7625.2012.00803.x>
- Hill, E. M., Turner, E. L., Martin, R. M., & Donovan, J. L. (2013). "let's get the best quality research we can": Public awareness and acceptance of consent to use existing data in health research: A systematic review and qualitative study. *BMC Medical Research Methodology*, 13. <https://doi.org/10.1186/1471-2288-13-72>
- Hillebrand, K., Hornuf, L., Müller, B., & Vrankar, D. (2023). The social dilemma of big data: Donating personal data to promote social welfare. *Information and Organization*, 33(1), 100452.
- Hofstede, G. (1984). *Culture's consequences: International differences in work-related values* (Vol. 5). sage.
- Holm, S., Kristiansen, T. B., & Ploug, T. (2020). Control, trust and the sharing of health information: the limits of trust. *Journal of Medical Ethics*. <https://doi.org/10.1136/medethics-2019-105887>
- Hong, S. J., & Cho, H. (2023). Privacy management and health information sharing via contact tracing during the COVID-19 pandemic: a hypothetical study on AI-based technologies. *Health Communication*, 38(5), 913-924.
- Hutchings, E., Loomes, M., Butow, P., & Boyle, F. M. (2020). A systematic literature review of health consumer attitudes towards secondary use and sharing of health administrative and clinical trial data: A focus on privacy, trust, and transparency. *Systematic Reviews*, 9(1). <https://doi.org/10.1186/s13643-020-01481-9>
- Jamal, L., Sapp, J. C., Lewis, K., Yanes, T., Facio, F. M., Biesecker, L. G., & Biesecker, B. B. (2014). Research participants' attitudes towards the confidentiality of genomic sequence information. *European Journal of Human Genetics*, 22, 964-968. <https://doi.org/10.1038/ejhg.2013.276>
- Jenkner, C. S., Ravi, N., Gabel, M., & Vogt, J. C. (2022). Trust in data-requesting organizations—A quantitative analysis on cultural antecedents and individual-level

- perceptions. *The Electronic Journal of Information Systems in Developing Countries*, e12208.
- Jiang, C. X., Chua, R. Y., Kotabe, M., & Murray, J. Y. (2011). Effects of cultural ethnicity, firm size, and firm age on senior executives trust in their overseas business partners: Evidence from China. *Journal of International Business Studies*, 42(9), 1150-1173. <https://doi.org/10.1057/jibs.2011.35>
- Jones, R. D., Krenz, C., Gornick, M., Griffith, K. A., Spence, R., Bradbury, A. R., De Vries, R., Hawley, S. T., Hayward, R. A., Zon, R., Bolte, S., Sadeghi, N., Schilsky, R. L., & Jagsi, R. (2020). Patient Preferences Regarding Informed Consent Models for Participation in a Learning Health Care System for Oncology. *JCO Oncology Practice*, 16, e977-e990. <https://doi.org/10.1200/jop.19.00300>
- Jones, R. D., Sabolch, A. N., Aakhus, E., Spence, R. A., Bradbury, A. R., & Jagsi, R. (2017). Patient perspectives on the ethical implementation of a rapid learning system for oncology care. *Journal of Oncology Practice*, 13, e163-e175. <https://doi.org/10.1200/JOP.2016.016782>
- Juga, J., Juntunen, J., & Koivumäki, T. (2021). Willingness to share personal health information: impact of attitudes, trust and control. *Records Management Journal*. <https://doi.org/10.1108/RMJ-02-2020-0005>
- Karwatzki, S., Trenz, M., & Veit, D. (2022). The multidimensional nature of privacy risks: Conceptualisation, measurement and implications for digital services. *Information Systems Journal*, 32(6), 1126-1157.
- Kerns, J. W., Krist, A. H., Longo, D. R., Kuzel, A. J., & Woolf, S. H. (2013). How patients want to engage with their personal health record: A qualitative study. *BMJ Open*, 3, 1-10. <https://doi.org/10.1136/bmjopen-2013-002931>
- Kettis-Lindblad, Å., Ring, L., Vibberth, E., & Hansson, M. G. (2006). Genetic research and donation of tissue samples to biobanks. What do potential sample donors in the Swedish general public think? *European Journal of Public Health*, 16, 433-440. <https://doi.org/10.1093/eurpub/cki198>
- Kim, D. J., Ferrin, D. L., Raghav Rao, H., Raghav, H., Kong Chian, L., & Raghav Rao, H. (2008). A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. 44(2), 544-564. <https://doi.org/10.1016/j.dss.2007.07.001>
- Korsgaard, M., Kautz, J., Bliese, P., Samson, K., & Kostyszyn, P. (2018). Conceptualising time as a level of analysis: New directions in the analysis of trust dynamics. *Journal of Trust Research*, 8(2), 142-165. <https://doi.org/10.1080/21515581.2018.1516557>
- Kramer, R. M. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. 50, 569-598. [www.annualreviews.org](http://www.annualreviews.org)
- Lee, S. B., Zak, A., Iversen, M. D., Polletta, V. L., Shadick, N. A., & Solomon, D. H. (2016). Participation in Clinical Research Registries: A Focus Group Study Examining Views From Patients With Arthritis and Other Chronic Illnesses. *Arthritis Care and Research*, 68, 974-980. <https://doi.org/10.1002/acr.22767>
- Lemke, A. A., Wolf, W. A., Hebert-Beirne, J., & Smith, M. E. (2010). Public and biobank participant attitudes toward genetic research participation and data sharing. *Public Health Genomics*, 13, 368-377. <https://doi.org/10.1159/000276767>
- Lewicki, R. J., McAllister, D. J., & Bies, R. I. (1998). Trust and distrust: New relationships and realities. *Academy of Management Review*, 23, 438-458. <https://doi.org/10.5465/AMR.1998.926620>
- Li, H., Gupta, A., Zhang, J., & Sarathy, R. (2014). Examining the decision to use standalone personal health record systems as a trust-enabled fair social contract. *Decision Support Systems*, 57, 376-386. <https://doi.org/10.1016/j.dss.2012.10.043>



- Lin, J., Carter, L., & Liu, D. (2021). Privacy concerns and digital government: exploring citizen willingness to adopt the COVIDSafe app. *European Journal of Information Systems*, 30(4), 389-402.
- Lind, E. A. (2018). Trust and fairness. In *The Routledge Companion to Trust* (pp. 183-196). Routledge.
- Lor, M., & Bowers, B. J. (2018). Hmong Older Adults' Perceptions of Insider and Outsider Researchers: Does It Matter for Research Participation? *Nursing research*, 67, 222-230. <https://doi.org/10.1097/NNR.0000000000000277>
- Lucero, R. J., Kearney, J., Cortes, Y., Arcia, A., Appelbaum, P., Fernández, R. L., & Luchsinger, J. (2015). Benefits and Risks in Secondary Use of Digitized Clinical Data: Views of Community Members Living in a Predominantly Ethnic Minority Urban Neighborhood. *AJOB Empirical Bioethics*, 6, 12-22. <https://doi.org/10.1080/23294515.2014.949906>
- Lysaght, T., Ballantyne, A., Xafis, V., Ong, S., Schaefer, G. O., Ling, J. M. T., Newson, A. J., Khor, I. W., & Tai, E. S. (2020). "Who is watching the watchdog?": ethical perspectives of sharing health-related data for precision medicine in Singapore. *BMC Medical Ethics*, 21. <https://doi.org/10.1186/s12910-020-00561-8>
- Lyu, S. C., & Ferrin, D. L. (2018). Determinants, consequences and functions of interpersonal trust within organizations: What is the empirical evidence? In R. Searle, A. M. Nienaber, & S. Sitkin (Eds.), *The Routledge companion to trust* (pp. 65-104). Routledge <https://doi.org/10.4324/9781315745572>
- Mählmann, L., Schee Gen Halfmann, S., Von Wyl, A., & Brand, A. (2018). Attitudes towards personal genomics and sharing of genetic data among older swiss adults: A qualitative study. *Public Health Genomics*, 20, 293-306. <https://doi.org/10.1159/000486588>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *Academy of Management Review*, 20(3), 709-734. <https://doi.org/https://doi.org/10.5465/amr.1995.9508080335>
- McAllister, D. J. (1995). *Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations*.
- McEvily, B., & Tortoriello, M. (2011). Measuring trust in organisational research: Review and recommendations. *Journal of Trust Research*, 1(1), 23-63. <https://doi.org/10.1080/21515581.2011.552424>
- McEvily, B., Perrone, V., & Zaheer, A. (2003). Trust as an organizing principle. *Organization Science*, 14, 91-103. <https://doi.org/10.1287/orsc.14.1.91.12814>
- McGuire, A. L., Hamilton, J. A., Lunstroth, R., McCullough, L. B., & Goldman, A. (2008). DNA data sharing: research participants' perspectives. *Genetics in medicine : official journal of the American College of Medical Genetics*, 10, 46-53. <https://doi.org/10.1097/GIM.0b013e31815f1e00>
- McKnight, D. H., & Chervany, N. L. (2001). Trust and Distrust Definitions: One Bite at a Time Technology Use View project Defining the Digital Enterprise View project Trust and Distrust Definitions: One Bite at a Time. [https://doi.org/10.1007/3-540-45547-7\\_3](https://doi.org/10.1007/3-540-45547-7_3)
- McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, 13(3). <https://doi.org/10.1287/isre.13.3.334.81>
- McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial Trust Formation in New Organizational Relationships. *Academy of Management Review*, 23(3), 473-490. <https://doi.org/https://doi.org/10.5465/amr.1998.926622>
- Mello, M. M., Lieou, V., & Goodman, S. N. (2018). Clinical Trial Participants' Views of the Risks and Benefits of Data Sharing. *New England Journal of Medicine*, 378, 2202-2211. <https://doi.org/10.1056/nejmsa1713258>

- Merson, L., Phong, T. V., Nhan, L. N. T., Dung, N. T., Ngan, T. T. D., Kinh, N. V., Parker, M., & Bull, S. (2015). Trust, Respect, and Reciprocity: Informing Culturally Appropriate Data-Sharing Practice in Vietnam. *Journal of empirical research on human research ethics : JERHRE*, 10, 251-263. <https://doi.org/10.1177/1556264615592387>
- Middleton, A., Milne, R., Almarri, M. A., Anwer, S., Atutornu, J., Baranova, E. E., Bevan, P., Cerezo, M., Cong, Y., Critchley, C., Fernow, J., Goodhand, P., Hasan, Q., Hibino, A., Houeland, G., Howard, H. C., Hussain, S. Z., Malmgren, C. I., Izhevskaya, V. L., Jędrzejak, A., Jinhong, C., Kimura, M., Kleiderman, E., Leach, B., Liu, K., Mascalonzi, D., Mendes, Á., Minari, J., Wang, N., Nicol, D., Niemiec, E., Patch, C., Pollard, J., Prainsack, B., Rivière, M., Robarts, L., Roberts, J., Romano, V., Sheerah, H. A., Smith, J., Soulier, A., Steed, C., Stefànsdóttir, V., Tandré, C., Thorogood, A., Voigt, T. H., West, A. V., Yoshizawa, G., & Morley, K. I. (2020). Global Public Perceptions of Genomic Data Sharing: What Shapes the Willingness to Donate DNA and Health Data? *American Journal of Human Genetics*, 107, 743-752. <https://doi.org/10.1016/j.ajhg.2020.08.023>
- Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and ‘the dark side’ of AI. *European Journal of Information Systems*, 31(3), 257-268.
- Möllering, G. (2019). Putting a spotlight on the trustor in trust research. *Journal of Trust Research*, 9(2), 131-135.
- Morin, D., Tourigny, A., Pelletier, D., Robichaud, L., Mathieu, L., Vézina, A., Bonin, L., & Buteau, M. (2005). Seniors' views on the use of electronic health records. *Informatics in Primary Care*, 13, 125-133. <https://doi.org/10.14236/jhi.v13i2.589>
- Mozersky, J., Parsons, M., Walsh, H., Baldwin, K., McIntosh, T., & DuBois, J. M. (2020). Research Participant Views regarding Qualitative Data Sharing. *Ethics & human research*, 42, 13-27. <https://doi.org/10.1002/eahr.500044>
- Okoli, C., & Schabram, K. (2012). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *SSRN Electronic Journal*, 10. <https://doi.org/10.2139/ssrn.1954824>
- Overby, C. L., Maloney, K. A., Alestock, T. D., Chavez, J., Berman, D., Sharaf, R. M., Fitzgerald, T., Kim, E.-Y., Palmer, K., Shuldiner, A. R., & Mitchell, B. D. (2015). Prioritizing Approaches to Engage Community Members and Build Trust in Biobanks: A Survey of Attitudes and Opinions of Adults within Outpatient Practices at the University of Maryland. *Journal of Personalized Medicine*, 5, 264-279. <https://doi.org/10.3390/jpm5030264>
- Papoutsis, C., Reed, J. E., Marston, C., Lewis, R., Majeed, A., & Bell, D. (2015). Patient and public views about the security and privacy of Electronic Health Records (EHRs) in the UK: Results from a mixed methods study Healthcare Information Systems. *BMC Medical Informatics and Decision Making*, 15. <https://doi.org/10.1186/s12911-015-0202-2>
- Platt, J., & Kardia, S. (2015). Public trust in health information sharing: Implications for biobanking and electronic health record systems. *Journal of Personalized Medicine*, 5(1), 3-21. <https://doi.org/10.3390/jpm5010003>
- Platt, J., Jacobson, P. D., & Kardia, S. L. R. (2018). Public Trust in Health Information Sharing: A Measure of System Trust. *Health Services Research*, 53(2), 824-845. <https://doi.org/10.1111/1475-6773.12654>
- Platt, J., Raj, M., & Kardia, S. L. R. (2019). The public’s trust and information brokers in health care, public health and research. *Journal of Health Organization and Management*, 33(7-8), 929-948. <https://doi.org/10.1108/JHOM-11-2018-0332>



- Platt, J., Raj, M., Büyüktür, A. G., Trinidad, M. G., Olopade, O., Ackerman, M. S., & Kardia, S. (2019). Willingness to Participate in Health Information Networks with Diverse Data Use: Evaluating Public Perspectives. *EGEMS (Washington, DC)*, 7, 33. <https://doi.org/10.5334/egems.288>
- Pool, J., Akhlaghpour, S., & Fatehi, F. (2020). Towards a contextual theory of Mobile Health Data Protection (MHDP): A realist perspective. *International Journal of Medical Informatics*, 141, 104229.
- Pratap, A., Allred, R., Duffy, J., Rivera, D., Lee, H. S., Renn, B. N., & Areán, P. A. (2019). Contemporary Views of Research Participant Willingness to Participate and Share Digital Data in Biomedical Research. *JAMA network open*, 2, e1915717. <https://doi.org/10.1001/jamanetworkopen.2019.15717>
- Raj, M., de Vries, R., Nong, P., Kardia, S. L. R., & Platt, J. E. (2020). Do people have an ethical obligation to share their health information? Comparing narratives of altruism and health information sharing in a nationally representative sample. *PLoS ONE*, 15, 1-14. <https://doi.org/10.1371/journal.pone.0244767>
- Rho, M. J., Jang, K. S., Chung, K.-Y., & Choi, I. Y. (2015). Comparison of knowledge, attitudes, and trust for the use of personal health information in clinical research. *Multimedia Tools and Applications*, 74, 2391-2404. <https://doi.org/10.1007/s11042-013-1772-6>
- Saunders, M. N., Dietz, G., & Thornhill, A. (2014). Trust and distrust: Polar opposites, or independent but co-existing?. *Human Relations*, 67(6), 639-665.
- Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An Integrative Model of Organizational Trust: Past, Present, and Future. *Academy of Management Review*, 32(2), 344-354. <https://doi.org/https://doi.org/10.5465/amr.2007.24348410>
- Serrano, K. J., Yu, M., Riley, W. T., Patel, V., Hughes, P., Marchesini, K., & Atienza, A. A. (2016). Willingness to Exchange Health Information via Mobile Devices: Findings From a Population-Based Survey. *Annals of Family Medicine*, 14, 34-40. <https://doi.org/10.1370/afm.1888>
- Shen, N., Bernier, T., Sequeira, L., Strauss, J., Silver, M. P., Carter-Langford, A., & Wiljer, D. (2019). Understanding the patient privacy perspective on health information exchange: a systematic review. *International journal of medical informatics*, 125, 1-12.
- Shen, N., Sequeira, L., Silver, M. P., Carter-Langford, A., Strauss, J., & Wiljer, D. (2019). Patient Privacy Perspectives on Health Information Exchange in a Mental Health Context: Qualitative Study. *JMIR Mental Health*, 6, e13306. <https://doi.org/10.2196/13306>
- Sitkin, S. B., & Bijlsma-Frankema, K. M. (2018). Distrust. In *The Routledge companion to trust* (pp. 50-61). Routledge.
- Skatova, A., Shiells, K., & Boyd, A. (2019). Attitudes towards transactional data donation and linkage in a longitudinal population study: Evidence from the Avon Longitudinal Study of Parents and Children. *Wellcome Open Research*, 4, 1-16. <https://doi.org/10.12688/wellcomeopenres.15557.1>
- Slegers, C., Zion, D., Glass, D., Kelsall, H., Fritschi, L., Brown, N., & Loff, B. (2015). Why Do People Participate in Epidemiological Research? *Journal of Bioethical Inquiry*, 12, 227-237. <https://doi.org/10.1007/s11673-015-9611-2>
- Spencer, K., Sanders, C., Whitley, E. A., Lund, D., Kaye, J., & Dixon, W. G. (2016). Patient perspectives on sharing anonymized personal health data using a digital system for dynamic consent and research feedback: A qualitative study. *Journal of Medical Internet Research*, 18. <https://doi.org/10.2196/jmir.5011>

- Stablein, T., Hall, J. L., Pervis, C., & Anthony, D. L. (2015). Negotiating stigma in health care: disclosure and the role of electronic health records. *Health Sociology Review, 24*, 227-241. <https://doi.org/10.1080/14461242.2015.1078218>
- Sterckx, S., Rakic, V., Cockbain, J., & Borry, P. (2016). "You hoped we would sleep walk into accepting the collection of our data": controversies surrounding the UK care.data scheme and their wider relevance for biomedical research. *Medicine, Health Care and Philosophy, 19*, 177-190. <https://doi.org/10.1007/s11019-015-9661-6>
- Stockdale, J., Cassell, J., & Ford, E. (2018). "Giving something back": a systematic review and ethical enquiry into public views on the use of patient data for research in the United Kingdom and the Republic of Ireland. *Wellcome open research, 3*.
- Stone, M. A., Redsell, S. A., Ling, J. T., & Hay, A. D. (2005). Sharing patient data: Competing demands of privacy, trust and research in primary care. *British Journal of General Practice, 55*, 783-789.
- Taylor, L. A., Nong, P. A. I. G. E., & Platt, J. O. D. Y. N. (2023). Fifty years of trust research in health care: a synthetic review. *The Milbank Quarterly, 101*(1), 126-178.
- Teixeira, P. A., Gordon, P., Camhi, E., & Bakken, S. (2011). HIV patients' willingness to share personal health information electronically. *Patient Education and Counseling, 84*, e9. <https://doi.org/10.1016/j.pec.2010.07.013>
- Tomlinson, E. C., Schnackenberg, A. K., Dawley, D., & Ash, S. R. (2020). Revisiting the trustworthiness–trust relationship: exploring the differential predictors of cognition-and affect-based trust. *Journal of Organizational Behavior, 41*(6), 535-550. <https://doi.org/10.1002/job.2448>
- Trinidad, S. B., Fullerton, S. M., Bares, J. M., Jarvik, G. P., Larson, E. B., & Burke, W. (2010). Genomic research and wide data sharing: Views of prospective participants. *Genetics in Medicine, 12*, 486-495. <https://doi.org/10.1097/GIM.0b013e3181e38f9e>
- van Knippenberg, D. (2018). Reconsidering affect-based trust: A new research agenda. In *The Routledge companion to trust* (pp. 3-13). Routledge.
- Vanneste, B. S., Puranam, P., & Kretschmer, T. (2014). Trust over time in exchange relationships: Meta-analysis and theory. *Strategic Management Journal, 35*(12), 1891-1902.
- Weidman, J., Aurite, W., & Grossklags, J. (2019). On sharing intentions, and personal and interdependent privacy considerations for genetic data: A vignette study. *IEEE/ACM Transactions on Computational Biology and Bioinformatics, 16*, 1349-1361. <https://doi.org/10.1109/TCBB.2018.2854785>
- Weng, C., Hao, T., Friedman, C., & Hurdle, J. (2019). Crowdsourcing public opinion for sharing medical records for the advancement of science. *Studies in health technology and informatics, 264*, 1393-1397. <https://doi.org/10.3233/SHTI190456>
- White, K., & Peloza, J. (2009). Self-benefit versus other-benefit marketing appeals: Their effectiveness in generating charitable support. *Journal of Marketing, 73*(4), 109-124.
- Willison, D. J., Steeves, V., Charles, C., Schwartz, L., Ranford, J., Agarwal, G., Cheng, J., & Thabane, L. (2009). Consent for use of personal information for health research: Do people with potentially stigmatizing health conditions and the general public differ in their opinions? *BMC Medical Ethics, 10*. <https://doi.org/10.1186/1472-6939-10-10>
- Willison, D. J., Swinton, M., Schwartz, L., Abelson, J., Charles, C., Northrup, D., Cheng, J., & Thabane, L. (2008). Alternatives to project-specific consent for access to personal information for health research: Insights from a public dialogue. *BMC Medical Ethics, 9*, 706-712. <https://doi.org/10.1186/1472-6939-9-18>
- Wolff, J., Pauling, J., Keck, A., & Baumbach, J. (2020). The economic impact of artificial intelligence in health care: systematic review. *Journal of medical Internet research, 22*(2), e16866.

Wu, K. W., Huang, S. Y., Yen, D. C., & Popova, I. (2012). The effect of online privacy policy on consumer privacy concern and trust. *Computers in Human Behavior*, 28(3), 889-897. <https://doi.org/10.1016/j.chb.2011.12.008>

## Appendix A

Search terms used for the three databases and articles found.

Database	Scopus	Web of Science	Pubmed
Search terms	( TITLE-ABS-KEY ( ( data OR information OR record* ) W/4 ( shar* OR link* OR mining OR disclos* OR reus* ) ) AND TITLE-ABS-KEY ( ( health* OR medical OR clinical OR patient ) W/4 ( record* OR data OR information ) ) ) AND TITLE-ABS-KEY ( willing* OR intent* OR attitud* OR accept* OR view* OR opinion* OR perspective* ) AND TITLE-ABS-KEY ( empiric* OR sampl* OR survey* OR qualitat* OR quantita* OR interview* OR longitudinal OR "case study" OR "cross-sectional" OR "focus group*" ) AND TITLE-ABS-KEY ( *trust* ) )	((data OR information OR record*) NEAR/4 (shar* OR link* OR mining OR disclos* OR reus*)) AND TOPIC: ((health* OR medical OR clinical OR patient) NEAR/4 (record* OR data OR information)) AND TOPIC: (willing* OR intent* OR attitud* OR accept* OR view* OR opinion* OR perspective*) AND TOPIC: (*trust*) AND TOPIC: (empiric* OR sampl* OR survey* OR qualitat* OR quantita* OR interview* OR longitudinal OR "case study" OR "cross-sectional" OR "focus group*"). Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI	("health information" OR "health data" OR "clinical data" OR "clinical information" OR "medical record*" OR "clinical record*" OR "data sharing" OR "patient information" OR "patient record*") AND (shar* OR disclos* OR link* OR reus*) AND (willing* OR intent* OR attitud* OR accept* OR view* OR opinion* OR perspective*) AND (empiric* OR sampl* OR survey* OR qualitat* OR quantita* OR interview* OR longitudinal OR "case study" OR "cross-sectional" OR "focus group*") AND *trust*
# Articles found	366	236	379

Note: TITLE-ABS-KEY searches the terms in the title, abstract, or keywords of each article. The NEAR/ and W/ are adjacency operators, and the number notes how many words are allowed between the two terms or sets of terms. Asterisks include all truncations of a word and quotation marks require the quoted terms to be found verbatim. For Web of Science, the indexes refer to the searched sub-databases.

## Appendix B

Coding scheme of the elements of the analytical lens found in the collected articles.

Code	Subcode	Example
Ability	Having expertise/Possessing the necessary knowledge/Having had sufficient training	“I trust that they [NHS] have those skills to keep it anonymized” (Spencer et al., 2016)
	Being generally competent	“[...] you trust people not to be careless in terms of entering information, in terms of leaving perhaps your records on a screen just for other people to look at [...]” (Stablein et al., 2015)
Benevolence	Having the trustor’s best interest at heart	“It’s not [as] if they’re up to jack [rob] you or nothing” (Greenhalgh et al., 2008)
	Not primarily seeking profit	“The idea that many fertility clinics are businesses, with financial motivations, negatively impacted on the trust that patients felt” (Carson et al., 2019)
Integrity	Telling the truth	“[...] sometimes they confuse you, and there are many interpreters that sometimes don’t tell you the truth [...]” (Lee et al., 2016)
	Not withholding information	“some participants discussed that it is important to provide complete disclosure about genetic research to reduce fears and build trust” (Lemke et al., 2010)
	Adhering to morals important to the trustor	“[...] if I had like a tumor removed and they were using it for a biobank for further research and they’re using it to do this sort of stuff, it’s just against my religion” (De Vries et al., 2019)
Propensity	General trust towards other people	“In particular since we live in a country where things work pretty well, where there is no problem in that respect” (Bosisio et al., 2021)
Confidentiality	Keeping information protected	“I trust them that it wouldn’t get out of hand. So I feel pretty comfortable. I feel pretty comfortable and because I think it would be safe kept.” (Jones et al., 2017)
	Privacy concerns	“[...] reassurance...it would just be few hands and eyes ... so it's not being passed around” (Grant et al., 2013)
Time	Frequency of contact/ Relationship length/ Familiarity/Experience	“[T]he head pharmacist, he's been working with me for the last like 25, 30 years and I always refer him to the pharmacist at whatever hospital I'm at [...]” (Shen et al., 2019)

Culture	Differences between cultures	“The “Your DNA, Your Say” project is a very large social sciences study conducted on global public attitudes toward genomic data sharing. [...] The results show patterns of both consistency and diversity across the globe.” (Middleton et al., 2020)
	Differences by minority status	“[...] ethnicity (except for Caucasian) is shown to be a significant influence on most measured dependents” (Weidman et al., 2019)
Distrust	Distrust	“In contrast, two patients tended to distrust these unknown researchers and expressed concerns relating to misuse and security of their data” (Broes et al., 2020)
	Mistrust	“Overall, money causes people’s mistrust, they argue that privacy and data protection are not guaranteed, since the private institutions might sell or distribute data for financial gains.” (Mählmann et al., 2018)
Third party influence	Trust transfer/Trust by proxy	“[...] many indicated they would trust the governmental authorities and branding of such would be a factor for trusting the system” (Alaqla et al., 2018)

# Project Materiality and Integrated Analytics in Sustainability Augmented Portfolio Management

Javed Mohammad, Laden Husamaldin and Ikram Ur Rehman

School of Computing and Engineering, University of West London, Ealing, London U.K.

Email: [21434312@student.uwl.ac.uk](mailto:21434312@student.uwl.ac.uk) , [laden.aldin@uwl.ac.uk](mailto:laden.aldin@uwl.ac.uk),

[ikram.rehman@uwl.ac.uk](mailto:ikram.rehman@uwl.ac.uk)

*Completed Research*

## **Abstract**

*The integration of sustainability within project environments has emerged as a central concern in contemporary project management literature and practice. This research paper examines the critical issue of sustainability integration and its impact on project success. Inadequate contextual knowledge often leads to ineffectual strategies, unsustainable outcomes, and unsuccessful projects, giving rise to practices like greenwashing and regulatory non-compliance. As the project environment (PE) transforms, sustainability is poised to become an enduring feature of project management (PM). This research closely examines tactical versus strategic, product versus project lifecycle approach, and PM versus project portfolio management (PPM) for a sustainable PE. Thus, this research paper introduces 'Sustainability Augmented Portfolio Management' (SAPoM) as a strategic approach leveraging Project Materiality assessment derived from sustainability reporting standards to create a sustainability-driven PE and enhanced PPM decision-making through integrated analytics. The SAPoM-based conceptual framework proposed here aligns projects with organisational objectives and sustainability principles, offering a strategic solution for the complex challenge of integrating sustainability within PEs.*

**Keywords:** Sustainability, Project portfolio management, Analytics, Project Materiality, Sustainability Reporting Standards

## 1.0 Introduction

Across the globe, international entities, governments, and industry organisations are playing an active role in shaping the landscape of sustainability reporting. Their collective aim is to encourage transparency and accountability in the way businesses conduct their activities. This movement towards sustainable reporting is not occurring in isolation; instead, it draws inspiration and borrows key concepts from financial reporting standards (Elliott and Elliott, 2022). Much like financial reports provide a structured and standardised way to communicate an organisation's financial health and performance, sustainability reporting is emerging as a similar framework that discloses an organisation's commitment to, and impact on the environment, economy and society.

The EU's Corporate Sustainability Reporting Directive (CSRD) and the UK's Sustainability Disclosure Requirements (SDR) are notable examples of regulations that are on the horizon (European Commission, 2023; Financial Conduct Authority (FCA), 2022). Compliance with these regulations is, as a matter of course, necessary for maintaining competitiveness and financial stability (Turan *et al.*, 2008). In a recent development, the Global Reporting Initiative (GRI) and the European Financial Reporting Advisory Group (EFRAG) now offer interoperable and complementing standards (EFRAG, 2023). This transformation will lead to a lasting reconfiguration of corporate strategy and operations, with a primary emphasis on prompting business leaders to recognise the significance of incorporating sustainability into the project environment (PE).

In contemporary project management (PM) research, sustainability is frequently seen as a comprehensive and context-sensitive notion. This poses challenges in developing universally applicable knowledge or approaches for its integration - a problem rooted in the hermeneutics of sustainability. Researchers struggle to define the boundaries of what constitutes sustainability and sustainable practices within project management. Existing research often views sustainable PM from a project-oriented standpoint, limiting sustainability integration. This dilemma leaves scholars and practitioners contemplating whether PM or project portfolio management (PPM) is more effective for enduring sustainability outcomes.

This research is grounded on the premise that materiality assessment derived from sustainability reporting standards provides a platform for smoothly integrating sustainability into the project environment creating the capacity and capability to remain in step with the evolving nature of sustainability while delivering an organisation's strategic objectives. This research paper



critically evaluates both tactical and strategic approaches, emphasising the importance of considering project outcomes from a broader product lifecycle perspective. Moreover, it incorporates analytics to support informed decision-making throughout the process. This holistic approach will ensure the delivery of oven-ready outcomes for disclosure and regulatory scrutiny.

The research paper is structured into five sections, commencing with this introduction following the abstract. The second section, Background, comprises two parts, exploring extant literature and its limitations relevant integration of sustainability within the project environment, and project materiality. The third section explains the methodology adopted for this research, while the fourth section extensively addresses the proposed conceptual framework of SAPoM, encompassing sustainability reporting standards, project materiality, the integration of analytics and stakeholder engagement with SAPoM. The fifth and final section examines the research's impacts and serves as the conclusion of the paper.

## 2.0 Background

### 2.1 The Evolving Notion of Sustainability

Sustainability is a versatile but sometimes contentious concept, drawing scepticism from both researchers and practitioners due to its lexical ambiguity (Sze, 2018). At its core, sustainability involves maintaining, supporting, enduring and enhancing the well-being of humanity within the constraints of available resources. While this ambiguity might risk rendering sustainability meaningless (Silvius, 2017), it has also broadened its application, inspiring more actions in its name in the absence of a rigid definition (Cadenasso and Pickett, 2018).

It is almost inconceivable to write about sustainability without mentioning the Brundtland Report by the World Commission on Environment and Development (WCED) which established a universally accepted definition underscoring the inseparable link between environment and development. It emphasises the imperative of ‘meeting the needs of the present generation without compromising the ability of future generations to meet their own needs’ (Brundtland, 1987, p.16). However, this foundational definition has spawned a multitude of interpretations in various contexts, spanning ethics, philosophy, strategy, climate change, and more. In 2012, Kilbert *et al.* documented at least seventy sustainability definitions, a number that surged to over three hundred by 2019 (Kilbert *et al.*, 2012; Per, 2019).

Sustainability embodies the Triple Bottom Line (TBL) principles of economic performance, environmental protection, and social equity (Elkington, 1997, 1999; Hubbard, 2009; Høgevold *et al.*, 2015; Padin *et al.*, 2016; Ghannadpour *et al.*, 2021). Taking a hermeneutical perspective of interpreting sustainability entails incorporating the Triple Bottom Line (TBL) principles collectively in research, rather than treating them as separate and isolated principles (Padin *et al.*, 2016; Al-Marri and Pinnington, 2022) While TBL aids in operationalising sustainability (Marcelino-Sadaba *et al.*, 2015; Silvius, 2017), encouraging organisations to assess and manage their performance holistically across these three dimensions. It helps organisations recognise that long-term success requires not only financial viability but also consideration of environmental and social impacts.

Measuring, reporting, and acting on economic, environmental, and social indicators, helps organisations to make informed decisions that contribute to sustainability and responsible business practices. From project management (PM) integration perspective, materiality assessment (Wu, Shao and Chen, 2018; Sardianou *et al.*, 2021; To and Chau, 2022), industry-

relevant indicators and criteria (Kuzemko and Britton, 2020), global applicability, qualitative and quantitative insights should be integral to the process. As it can enhance the necessary changes for achieving sustainable development, therefore, bridging the gap between the two domains.

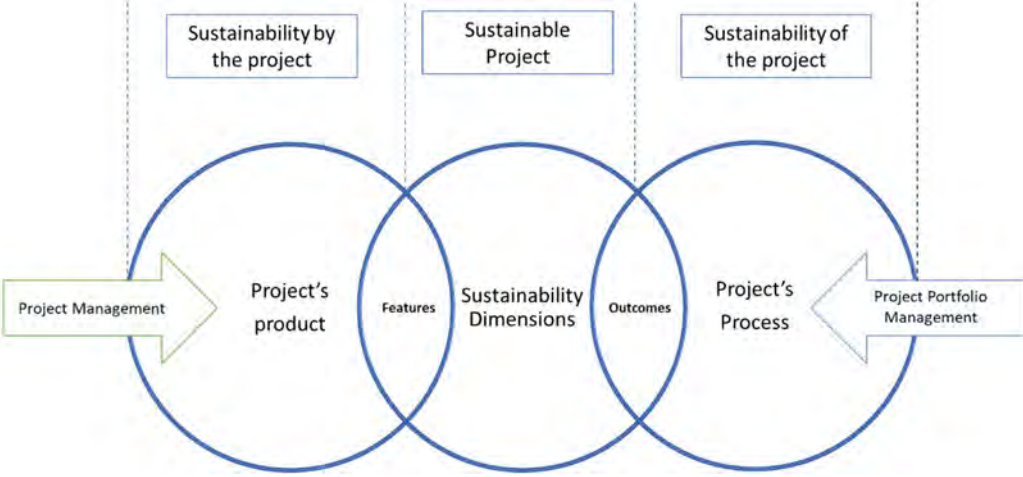
## **2.2 Sustainability at the Crossroads of Project and Portfolio Management**

This research paper examines projects by focusing on two domains: project portfolio management (PPM) and project management (PM). The primary aim is to acquire a comprehensive insight into the rationale for incorporating sustainability into the project environment. The term ‘project environment’ is central to this discussion, encompassing several elements, including project teams, stakeholders, resources, constraints, risks, governance, interdependencies, external environment, culture, policies, and organisational support.

Project management (PM) as defined by the Project Management Institute, (2021) is a temporary endeavour that entails the application of knowledge, skills, tools, and techniques to meet project requirements within the defined constraints of time, cost and scope or quality (Martinsuo and Lehtonen, 2007). Hence, PM is primarily centred on the product itself (Pensenstadler and Femmer, 2013; Doorasamy, 2017), concentrating on project objectives and ‘sustainability by the project’ (Huemann and Silvius, 2017). Project practitioners recognise that once these objectives are met, the project serves no further purpose. The Association of Project Management underscores that tactical projects should only be undertaken if they align with strategic objectives (APM Portfolio Management SIG, 2019). Sustainability often involves much broader, ethical and long-term questions about responsible resource use, equity, and the moral obligations of organisations. These have limited relevance to the project’s immediate goals. Achieving these aims is contingent upon projects operating within the more strategic domain of project portfolio management (PPM).

Project portfolio management (PPM), as defined by the Project Management Institute, (2021) is the organisation of a collection of projects and programs aimed at achieving strategic objectives (Clegg et al., 2018). Cooper et al., (2001) underscores the functional role of PPM in selecting, controlling, prioritising, and monitoring projects to align resource allocation and managing risks to gain strategic advantage. Crucially, PPM's process-centric and ‘sustainability of the project’ approach (Huemann and Silvius, 2017) transcends individual project lifecycles and PM, emphasising a broader strategic outlook. Additionally, a notable research gap exists regarding the utilisation of PPM as a process that can effectively be applied across the entire

product lifecycle. Such research would be pivotal in the seamless integration of sustainability into the project environment as presented in Figure 1.



**Figure 1. Product versus Process sustainability in the project environment**

Sustainability adopts a strategic perspective, with far-reaching consequences spanning environmental, economic, and social domains (Silvius and Marnewick, 2022). Sustainability goes beyond individual project timelines, requiring alignment with the organisation's overarching sustainability strategy. If conventional project management falls short in addressing sustainability due to its temporal constraints, inflexibility, and limited impact on deliverables, it becomes imperative to seek a more strategic solution free from these confines. Enter Sustainability Augmented Portfolio Management or SAPoM.

Before this research paper describes SAPoM as a solution for integrating sustainability in the project environment (PE), it would like to present Project Materiality as a key element that bridges the gap between sustainability and portfolio management.

**2.3 Project Materiality: The Key to Sustainability and Portfolio Management Integration**

In sustainability-related academic literature, there exists a substantial body of research dedicated to exploring the concept of materiality. Materiality is widely acknowledged as a pivotal element in sustainability, intimately connected with the norms and guidelines governing sustainability reporting. In traditional project portfolio management, it is a common practice to

define a set of decision criteria and attributes for screening, prioritising, selecting, and discontinuing projects within a portfolio, while also ensuring their alignment with the organisation's strategic goals (Bible and Bivins, 2011; Petit, 2012; Lister, 2015; Moustafaev, 2017a; Project-management.pm, 2017; Clegg *et al.*, 2018; Martinsuo and Geraldi, 2020). Diverse project evaluation techniques align with organisational strategies. Senior executives promote sustainability in goal setting. Resources are allocated judiciously to meet sustainability prerequisites, ensuring transparency and accountability in reporting for Corporate Social Responsibility (CSR) (Delbard, 2008; Ashrafi *et al.*, 2018; Paun, 2018; Christensen, Hail and Luez, 2019; Bahu, 2020; Al-Marri and Pinnington, 2022). A significant research gap exists in the effort to introduce materiality as a comprehensive substitute for the traditional decision criteria approach.

This research paper introduces the concept of project materiality as a means to incorporate sustainability into project portfolio management. Conventional decision-making criteria, in the context of sustainability integration, do not adequately account for social and environmental factors, market dynamics, risk evaluation, financial viability, stakeholder implications, and alignment with a sustainability-focused approach. This paper, therefore, conceptualises project materiality as integral to sustainable project portfolio management. As project materiality comprehensively covers economic, environmental and social aspects, it replaces the traditional decision criteria process currently practised.

Materiality is a concept borrowed by the Global Reporting Initiative (GRI) and other sustainability reporting standards from financial reporting (Wu, Shao and Chen, 2018; Kuzemko and Britton, 2020). It reflects the significant economic, environmental, and social impacts of a company and how these affect the assessments and decisions of stakeholders. Distinguishing materiality within the context of financial reporting from materiality in sustainability reporting is essential for comprehending the rationale behind its adoption in sustainability reporting, drawing on principles borrowed from the financial realm.

This differentiation is necessary as it sheds light on why materiality concepts, traditionally associated with financial matters, have been integrated into sustainability reporting. It enables the appreciation of underlying motivations and benefits of incorporating financial reporting practices into sustainability frameworks. Materiality, a concept deeply rooted in financial contexts, has been repurposed to better address the multifaceted environmental, social, and

economic concerns inherent to sustainability, emphasising the importance of understanding this transition in reporting approaches.

In financial reporting, information is considered material if its absence or misrepresentation could impact the financial decisions made by users relying on financial statements, as outlined by the IASB Framework (International Accounting Standards Board (IASB), 2018). Conversely, in the context of sustainability, materiality pertains to issues that can exert substantial effects on a company, encompassing both advantageous and detrimental outcomes (Rifkin, 2019). Materiality is fundamental to sustainability, facilitating impartial decision-making and the pursuit of ethical and genuine actions (To and Chau, 2022).

Additionally, in the context of Sustainability Augmented Portfolio Management (SAPoM), it delineates the issues that are significant to an organisation while assessing the complexity and subjectivity of sustainability within project portfolio management (PPM). It also raises questions about the consistency and comparability of sustainability within projects within the portfolio, as different stakeholders may interpret materiality differently. This lack of standardisation is a significant issue for sustainability integration and requires ongoing critical consideration by stakeholders.

For this research, project materiality is defined as:

*Project materiality refers to a set of factors within the project environment (PE) that have substantial influence in terms of strategic, economic, environmental, and social consequences. These factors significantly influence project and portfolio management and shape stakeholders' determinations concerning an organisation's strategic direction and its capacity to withstand sustainability-resilience challenges and uncertainties.*

Sustainability resilience from a PE perspective refers to the project's ability to withstand and adapt to various sustainability challenges and changes over time while still achieving its long-term sustainable objectives (Espiner, Orchiston and Higham, 2017; Nüchter *et al.*, 2021). The importance of sustainability-resilience within the Project Environment (PE) focuses on the lifecycle of deliverables and not just the project itself. Project practitioners are, therefore, required to revisit and evaluate materiality factors continually in the face of environmental, regulatory and strategic shifts, anticipate and mitigate risks, and account for ethical and social factors. Sustainability-resilience should be seen as a paradigm shift in project and portfolio management.

Further, for this research, Project Materiality Assessment is defined as:

*Project materiality assessment is the systematic evaluation and analysis of the factors within the project environment (PE) that possess substantial influence concerning strategic, economic, environmental, and social outcomes. This assessment aims to comprehensively understand and quantify how these factors impact projects, programmes and portfolios, stakeholders' determinations concerning an organisation's strategic orientation and its ability to navigate sustainability-resilience challenges and uncertainties.*

#### **2.4 Highlighting the Challenges and the Limitations**

In this research paper, the challenge encompasses three main dimensions. Firstly, it involves the integration of sustainability principles into the project environment, requiring organisations to align and adapt their project management practices with sustainability objectives. Secondly, it necessitates the application of analytics to enhance decision-making processes, bridging the gap between sustainability and project portfolio management (PPM). This aims to develop advanced prediction models, data-driven decision-making, and the delivery of sustainable project outcomes. Finally, organisations are confronted with the task of restructuring PPM practices within their operational frameworks.

These challenges are further compounded by the emergence of new and forthcoming regulatory requirements mandating organisations to disclose their activities and their environmental and social impacts. This, in turn, requires organisations to innovate and shift their existing practices towards sustainability principles right from the project's inception. PPM must evolve to become data-driven and focused on delivering sustainable outcomes through projects. Therefore, this paper seeks to examine these challenges, and their implications, and propose strategies for organisations to effectively respond to the demands for sustainable project outcomes while ensuring compliance with evolving sustainability standards and regulations.

Finally, summarising the limitations within the existing literature and practice can be delineated as follows:

- The current body of research fails to adequately address the incorporation of analytics within the realm of sustainable project portfolio management (PPM).

- Ambiguities persist regarding the most effective means of achieving sustainable outcomes, whether through the product-oriented approach of Project Management (PM) or the process-oriented approach of Project Portfolio Management (PPM). This necessitates further investigation.
- The translation and measurement of sustainability-related quantitative data (including economic and environmental aspects) and qualitative data (of social factors) concerning an organisation's short and long-term strategic objectives require clarification and exploration.
- Bridging the gap between sustainability reporting tools, such as the Global Reporting Initiative (GRI), and strategic business processes to facilitate practitioners in the quantitative and qualitative evaluation of sustainability remains a challenge. The absence of any materiality assessment from a project environment (PE) perspective creates further complexities. The question arises as to how such an analytics-based framework can enhance reporting accuracy and support more effective sustainability assessments.

### **3.0 Methodology**

The literature search for the four interconnected concepts of sustainability, project portfolio management, materiality and analytics was conducted using a systematic approach to ensure comprehensive coverage of relevant research. After clearly defining the limitations of the current research, the following steps were followed:

1. Identifying Keywords and Search Terms: A set of keywords and search terms related to sustainability, project portfolio management, and analytics were identified. In the first step keywords were chosen to capture the core concepts and their various aspects. Next, literature with the concepts combined were selected e.g., sustainable project management, sustainable project portfolio management (PPM), sustainable project reporting, analytics in sustainable project decision-making, materiality in financial sector and sustainability reporting, stakeholder engagement in sustainable project management, sustainability reporting and project management, sustainability reporting and project portfolio management.
2. Selecting Databases and Resources: Multiple academic databases, such as EBSCO, Science Direct, ProQuest, JSTOR and Google Scholar, were selected to conduct the



literature search. Additionally, industry-specific platforms, organisational websites, and conference proceedings were considered for relevant literature.

3. Applying Inclusion and Exclusion Criteria: Inclusion and exclusion criteria were established to ensure that only the most relevant and high-quality sources were included in the search results. The criteria included publication date, language, study type, and relevance to the research questions. The aim was to find literature that included all four areas of study; however, as can be seen in Figure 2 below, out of 209 articles searched, only three articles covered sustainable project management and analytics while no article was found that covered the application of analytics in sustainable project portfolio management.
4. Conducting the Search: The identified keywords and search terms as seen in Figure 2 were used to query the selected databases and resources. The search was conducted systematically and consistently to retrieve relevant articles and publications. Other affiliated concepts which included definitions, best practices, business maturity models, business processes, regulatory consultations, web-based organisational resources, stakeholder management, and knowledge management were also studied for a more thorough and inclusive research.

Following is a breakdown of the research articles studied:

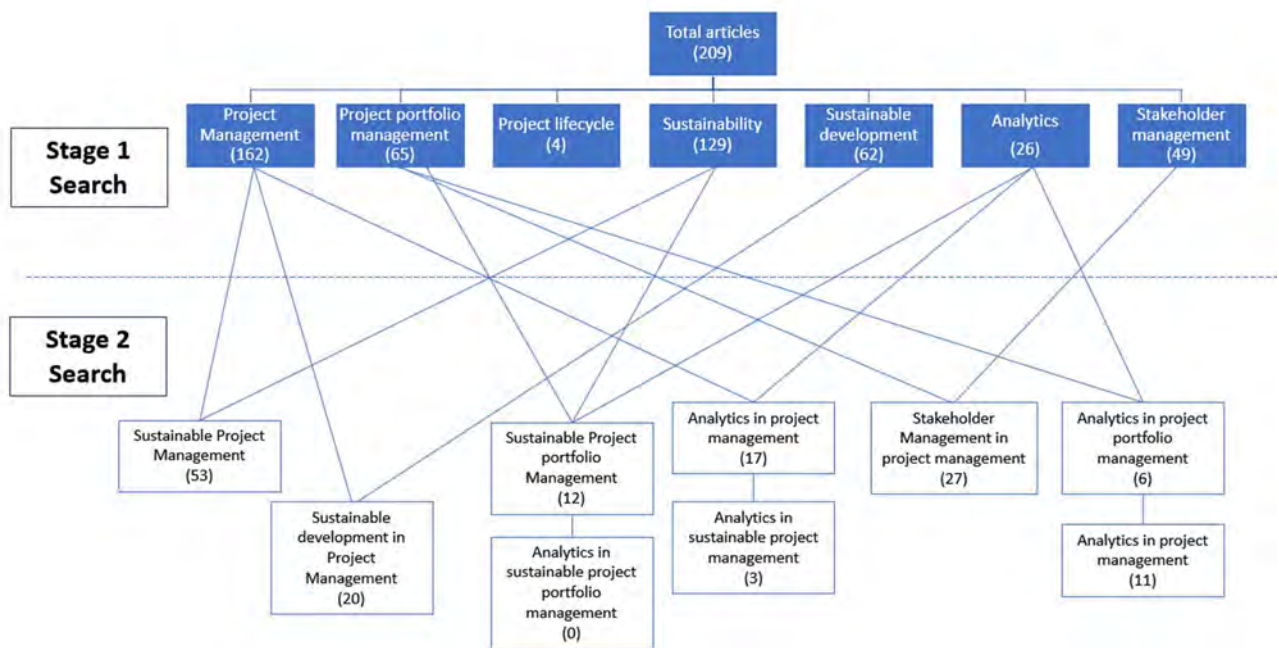


Figure 2. Article search tree map

5. **Screening Titles and Abstracts:** The retrieved articles' titles and abstracts were screened to determine their potential relevance to the research questions. Core concepts have been derived from the stage 1 searches, while blended and integrated concepts were part of the stage 2 searches as depicted in Figure 2 above. Stage 2 search applied Boolean operators (AND, OR, NOT) to combine terms and concepts effectively while irrelevant or unrelated sources were excluded at this stage. Sources with a more universal acceptance were preferred to those with regional scope.
6. **Reviewing Full-Text Articles:** The full-text articles of the selected studies were reviewed to assess their suitability for inclusion in the literature review. Articles that met the inclusion criteria were considered for analysis.
7. **Data Extraction and Synthesis:** Data from the selected articles were extracted and organised to identify key findings, themes, and interconnections between sustainability, sustainability reporting and analysis, project management, project portfolio management, and analytics. As the research aims to integrate different concepts, the synthesis process involves examining the relationships and interactions between these concepts to gain a corroborative and comprehensive understanding.
8. **Analysing and Reporting the Findings:** The extracted data and synthesised findings were analysed to draw meaningful conclusions and insights. The results have been presented clearly and coherently in the literature review, discussing the interrelationships and implications of interconnecting concepts, which concepts have more synergies while gaps in the literature have also been identified.

By following this systematic process, the literature search aimed to ensure a comprehensive exploration of the connections between sustainability, project portfolio management, and analytics while maintaining rigour and transparency in the review process.

## **4.0 Sustainability Augmented Portfolio Management (SAPoM)**

### **4.1 Conceptual Framework**

The fundamental considerations that underpin this research are:

- Organisations can now face rigorous assessments of their operational impact and their commitment to sustainability standards within defined reporting periods (Courtneil, 2019; Global Sustainability Standards Board, 2020; Global Reporting Initiative, 2021).
- Leveraging existing sustainability reporting standards eliminates the need for unnecessary reinvention. These industry-specific, sustainability-driven guidelines can be effortlessly integrated into the project environment (PE), streamlining the processes.
- The continuous evolution of reporting standards underscores their ability to stay in step with upcoming regulatory changes. This adaptability is essential as it allows these standards to not only keep up with but also proactively respond to the shifting regulatory landscape. They remain effective and relevant by adjusting and aligning with the changing requirements, ensuring that organisations can meet their reporting obligations in a dynamic and ever-changing business environment. This evolution serves as a proactive strategy to keep reporting practices in tune with the demands of the present and future regulatory frameworks.
- Project practitioners can be assured that they are always up to date when carrying out project materiality assessments of their projects and portfolios using these standards.

In short, the reporting standards employed by businesses have the power to drive sustainable results within the project environment. Consequently, SAPoM deliverables can be designated as products and services poised for sustainability, ready to make a meaningful impact.

### **4.2 SAPoM: The Imperative For An Integrated Approach**

SAPoM is a composite framework that aims to integrate analytics and project materiality into project portfolio management (PPM) planning and processes with minimal disruption. The conceptual framework is depicted in Figure 3. Similarly, the feedback can be fed back for further analysis to build best practices and maturity in planning and processes. The building blocks of this composite framework are stated below:

- **Integrated Analytics and Project Materiality**
  - Step 1: Sustainability criteria and metrics derived from sustainability reporting standards e.g., Global Reporting Initiative (GRI) or Sustainability Accounting Standards Board (SASB)
  - Step 2: Project Materiality Assessment
  - Step 3: Analytical tools to support insights to support materiality assessment and decision-making in processes within the PPM framework
- **Project portfolio management (PPM)**

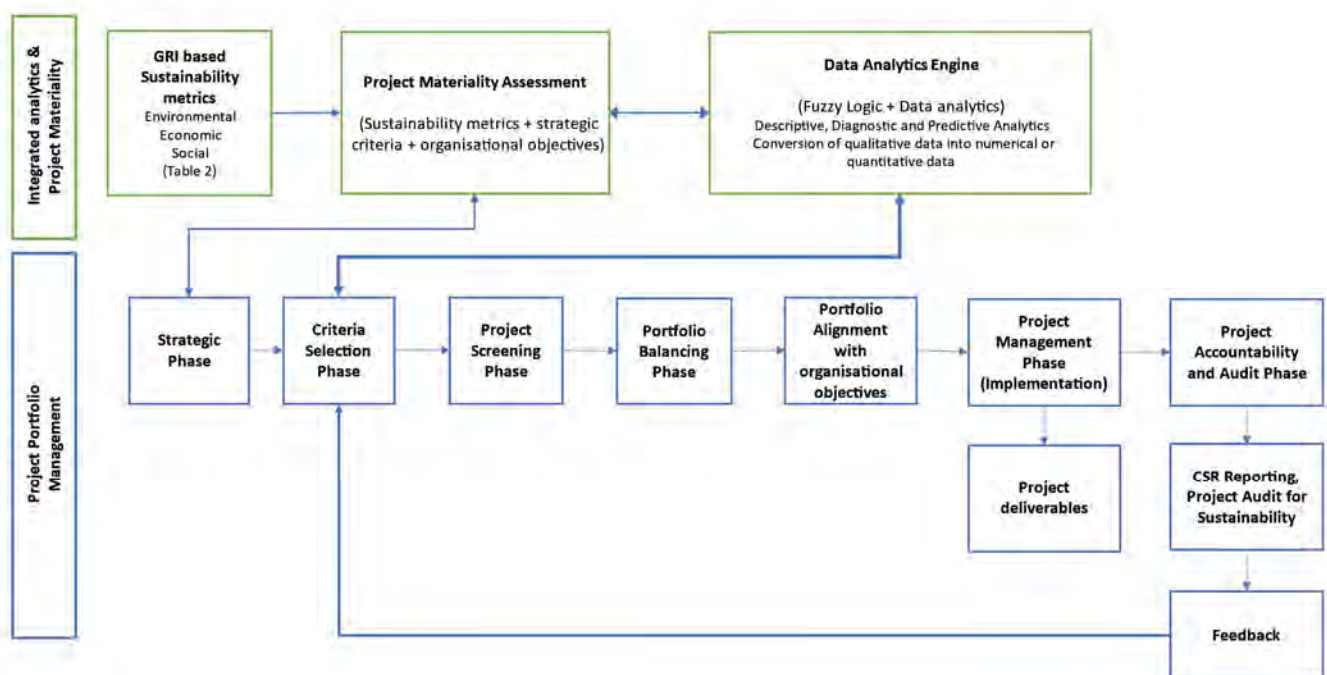


Figure 3. Conceptual Framework - Sustainability Augmented Portfolio Management (SAPoM)

## Integrated Analytics and Project Materiality

### Step 1: Employing Sustainability Reporting Standards into SAPoM

Sustainability criteria within reporting standards do not directly address the challenges of sustainability in the context of projects. These standards lack the incorporation of analytics for informed decision-making, which is essential. Regulatory frameworks primarily focus on reporting and labelling to meet compliance requirements. Noteworthy among them are:

- Global Reporting Initiative (GRI) (Dumay, Guthrie and Farneti, 2010; Courtnell, 2019; Global Reporting Initiative, 2021)
- Carbon Disclosure Project (CDP) (Carbon Disclosure Project, 2023)
- Task Force on Climate-related Financial Disclosures (Task Force on Climate-related Financial Disclosures (TCFD), 2022)
- Sustainability Accounting Standards Board (SASB) (Sustainability Accounting Standards Board, 2023)
- Integrated Reporting Framework by the International Sustainability Standards Board (ISSB)
- ESG Scoring (Tarmuji, Maelah and Tarmuji, 2016)

The challenge lies in bridging the gap between sustainability criteria and PPM, ensuring measurability, clarity, stakeholder engagement, and transparency without compromising project integrity. This calls for expertise in statistical analysis, strategic management, ethical sustainability practices, and data management to effectively apply these criteria within project frameworks. Reporting standards provide guidelines and standards for sustainability reporting, helping organisations assess and communicate their sustainability performance. Some of the key features of GRI reporting (Dumay, Guthrie and Farneti, 2010; Courtnell, 2019; Global Sustainability Standards Board, 2020) are:

- **Materiality Assessment:** This involves identifying the sustainability topics and issues that are most significant for the organisation and its stakeholders. Materiality helps organisations focus their reporting on what truly matters in terms of economic, environmental, and social impacts.
- **Reporting Principles:** These principles include transparency, materiality, completeness, and stakeholder inclusiveness. They emphasise the importance of reporting relevant, accurate, and comprehensive sustainability information.
- **Reporting Frameworks:** Sustainability reporting standards have developed several reporting frameworks that organisations can use to structure their sustainability reports. The most widely used is the GRI Standards, which cover a range of sustainability topics such as governance, ethics, labour practices, environmental impact, and societal contributions.

- **Indicators and Disclosures:** These include specific indicators and disclosures that organisations can use to measure and report on their sustainability performance. These indicators are organised by topic, making it easier for organisations to address the most relevant sustainability issues.
- **Stakeholder Engagement:** GRI emphasises the importance of engaging with stakeholders throughout the reporting process. Organisations are encouraged to identify their stakeholders, understand their concerns, and incorporate their feedback into sustainability reporting.
- **Integration with Other Standards:** GRI recognises that organisations may use other sustainability and reporting standards. The GRI framework is designed to be compatible with other standards, such as the United Nations Sustainable Development Goals (SDGs) (United Nations, 2001, 2015; Tsalis *et al.*, 2020) and the principles of the International Integrated Reporting Council (IIRC).
- **Assurance and Verification:** GRI encourages organisations to seek external assurance or verification of their sustainability reports. This adds credibility to the reported information and ensures that it has been independently assessed for accuracy and completeness.
- **Continuous Improvement:** GRI promotes a process of continuous improvement in sustainability reporting. Organisations are encouraged to set targets, track progress, and update their reports regularly to demonstrate their commitment to sustainability.
- **Global Applicability:** GRI's framework is designed to apply to organisations of all sizes and sectors, making it a globally recognised standard for sustainability reporting.

GRI examines sustainability by providing a structured framework for organisations to assess and report on their economic, environmental, and social impacts. It emphasises the importance of materiality, stakeholder engagement, transparency, and continuous improvement in the reporting process, making it a widely used and respected tool in the field of sustainability reporting. These standards can, however, be seamlessly integrated into project portfolio

management ensuring that projects align with sustainability tenets, regulatory requirements, and an organisation's sustainability strategy from the outset.

**Step 2: Deriving Project Materiality Criteria from Sustainability Reporting Standards**

Literature reveals that the concept of materiality within sustainability draws its origins from the principles of financial materiality. This paper introduces the same influence within the domain of the project environment (PE) and more specifically Sustainability Augmented Portfolio Management (SAPoM). To gain a comparative understanding of the intricate evolutions in these contexts, it is imperative to present a comparative analysis illustrated in Table 1.

For this research paper, materiality topics from the Global Reporting Initiative (GRI) guidelines have been used as the basis for the study. This process involves systematically identifying the most significant sustainability topics for an organisation, the portfolio and the projects. Project Materiality topics are those issues that have the most significant impact on the portfolio and projects and are of significant concern to the stakeholders. Sustainability reporting standards like GRI provide a wealth of materiality-related topics which can be very helpful. Following are the steps to help derive materiality topics from GRI guidelines (Dumay, Guthrie and Farneti, 2010; Courtneil, 2019; Global Sustainability Standards Board, 2020).

Aspect	Financial Materiality	Sustainability Materiality	Project Materiality
<b>Definition</b>	Relates to financial impacts, focusing on financial consequences and disclosure in financial reporting	Pertains to environmental, social, economic aspects, particularly in sustainability reporting contexts	Involves influential factors within the project environment (PE) that impact organisational strategy as well as sustainability dimensions.
<b>Stakeholder Focus</b>	Primarily concerns investors, creditors and financial analysts	Addresses a broad spectrum of stakeholders including customers, employees, regulators and the public, reporting standards, regulatory fulfilment.	Focused on stakeholders' determinations regarding the organisation's strategic direction as well as the project and portfolio's management and sustainability-resilience capabilities.
<b>Purpose</b>	To assess the materiality of financial information in financial statements for decision-making	To identify significant sustainability issues for disclosure and reporting	To build a set of criteria that help in the selection, screening and termination of projects within a portfolio and to balance and align portfolios with organisational strategies
<b>Primary Focus</b>	Economic impacts, revenue, expenses, and financial performance indicators	Environmental, social, and economic consequences of business activities	Project-specific factors that affect projects, portfolios and the organisation's ability to manage projects to successfully deliver sustainable outcomes

**Table 1. A comparison between financial, sustainability and project materiality**

- **Identify Relevant GRI Indicators:** Understand and review the list of GRI indicators within the guidelines that provide a framework for reporting sustainability. These indicators cover a wide range of economic, environmental, and social topics. Identify the indicators that are most relevant to an organisation's industry, operations, and stakeholder interests.

Category	Economic		Environmental							
Aspects	EC1	Economic performance	EN1	Materials						
	EC2	Market presence	EN2	Energy						
	EC3	Indirect economic impacts	EN3	Water						
	EC4	Procurement practices	EN4	Biodiversity						
			EN5	Emissions						
			EN6	Effluents and waste						
			EN7	Products and services						
			EN8	Compliance						
			EN9	Transport						
			EN10	Overall						
			EN11	Supplier environmental assessment						
			EN12	Environmental grievance mechanisms						
Category	Social									
Sub-Category	Labor practices and decent work			Human rights			Society		Product responsibility	
Aspects	LP1	Employment	HR1	Investment	S1	Local communities	PR1	Customer health and safety		
	LP2	Labor/management relations	HR2	Non-discrimination	S2	Anti-corruption	PR2	Product and service labeling		
	LP3	Occupational health and safety	HR3	Freedom of association and collective bargaining	S3	Public policy	PR3	Marketing communications		
	LP4	Training and education	HR4	Child labor	S4	Anti-competitive behavior	PR4	Customer privacy		
	LP5	Diversity and equal opportunity	HR5	Forced or compulsory labor	S5	Compliance	PR5	Compliance		
	LP6	Equal remuneration for women and men	HR6	Security practices	S6	Supplier assessment for impacts on society				
	LP7	Supplier assessment for labor practices	HR7	Indigenous rights	S7	Grievance mechanisms for impacts on society				
	LP8	Labor practices grievance mechanisms	HR8	Assessment						
			HR9	Supplier human rights assessment						
			HR10	Human rights grievance mechanisms						

**Table 2. GRI Guidelines - Economic, Environmental and Social categories (Calabrese et al., 2019; Rifkin, 2019; Global Sustainability Standards Board, 2020)**

- **Conduct Stakeholder Engagement:** Engage with the organisation's stakeholders to understand their concerns, expectations, and priorities regarding sustainability issues. This can be done through surveys, interviews, focus groups, or other forms of dialogue.
- **Prioritise Sustainability Topics:** Use the information gathered from stakeholder engagement to prioritise the sustainability topics that are of greatest concern to the stakeholders. Consider their input in determining which issues are most material to the organisation.



- **Assess Impact and Significance:** Evaluate the impact and significance of each prioritised sustainability topic on the organisation. Consider factors such as potential risks, opportunities, reputational impact, regulatory requirements, and financial implications.
- **Rank Materiality Topics:** Rank the sustainability topics based on their materiality to the organisation. Focus on those topics that have the most significant impact and are of the highest concern to both the organisation and stakeholders. Materiality matrices can be applied to manage stakeholder expectations from the very outset of the process.

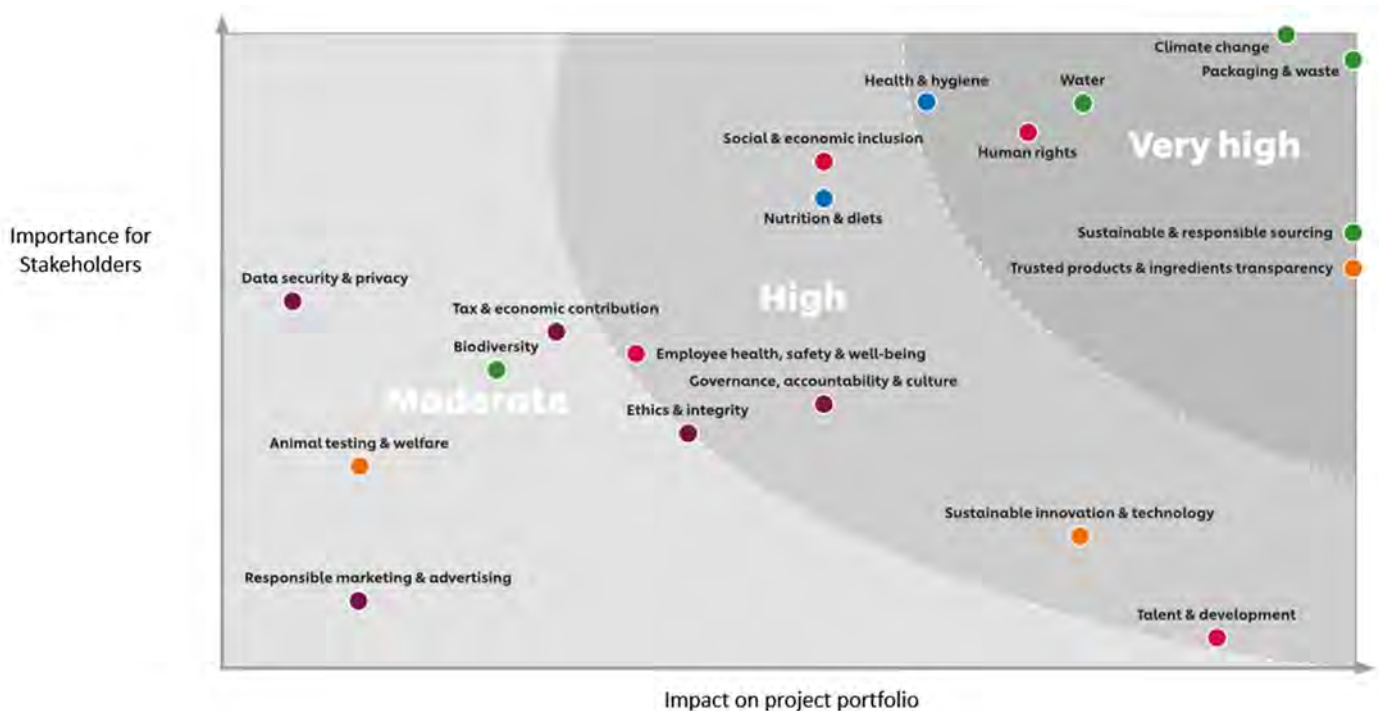


Figure 4. Project Materiality impact – a scatter graph representation

- **Validate with Key Stakeholders:** Validate the materiality assessment with key stakeholders, including internal and external parties. Ensure that the identified materiality topics align with stakeholder expectations and concerns.
- **Document the Materiality Assessment:** Document the materiality assessment process, including the criteria used, stakeholder input, and the rationale for prioritising specific material topics.
- **Integrate Materiality into Reporting:** Incorporate the identified materiality topics into an organisation's sustainability reporting. Report on these topics transparently and comprehensively, following the GRI reporting framework.

- **Regularly Review and Update:** Materiality is not static; it can change over time due to shifts in stakeholder priorities, industry trends, or organisational developments. Periodically review and update the materiality assessment to ensure its relevance.

By following these steps, one can derive materiality topics from the GRI guidelines that are aligned with an organisation's strategic context and stakeholder expectations. This process ensures that the principles, attributes and indicators of sustainability focus on the issues that matter most to the projects, portfolio, organisational strategy and its stakeholders, thereby enhancing transparency, accountability and meaningful buy-ins.

After conducting an extensive review of the existing literature on project management, project portfolio management, materiality and sustainability, it becomes evident that project materiality offers a range of significant advantages. The following advantages have been deduced from the insights gained by analysing and deducing relationships between these interrelated fields:

- **Leveraging Materiality Knowledge:** The materiality assessment concept, while firmly established in the financial sector and sustainability reporting standards, offers a valuable avenue for knowledge transfer into the project environment. Project practitioners can harness a wealth of existing knowledge and methodologies from these domains, leveraging the well-established principles of identifying and prioritising material issues. By integrating materiality assessment techniques into project management processes, practitioners can enhance their ability to identify and prioritise project-specific sustainability concerns and align them with organisational objectives, thus promoting more sustainable project outcomes. This knowledge transfer not only facilitates the incorporation of sustainability principles into project portfolio management but also fosters a more holistic and socially responsible approach to project execution, with the potential to yield positive impacts on both the project and the broader organisational sustainability agenda.
- **Alignment with organisational goals:** Materiality assessment has a broad reach covering not only an organisation's strategic objectives but also its sustainability objectives under one framework of assessment.
- **Enhanced Decision-making:** Materiality assessment supports data-driven decision-making. Organisations can allocate resources within portfolios more efficiently, prioritise sustainability initiatives, and ensure that the processes have a meaningful

impact. This, in turn, contributes to the organisation's long-term sustainability and competitive advantage.

- **Sustainable Strategy:** Helps build a sustainability strategy that looks beyond the project lifecycle to a more product-centric and product lifecycle approach.
- **Externalities and Risk Management:** Provide a better understanding of any negative externalities and an appropriate response from risk management.
- **Sustainable Deliverables:** Although the primary objective of the portfolio is not sustainability reporting, project deliverables and outcomes that have been materially assessed will be ready for sustainability reporting standards and regulatory scrutiny.
- **Creating Value:** Material assessment of projects will help create long-term net positive value for the portfolio, the organisation, markets and society.
- **Transparency and Stakeholder Management:** Effective communication with stakeholders is fundamental to sustainable portfolio management. Organisations conducting materiality assessments gain insights into what matters most to their stakeholders. Consequently, they can tailor their communication strategies to address these key concerns, thereby enhancing engagement, transparency, and trust.
- **Trade-offs and conflicts:** Sustainability often involves trade-offs and conflicts between strategic, economic, environmental, and social goals. Insights from materiality assessments will provide clarity to considerations associated with these trade-offs.
- **Cultural and Contextual Variations:** Sustainability practices and philosophies can vary significantly across cultures and regions. Materiality assessment can influence sustainability approaches and could add depth to the analysis.

### **Step 3: Applying Analytics in SAPoM**

Project failures often stem from human bias, unrealistic goals, changing expectations, unclear requirements, lack of executive support, cost overruns, and schedule delays, among other factors (Peddada and Sharma, 2020). Data-driven analytics or Business Intelligence (BI) can introduce objectivity into decision-making processes (Marques, Gourc and Lauras, 2011; Aldea *et al.*, 2019).

The absence of data insights in project environments can lead to poor decision-making and project failures (Sharda, 2018). Data-based risk management, driven by statistical probabilities, can be challenging to estimate but is crucial for successful project outcomes. As Niederman (2021) highlighted, analytics can revolutionise project management by providing incremental

and disruptive advancements. Sharda (2018) emphasises the importance of data-based risk management, allowing for quick and economic scenario building and impact assessment (Gachie, 2019; Barghi and Shadrokh Sikari, 2020; Hartwig and Mathews, 2020; Peddada and Sharma, 2020).

Sustainability, especially the social elements, often subjective, should be assessed as a risk within SAPoM. Analytical techniques like Fuzzy logic (Dursun, Goker and Mutlu, 2022) can translate qualitative sustainability data into quantitative formats, enabling measurable data, empirical evidence, and objective assessments alongside other data (Papadopoulos and Balta, 2022). Figure 5 (*The SAPoM Analytics Decision Engine using Fuzzy Logic*) depicts the decision process using the criteria for project materiality i.e., the strategic criteria based on an organisation’s objectives and sustainability-related criteria derived from GRI (Global Reporting Initiative) guidelines to include environmental, social and economic criteria.



**Figure 5. The SAPoM Analytics Decision Engine flowchart applying Fuzzy Logic**

Applying a mathematical approach like Fuzzy Logic can assist with approximate reasoning and decision-making under uncertainty. This allows values to range between true and false rather than adhering to strict binary logic. Such a method can be hugely beneficial when considering the social criteria of sustainability which can have unclear or ‘Fuzzy’ boundaries. Analytics

integrated into project management, as highlighted by Nayebi et al., (2015), encompass domains, data access, validation, and result reuse. The three key types of analytics - descriptive, predictive, and prescriptive - enable data-driven decision-making and forecasting in project management. These analytics tools can uncover hidden risks and gaps in project assumptions, ultimately improving project success (Hartwig and Mathews, 2020; Luk *et al.*, 2021). Other data analytics tools like Microsoft Power BI and IBM-SPSS can aid practitioners in analysing, describing, and predicting project outcomes, enhancing project management practices.

**Project Portfolio Management (PPM) Framework**

Several frameworks for project portfolio management (PPM) are currently in use. In the context of this research paper, provides a high-level overview of the key stages in the process, which are briefly outlined. The SAPoM conceptual framework can be segmented into the following phases, as illustrated in Figure 3, and a more comprehensive structure is presented in Figure 4:

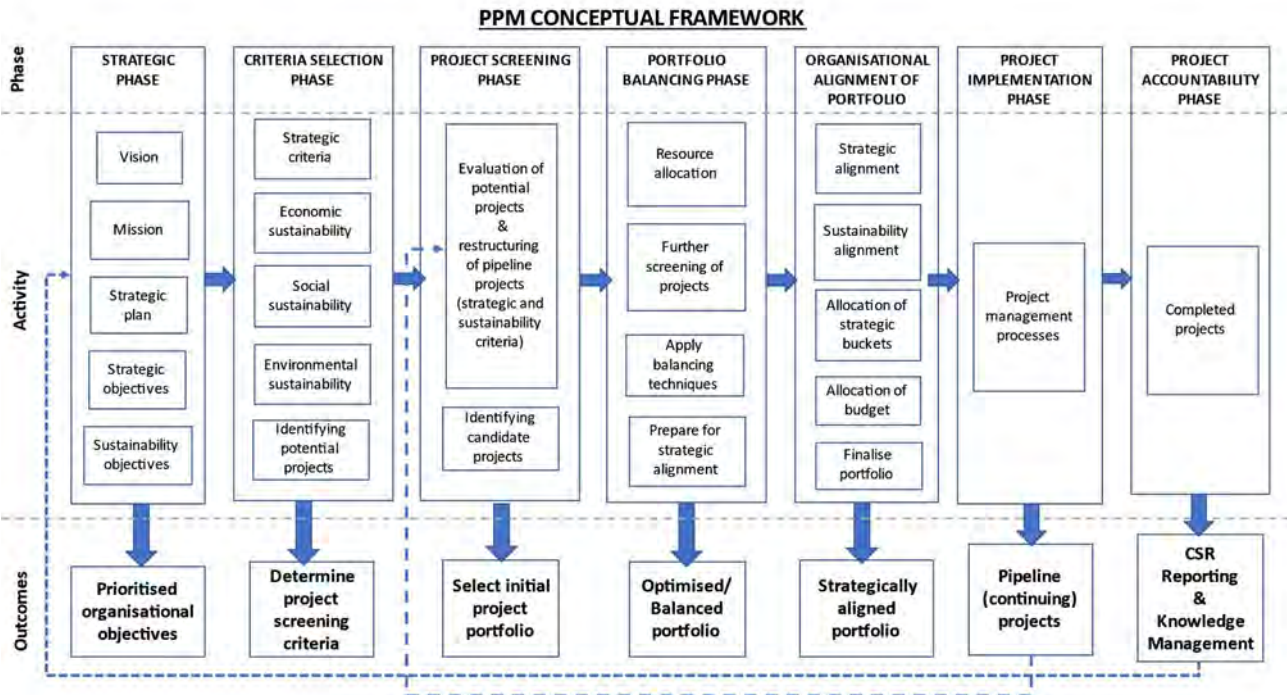


Figure 6. A Typical Project Portfolio Management Framework with Sustainability integrated in different phases (Mohammad and Pan, 2021)

- **Strategic:** In this phase, senior executives define the organisation's vision, mission, and approved strategic plan, which must incorporate both strategic and sustainability objectives. The goal is to prioritise organisational objectives effectively.
- **Criteria selection:** This stage aims to create criteria for screening candidate projects, involving key stakeholders who assess criteria aligned with organisational strategic and sustainability goals established in the prior phase. It is recommended that sustainability criteria can be derived from the sustainability reporting standards e.g., GRI and SASB.
- **Project screening:** To screen projects, each strategic and sustainability criterion must have a measurable or quantifiable basis. The result of this phase is the selection of an initial project portfolio using the combined criteria. Candidate projects include potential projects, as well as previously postponed ones that may require restructuring or modifications and are brought back for screening in this phase.
- **Portfolio balancing:** The initial portfolio must undergo an evaluation to ensure it maximises benefits for the organisation, considering resource constraints (Bible and Bivins, 2011). This phase also addresses situations where an organisation has a mix of good projects alongside many smaller or less relevant ones, or large projects with low risk that fail to drive growth and high returns for the organisation.
- **Strategic alignment:** This process focuses on the organisation's rationale for selecting particular projects and assessing their alignment with broader objectives. This ensures that selected projects do not unnecessarily strain organisational resources and have a meaningful association with the overall objectives (Moustafaev, 2017b).
- **Project implementation:** This phase is related purely to project management processes and methodologies. A typical format is depicted in Figure 3 below.
- **Project accountability and audit:** Completed projects will move to the next phase and undergo assessment in the accountability phase to learn lessons and reporting. Projects still in the pipeline will be re-evaluated to gauge their overall progress and alignment

with organisational objectives. These projects will go through the project screening phase within the framework of the PPM process outlined in this conceptual framework.

#### **4.3. Engaging Stakeholders with SAPoM**

The perception and interaction of different stakeholders with the SAPoM (Sustainability Augmented Portfolio Management) framework can vary based on their roles, interests, and priorities. An overview of how different stakeholders might engage with SAPoM is as follows:

- **Project Managers and Project Teams:** Project managers may view SAPoM as a comprehensive tool that helps them integrate sustainability into project planning and execution. They will likely engage with the framework to align project goals with sustainability objectives, ensuring their projects contribute to broader organisational sustainability. SAPoM is designed to be flexible to integrate into existing practices and processes with the least disruption and change can be made incrementally.
- **Business leaders, Decision-Makers and Sponsors:** Decision-makers may see SAPoM as a strategic approach for aligning project portfolios with overall business strategies and sustainability goals. They might use SAPoM for decision-making, resource allocation, and ensuring that the organisation's portfolio reflects a commitment to sustainability. As mandatory regulation on sustainability reporting is enforced business leaders will be able to report with confidence the processes they have adopted to ensure sustainable deliverables through projects.
- **Sustainability Officers:** Sustainability officers may appreciate SAPoM as a tool that facilitates the systematic integration of sustainability principles into project management processes. They could actively contribute to defining sustainability criteria, ensuring that projects align with environmental, social, and economic sustainability goals.
- **Financial Analysts:** Financial analysts may see SAPoM as a way to evaluate the financial viability of projects within the context of sustainability. They might engage with the framework to assess the economic impact of projects and ensure alignment with financial objectives.
- **Stakeholders and Community Representatives:** External stakeholders may view SAPoM as a transparency tool, providing insights into how projects contribute to sustainability and community well-being. They may engage with the framework to

access information on the social impact of projects and hold the organisation accountable for sustainable practices.

- **Data Analysts:** Data analysts could see SAPoM as an opportunity to leverage analytics for better decision support. They might work on implementing and refining analytics components of the framework, ensuring that data-driven insights contribute to effective decision-making.
- **Regulators and Compliance Officers:** SAPoM can highlight any greenwashing practices within project management from the initiation of projects. Regulators may appreciate SAPoM as a tool that supports organisations in meeting sustainability reporting requirements. They may engage with the framework to ensure that projects adhere to regulatory standards and contribute to overall compliance.

Understanding and addressing the diverse needs and perspectives of these stakeholders is crucial for the successful implementation and acceptance of the SAPoM framework. Regular communication, training, and feedback mechanisms can enhance stakeholder engagement and collaboration.

## **5.0 Discussing Research Impacts and Conclusion**

Sustainability will inevitably be a permanent feature in the project environment for the foreseeable future. Project management cannot continue in its current format and must embrace the principles of sustainability as an integral part of the strategies, practices, processes, and methodologies. This research delivers a platform for project professionals to accept the challenge of integrating sustainability within their processes and applying sustainability-led, data-driven decision-making for greater objectivity in their work.

The evaluation of project materiality within the project environment (PE) goes beyond merely ensuring alignment with an organisation's overarching objectives and strategies. It extends to embracing the tenets of sustainability, equipping the organisation to proactively address the impending sustainability reporting regulations introduced by governmental and international entities. The incorporation of analytics into project materiality assessments holds the potential to transform qualitative information derived from environmental and social sustainability dimensions into quantitative data, thus providing a standardised framework for practitioners to



comprehensively evaluate all materiality aspects. Analytics will improve forecasting of project outcomes and weed out unsustainable projects to avoid ‘greenwashing’, ‘conspicuous consumption’ and other unsustainable practices.

This research aims to establish a connection between sustainability and projects, facilitating their integration through a data-driven, impartial decision-making mechanism. The teams engaged in this process will enhance their understanding of sustainability principles within the project context and ready themselves to apply these principles across various aspects of their respective roles, encompassing routine business operations or business-as-usual (BAU) processes. The proposed research will have the potential to be developed as a Software as a Service (SaaS) for the benefit of project practitioners and strategy builders alike.

As sustainability disclosures become obligatory due to regulations, organisations can prepare themselves for impending changes. A mindset of 'sustainability in, sustainability out' should be developed into practitioners' approach to organisational operations. While it is undeniable that organisations will need to engage in post-activity sustainability reporting, the emphasis lies in instilling sustainability into processes and functions right from their inception. This proactive approach not only ensures that sustainability is addressed in a timely and comprehensive manner but also paves the way for a more holistic and sustainable organisational culture.

The research will serve as a catalyst for innovation and exploration within the field of sustainable project management, providing fresh opportunities for academic and research endeavours, and practical application. It aims to unlock uncharted pathways that can benefit scholars, researchers, and industry practitioners. This research is poised to expand the horizons of knowledge and practice in sustainable project management.

Consumer confidence and trust in organisations delivering sustainable products and services will be enhanced. As sustainability ‘labelling’ becomes part of government-led regulatory requirements, products and services delivered through projects can be labelled as ‘Sustainable projects’ ensuring that organisations do not conflict with anti-greenwashing regulations.

It would also be prudent to highlight some of the challenges and limitations in implementing SAPoM (Sustainability Augmented Portfolio Management) in the real world:

- The need for organisations to adapt their existing project management processes to integrate sustainability seamlessly. This involves a cultural shift and may encounter resistance from established practices.

- Due to the inherent complexity and context-sensitive nature of sustainability, stakeholder opinions, interpretations and value perceptions, subjectivity may not be eliminated from the decision-making process.
- The implementation of SAPoM requires a robust data infrastructure and analytics capabilities, which may be lacking in some organisations.
- Obtaining accurate and comprehensive sustainability data for project materiality assessment could be another hurdle, especially if such data is not readily available or if organisations have not traditionally focused on sustainability reporting.
- The successful implementation of SAPoM also depends on the willingness of stakeholders to embrace sustainability as a core aspect of project decision-making.
- The effectiveness of SAPoM may vary across different industries and project types, requiring tailored approaches for diverse contexts.
- Addressing these challenges necessitates a strategic and phased approach to implementation, considering both organisational and project-specific factors.

In conclusion, the incorporation of sustainability principles into project management strategies and processes is a vital step in navigating the ever-evolving landscape of sustainable project management. SAPoM, with its data-driven approach, serves as a valuable platform, promoting objectivity in project environments while aligning them with organisational objectives and anticipated disclosure regulations. This paradigm shift towards sustainability-oriented practices not only encourages organisations to embrace sustainability reporting but also positions them as frontrunners in their respective markets, gaining a competitive edge.

Moreover, the research that bridges sustainability and projects through data-driven processes stimulates innovation and opens doors for scholarly exploration in sustainable project management. This, in turn, boosts consumer confidence and trust in organisations that offer sustainable products and services. With government-mandated sustainability labelling becoming more prevalent, projects can play a pivotal role in ensuring compliance with anti-greenwashing regulations, further enhancing the competitive position of organisations in their markets. The acceptance of sustainability-driven practices leads to a wider adoption of sustainability disclosure, underscoring the essential role of sustainability in the future of project management. It not only positions organisations for success in a sustainability-conscious world but also reinforces the significance of transparency, responsibility, and sustainability in today's business environment.

## References

- Aldea, A. *et al.* (2019) 'Multi-Criteria and Model-Based Analysis for Project Selection: An Integration of Capability-Based Planning, Project Portfolio Management and Enterprise Architecture', *2019 IEEE 23rd International Enterprise Distributed Object Computing Workshop (EDOCW)* [Preprint]. Available at: <https://doi.org/10.1109/EDOCW.2019.00032>.
- Al-Marri, M. and Pinnington, A.H. (2022) 'Managing Sustainability Projects for Social Impact from a Corporate Social Responsibility Perspective', *Sustainability (Switzerland)*, 14(11). Available at: <https://doi.org/10.3390/su14116418>.
- APM Portfolio Management SIG (2019) *Portfolio Management - A Practical Guide*. 1st edn. Princes Risborough: APM.
- Barghi, B. and Shadrokh Sikari, S. (2020) 'Qualitative and quantitative project risk assessment using a hybrid PMBOK model developed under uncertainty conditions', *Heliyon*, 6(1). Available at: <https://doi.org/10.1016/j.heliyon.2019.e03097>.
- Bible, M.J. and Bivins, S.S. (2011) *Mastering Project Portfolio Management - A Systems Approach to Achieving Strategic Objectives*. 1st edn. Fort Lauderdale, FL: J. Ross Publishing.
- Brundtland, G.H. (1987) *Report of the World Commission on Environment and Development: Our Common Future, United Nations general Assembly Document A42/427*. Oslo.
- Cadenasso, M.L. and Pickett, S.T.A. (2018) 'Situating sustainability from an ecological science perspective', in J. Sze (ed.) *Sustainability: Approaches to Environmental Justice and Social Power*. New York: New York University Press, pp. 29–52.
- Calabrese, A. *et al.* (2019) 'Materiality analysis in sustainability reporting: A tool for directing corporate sustainability towards emerging economic, environmental and social opportunities', *Technological and Economic Development of Economy*, 25(5), pp. 1016–1038. Available at: <https://doi.org/10.3846/tede.2019.10550>.
- Carbon Disclosure Project (2023) *Carbon Disclosure Project, cdp.net*. Available at: <https://www.cdp.net/en/info/about-us> (Accessed: 29 January 2023).
- Clegg, S. *et al.* (2018) 'Practices, projects and portfolios: Current research trends and new directions', *International Journal of Project Management*, 36(5), pp. 762–772. Available at: <https://doi.org/10.1016/j.ijproman.2018.03.008>.
- Cooper, R., Edgett, S. and Kleinschmidt, E. (2001) 'Portfolio Management for New Product Development: Results of an Industry Practices Study', *R and D Management*, 31(4), pp. 361–380. Available at: <https://doi.org/10.1111/1467-9310.00225>.
- Courtneil, J. (2019) *GRI Standards for Sustainability Reporting: What They Are and Why They Matter, Process Street*. Available at: <https://www.process.st/gri-standards/> (Accessed: 12 May 2021).
- Doorasamy, M. (2017) 'Product Portfolio Management Best Practices for New Product Development: A Review of Models', *Foundations of Management*, 9, pp. 139–148. Available at: <https://doi.org/10.1515/fman-2017-0011>.

- Dumay, J., Guthrie, J. and Farneti, F. (2010) 'GRI sustainability reporting guidelines for public and third sector organizations: A critical review', *Public Management Review*, 12(4), pp. 531–548. Available at: <https://doi.org/10.1080/14719037.2010.496266>.
- Dursun, M., Goker, N. and Mutlu, H. (2022) 'Evaluation of Project Management Methodologies Success Factors Using Fuzzy Cognitive Map Method: Waterfall, Agile, And Lean Six Sigma Cases', *Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE*, 10(1), pp. 35–43. Available at: <https://doi.org/10.1039/b000000x>.
- EFRAG (2023) 04/09/2023 - EFRAG-GRI JOINT STATEMENT OF INTEROPERABILITY, [efrag.org](https://www.efrag.org). Available at: <https://www.efrag.org/News/Public-444/EFrag-GRI-Joint-statement-of-interopability-> (Accessed: 16 September 2023).
- Elkington, J. (1997) *Cannibals with Forks: The Triple Bottom-line of 21st Century Business*. 1st edn. Oxford: Capstone Press.
- Elkington, J. (1999) 'Triple bottom-line reporting: Looking for balance', *Australian CPA*, 69(2), pp. 18–21.
- Elliott, B. and Elliott, J. (2022) 'Integrated reporting: sustainability, environmental and social', in *Financial Accounting and Reporting*. 20th edn. Harlow, UK: Pearson Education Limited.
- Espiner, S., Orchiston, C. and Higham, J. (2017) 'Resilience and sustainability: a complementary relationship? Towards a practical conceptual model for the sustainability–resilience nexus in tourism', *Journal of Sustainable Tourism*, 25(10), pp. 1385–1400. Available at: <https://doi.org/10.1080/09669582.2017.1281929>.
- European Commission (2023) *EUR-Lex - Directive 32022L2464*. Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022L2464> (Accessed: 31 January 2023).
- Financial Conduct Authority (FCA) (2022) *Sustainability Disclosure Requirements (SDR) and investment labels*. London. Available at: [www.fca.org.uk/cp22-20-response-form](http://www.fca.org.uk/cp22-20-response-form).
- Gachie, W. (2019) 'Project sustainability management: risks, problems and perspective', *Problems and Perspectives in Management*, 17(1), pp. 313–325. Available at: [https://doi.org/10.21511/ppm.17\(1\).2019.27](https://doi.org/10.21511/ppm.17(1).2019.27).
- Ghannadpour, S.F. *et al.* (2021) 'Appraising the triple bottom line utility of sustainable project portfolio selection using a novel multi-criteria house of portfolio', *Environment, Development and Sustainability*, 23(3), pp. 3396–3437. Available at: <https://doi.org/10.1007/s10668-020-00724-y>.
- Global Reporting Initiative (2021) 'A Short introduction to the GRI Standards', *GRI Standards* [Preprint]. Amsterdam: Global Reporting Initiative. Available at: [www.globalreporting.org](http://www.globalreporting.org) (Accessed: 10 October 2023).
- Global Sustainability Standards Board (2020) 'Consolidated Set of GRI Sustainability Reporting Standards 2020'. Amsterdam: Global Sustainability Standards Board. Available at: <https://www.globalreporting.org/how-to-use-the-gri-standards/gri-standards-english-language/>.

- Hartwig, S. and Mathews, S. (2020) 'Innovation Project Risk Analytics: A Preliminary Finding: Use risk analytics for new product development for high-risk innovation projects.', *Research Technology Management*, 63(3), pp. 19–23. Available at: <https://doi.org/10.1080/08956308.2020.1733901>.
- Høgevold, N.M. *et al.* (2015) 'A triple bottom line construct and reasons for implementing sustainable business practices in companies and their business networks', *Corporate Governance*, 15(4), pp. 427–443. Available at: <https://doi.org/10.1108/CG-11-2014-0134>.
- Hubbard, G. (2009) 'Measuring Organizational Performance: Beyond the Triple Bottom Line', *Business Strategy and the Environment*, 18, pp. 177–191. Available at: <https://doi.org/10.1002/bse.564>.
- Huemann, M. and Silvius, G. (2017) 'Projects to create the future: Managing projects meets sustainable development', *International Journal of Project Management*. Elsevier Ltd, pp. 1066–1070. Available at: <https://doi.org/10.1016/j.ijproman.2017.04.014>.
- International Accounting Standards Board (IASB) (2018) *Conceptual Framework for Financial Reporting*.
- Kilbert, C.J. *et al.* (2012) *Working toward sustainability : ethical decision making in a technological world*. Wiley.
- Kuzemko, C. and Britton, J. (2020) 'Policy, politics and materiality across scales: A framework for understanding local government sustainable energy capacity applied in England', *Energy Research and Social Science*, 62. Available at: <https://doi.org/10.1016/j.erss.2019.101367>.
- Lister, G. (2015) *Mastering Project, Program and Portfolio Management: Models for Structuring and Executing the Project Hierarchy*. 2nd edn. Upper Saddle River, NJ: Pearson Education Limited.
- Luk, J.W. *et al.* (2021) 'From everyday life predictions to suicide prevention: Clinical and ethical considerations in suicide predictive analytic tools', *Journal of Clinical Psychology*, 78(2), pp. 137–148. Available at: <https://doi.org/10.1002/jclp.23202>.
- Marcelino-Sadaba, S., Felipe Gonz Alez-Jaen, L. and P Erez-Ezcurdia, A. (2015) 'Using project management as a way to sustainability. From a comprehensive review to a framework definition', *Journal of Cleaner Production*, 99, pp. 1–16. Available at: <https://doi.org/10.1016/j.jclepro.2015.03.020>.
- Marques, G., Gourc, D. and Lauras, M. (2011) 'Multi-criteria performance analysis for decision making in project management', *International Journal of Project Management*, 29(8), pp. 1057–1069. Available at: <https://doi.org/10.1016/j.ijproman.2010.10.002>.
- Martinsuo, M. and Geraldi, J. (2020) 'Management of project portfolios: Relationships of project portfolios with their contexts', *International Journal of Project Management*, 38(7), pp. 441–453. Available at: <https://doi.org/10.1016/j.ijproman.2020.02.002>.
- Martinsuo, M. and Lehtonen, P. (2007) 'Role of single-project management in achieving portfolio management efficiency', *International Journal of Project Management*, 25, pp. 56–65. Available at: <https://doi.org/10.1016/j.ijproman.2006.04.002>.

- Mohammad, J.A. and Pan, Y.-C. (2021) 'Sustainability, the Fourth Pillar of Project Portfolio Management - A Holistic Approach', *Journal of Modern Project Management*, 9(2). Available at: <https://doi.org/10.19255/JMPM02710>.
- Moustafaev, J. (2017a) 'Introduction to Project Portfolio Management', in *Project Portfolio Management in Theory and Practice*. 1st edn. Boca Raton, FL: CRC Press, pp. 3–19.
- Moustafaev, J. (2017b) 'The Three Pillars of Project Portfolio Management', in *Project Portfolio Management in Theory and Practice*. 1st edn. Boca Raton, FL: CRC Press, pp. 21–47.
- Nayebi, M. *et al.* (2015) 'Analytics for software project management - Where are we and where do we go?', in *30th IEEE/ACM International Conference on Automated Software Engineering Workshops, ASEW 2015*. Lincoln, NE: IEEE, pp. 18–21. Available at: <https://doi.org/10.1109/ASEW.2015.28>.
- Nüchter, V. *et al.* (2021) 'The concept of resilience in recent sustainability research', *Sustainability (Switzerland)*, 13(5), pp. 1–21. Available at: <https://doi.org/10.3390/su13052735>.
- Padin, C. *et al.* (2016) 'Validating a triple bottom line construct and reasons for implementing sustainable business practices in companies and their business networks', *Corporate Governance*, 16(5), pp. 849–865. Available at: <https://doi.org/10.1108/CG-12-2015-0163>.
- Papadopoulos, T. and Balta, M.E. (2022) 'Climate Change and big data analytics: Challenges and opportunities', *International Journal of Information Management*, 63. Available at: <https://doi.org/10.1016/j.ijinfomgt.2021.102448>.
- Peddada, K. and Sharma, T.K. (2020) 'Project Risk Management Using Analytical Hierarchy Process : Illustrative Case Study', *Abhigyan*, XXXVIII(1), pp. 40–50.
- Pensenstadler, B. and Femmer, H. (2013) 'A Generic Model for Sustainability with Process and Product-specific Instances', in *GIBSE 13*. Fukuoka.
- Per, G. (2019) *The many definitions of sustainability and sustainable development*, *Tiime.org*. Available at: <https://www.tiime.org/single-post/2019/10/17/the-many-definitions-of-sustainability-and-sustainable-development> (Accessed: 9 May 2023).
- Petit, Y. (2012) 'Project portfolios in dynamic environments: Organizing for uncertainty', *International Journal of Project Management*, 30, pp. 539–553. Available at: <https://doi.org/10.1016/j.ijproman.2011.11.007>.
- Project Management Institute (2021) *A Guide to the Project Management Body of Knowledge (PMBOK® Guide)-and The Standard for Project Management*. 7th edn. Newtown Square, Pennsylvania: Project Management Institute.
- Project-management.pm (2017) *Portfolio Governance Management | Project-Management.pm, project-management.pm*. Available at: <https://www.project-management.pm/portfolio-management/portfolio-governance-management/> (Accessed: 12 May 2021).
- Rifkin, S. (2019) 'Sustainability Materiality Matrices Explained', *NYU Stern* [Preprint]. New York, NY.

- Sardianou, E. *et al.* (2021) 'A materiality analysis framework to assess sustainable development goals of banking sector through sustainability reports', *Sustainable Production and Consumption*, 27, pp. 1775–1793. Available at: <https://doi.org/10.1016/j.spc.2021.04.020>.
- Sharda, R. (2018) *Business intelligence, analytics, and data science : a managerial perspective*. 4th Edition. Pearson.
- Silvius, G. (2017) 'Sustainability as a new school of thought in project management', *Journal of Cleaner Production*, 166, pp. 1479–1493. Available at: <https://doi.org/10.1016/j.jclepro.2017.08.121>.
- Silvius, G. and Marnewick, C. (2022) 'Interlinking Sustainability in Organizational Strategy, Project Portfolio Management and Project Management A Conceptual Framework', in *Procedia Computer Science*. Elsevier B.V., pp. 938–947. Available at: <https://doi.org/10.1016/j.procs.2021.12.095>.
- Sustainability Accounting Standards Board (2023) *About Us - SASB*, [sasb.org](https://www.sasb.org). Available at: <https://www.sasb.org/about/> (Accessed: 30 January 2023).
- Sze, J. (2018) *Sustainability: approaches to environmental justice and social power*. Edited by J. Sze. New York: New York University Press.
- Tarmuji, I., Maelah, R. and Tarmuji, N.H. (2016) 'The Impact of Environmental, Social and Governance Practices (ESG) on Economic Performance: Evidence from ESG Score', *International Journal of Trade, Economics and Finance*, 7(3), pp. 67–74. Available at: <https://doi.org/10.18178/ijtef.2016.7.3.501>.
- Task Force on Climate-related Financial Disclosures (TCFD) (2022) *Task Force on Climate-Related Financial Disclosures | TCFD*, [fsb-tcfd.org/](https://www.fsb-tcfd.org/). Available at: <https://www.fsb-tcfd.org/> (Accessed: 30 January 2023).
- To, C.K.M. and Chau, K.P. (2022) 'Characterizing sustainability materiality: ESG materiality determination in technology venturing', *Sustainable Technology and Entrepreneurship*, 1(3). Available at: <https://doi.org/10.1016/j.stae.2022.100024>.
- Tsalis, T.A. *et al.* (2020) 'New challenges for corporate sustainability reporting: United Nations' 2030 Agenda for sustainable development and the sustainable development goals', *Corporate Social Responsibility and Environmental Management*, 27(4), pp. 1617–1629. Available at: <https://doi.org/10.1002/csr.1910>.
- Turan, F.K. *et al.* (2008) 'Organizational Sustainability: A New Project Portfolio Management Approach that Integrates Financial and Non-Financial Performance Measures', in *Industrial Engineering Research Conference*.
- United Nations (2001) *Indicators of Sustainable Development: Guidelines and methodologies*.
- United Nations (2015) *Transforming our world: the 2030 Agenda for Sustainable Development*.
- Wu, S.R., Shao, C. and Chen, J. (2018) 'Approaches on the screening methods for materiality in sustainability reporting', *Sustainability (Switzerland)*. MDPI. Available at: <https://doi.org/10.3390/su10093233>.

# A soft ecosystems methodology of digital innovation through a case study of the insurance industry's response to connected cars.

Anushri Gupta<sup>1</sup>, Roser Pujadas<sup>2</sup>, Will Venters<sup>1</sup>

<sup>1</sup>London School of Economics; <sup>2</sup>University College London

## Abstract

*Digital innovations are increasingly the result of the combination of resources and skills beyond organisational boundaries and/or industries, resulting in the rise of a new organisational form – the digital innovation ecosystem. Studies of ecosystems have considered the structural, conceptual, and temporal aspects such as emergence, however, understanding of ecosystems in flux remains nascent. In particular, existing research focuses on change in the ecosystem rather than imagined (anticipated) change and its impact on the ecosystem's future. To address this, we adopt a systems thinking perspective that builds on soft systems methodology and British Cybernetics to propose what we refer to as 'soft ecosystems methodology' and introduce the notion of ecosystem-as-was, ecosystem-as-is, and ecosystem-to-be. We discuss this emergent methodology in light of the disruptive digital innovation faced by the car insurance industry.*

**Keywords:** Systems thinking, Ecosystems, Digital innovation, ecosystem-as-was, ecosystem-as-is, ecosystem-to-be, imagining, soft ecosystem methodology.

## 1.0 Introduction

Digital innovations are increasingly the result of leveraging resources and interacting with entities beyond a single organisation's boundary (Seo 2017). This has led to a "new organisational form of digital innovation" (Wang 2021) – the digital ecosystem. Such digital ecosystems form a collective aggregate of loosely coupled autonomous actors lacking hierarchical control, but collectively undertaking activities around the development and implementation of digital technologies (ibid).

Existing research has addressed the structuring roles of actors within digital innovation ecosystems leading to an intense firm-centric focus on platform ecosystem dynamics in which one actor (the "platform" or lead firm e.g. Apple or SAP) is dominant (e.g. Parker et al., 2017; Qiu et al., 2017; Schreieck et al., 2021).

Yet the term ecosystem draws upon an ecological metaphor of symbiotic relations between actors (Moore 1993) in which such dominance of an actor is rarely seen.



Furthermore, recent research has highlighted a lack of research into integration (the whole) of an ecosystem in contrast to the dominant focus on key actors (the parts) – a focus which inherently overlooks the wider ecosystem dynamics (Wang, 2021). Wang (2021) usefully addresses this gap through an ecosystemic lens that builds upon ecology, and in particular part-whole relations (ibid). Digital ecosystems are thus constituted from parts which interact to form a whole, albeit, we argue, those parts may vary from simple API services, through companies of various forms, to entire digital infrastructures (such as the Internet).

Literature predominantly adopts a static view of ecosystems. While some researchers have studied ecosystems' evolution (Pujadas et al., 2024), the collective actions undertaken for that change to happen are understudied. Particularly, how knowledge and the imagined futures about the ecosystem shape both ecosystems and actors' evolution. We believe systems thinking can help address this gap. We introduce and build upon a socio-technical, British and soft systems (Checkland, 1981) perspective to the study of digital innovation ecosystems – one which, we believe, better accounts for digital innovation ecosystems' ontology.

We thus seek to contribute a systems perspective to the study of digital innovation ecosystems and in particular digital ecosystems in flux. From these ideas we intend to build a practical soft ecosystems methodology (based on SSM) which has practical benefit. We are starting to research this empirically through a design science informed study of the insurance industry as it faces the disruption of increasingly digitised (and indeed self-driving) cars and transport – disruption that is seriously affecting many actors within the ecosystem. In designing our **soft ecosystems methodology**, we seek to develop a lens by which we can illuminate the emergence and transformation of the actors and the ecosystem they inhabit. Our design activity is informed by Design Science (Hevner et al., 2004) using soft systems thinking as a *kernel theory* which is elaborated through cycles of empirical research within the insurance industry as we build our theory. Our over-arching aim is to build “tough, analytic, partly formalizable, partly empirical, partly teachable doctrine” (Simon 1996, p.113) to address the strategic challenges of digital innovation ecosystems by creating an artefact that can apply, test, modify, and extend ‘kernel theories’ (Markus et al. 2002; Walls et al. 1992). Our overarching aim – in keeping with other design

science is relevance to practitioners in the field (Straub and Ang 2011) and for this reason we are working closely with an insurance provider facing the challenge of connected cars and digital transformation influencing its digital ecosystem.

Our aim then is to address a broad research question of “*How can practitioners understand the emergence of a digital ecosystem as a sociotechnical process, and how can systems thinking assist in this understanding?*” We do this as follows. First, we review the ecosystem literature in greater detail as we seek to identify how it might align with our systems perspectives. Second, we introduce our systems thinking approach which is informed by European research on systems thinking. Third we briefly outline our case study from the insurance industry in which we aim to test our designed soft ecosystem methodology. Finally, we outline our soft ecosystem methodology and discuss our future research plans for testing and further refining our design.

## **2.0 Ecosystems research through the lens of systems perspective**

Adopting the term business ecosystems, coined initially by Moore (1993), ecosystems are “*literally and phenomenologically systems*” (Phillips & Ritala, 2019, p. 2), that illustrate the dynamic interactions between, and co-evolution of various actors in ecosystems. Scholars have studied this in context of varied ecosystems like technology, knowledge, business, innovation, platform, and entrepreneurial ecosystems, [see for instance (Gawer, 2021; Gomes et al., 2021; Meynhardt et al., 2016; Scaringella & Radziwon, 2018)].

Drawing on above such studies and synthesising seminal papers on ecosystems, Hou & Shi, (2021) and Phillips & Ritala, (2019) highlight that our understanding of ecosystems is limited to – 1) Ecosystems’ boundaries, 2) ecosystem actors and the activities and relationships between them, 3) the temporal evolution and emergence of ecosystems. Yet further research is needed to understand the mechanisms which underpin such emergence (Baldwin 2019, Kapoor 2018) in general.

Phillips & Ritala (2019, p. 3) underscore the importance to draw on systems thinking to address this by emphasising the “*need to link our (ontological) knowledge about systems to systems thinking (which is conceptual and epistemological)*”. Where

ecosystems literature has adopted systems thinking perspective, Badinelli et al. (2012, p. 499) note that scholars have looked at studying the structure of ecosystems by identifying actors, activities, and interconnectedness of entities, but have lacked applying “*system thinking principles, which often disrupt the traditional thinking.*”<sup>1</sup>

Ecosystems are best understood as systems and better studied using systems theory because it “*provides insights into the emergence.... [of] ... the complex constellation of connections among ecosystem components*” (Roundy et al., 2018, p. 2). Distinguishing between ‘general systems’ – systems which do not exhibit complex dynamic interactions and non-linearity and are therefore studied using simple systems theory that isolate and parametrise individual components – and ‘complex systems’ as evidenced in biology, management etc., Roundy et al. (2018, p. 2) claim that ecosystems are “*complex adaptive systems (CAS) - that cannot be explained using general systems theory.*”

Barnes et al. (2003, p. 276) define Complex Adaptive Systems as “*open systems in which different elements interact dynamically to exchange information, self-organize, and create many different feedback loops, relationships between causes and effects are nonlinear, and the systems as a whole have emergent properties that cannot be understood by reference to the component parts.*” Wollmann & Steiner (2017, p. 2) echo this by adding that “*the evolution of the system is the result of interactions between agents, where each of them acts in response to the behaviour of the other agents in the system, which ensure it has its own dynamic.*”

Of late, scholars have been adopting CAS principles to study ecosystems. Phillips & Ritala (2019) adopt a CAS lens to propose a methodology to study ecosystems. In doing so, they provide an epistemological tool to study ecosystems by focusing on either conceptual, structural, or temporal dimensions. Looking at data ecosystems as CAS, Brous et al. (2019, p. 3) discuss how interactions between ecosystem elements

---

<sup>1</sup> Interestingly in a professional development workshop titled “Fostering Rigor in Innovation and Entrepreneurial Ecosystem Research: Concepts, Methods and Theory” at the 2017 Academy of Management Annual Conference, it turned out that the majority (44% of scholars) supported the need for aligning research on ecosystems with systems thinking principles (Ritala & Gustafsson, 2018).

take place “*without any singular entity deliberately managing or controlling them. CASs are “dynamic systems” which are able to adapt and evolve to changing circumstance.*” CAS are systems in which macro level behaviour emerge from and is influenced by interactions between system elements at the micro level (Roundy et al., 2018).

These dynamic interactions shed light on the unique behaviour, or rather characteristics of ecosystems understood as CAS. Synthesising the works of (Brous et al., 2019; Roundy et al., 2018), we note that the characteristics are – 1) *Non-linearity*: nonproportionate response to stimuli; interactions between agents is complicated and that the behaviour cannot be predicted by simply understanding how each component and agent behave; 2) *Self-organisation and feedback* as a means of co-ordination and knowledge/information exchange between agents in the system across multiple levels. Feedback can be further understood as ‘*positive feedback loop*’ – *described as autocatalytic* (Morrison, 2008), *that result in system behaviour to increase or decrease indefinitely* (Roundy et al., 2018); or ‘*negative feedback loop*’ that moves the ecosystem *to a steady or equilibrium state, even if temporarily* (Lichtenstein & Plowman, 2009); 3) *Aggregation*- means by which agents cluster or combine in groups that contributes to the system’s identity; 4) *(Co)-evolution* where activity or event at one element affects the rest in the system; and finally, 5) *open with ill-defined boundaries* where unlike closed systems, such ecosystems do not follow a predictable path and are far from equilibrium, engaging and reacting to disturbances with elements internal and external to the ecosystem (Bhardwaj et al., 2023; Roundy et al., 2018).

Other integral ecosystem features are knowledge transfer, and collaboration and competition between ecosystem actors - key foundations governing ecosystem functioning (Scaringella & Radziwon, 2018). In what the authors refer to as a ‘territorial ecosystem’, that builds on literature on business and innovation ecosystems and the triple/quadruple helix (Schütz et al., 2019), they argue that an understanding of transfer of tacit and explicit knowledge between ecosystem actors is essential, and that it depends on the proximity between them leading to ecosystem dynamics.

Through the above properties, although integral to understanding of ecosystems, we realise that our understanding of such *ecosystems in flux* remains limited. A field that is embedded in complexity science (Benbya & McKelvey, 2006), – i.e. how system elements interact with each other and the embedded environment to adapt and create new structures, processes and system wide properties across levels (Lissack & Letiche, 2002; Werder & Maedche, 2018) – needs further attention. To elaborate (Choi et al., 2001) note that the most productive stage for a CAS is the ‘*quasi-equilibrium state*’ that maintains balance between complete order and incomplete disorder – what Sapir (2019) refers to as the ‘*edge of chaos*’. Our understanding of such stages of flux and how new knowledge processes lend to accommodate and adapt to these changes demands further investigation.

### **3.0 British soft systems perspectives brought into digital innovation ecosystems.**

The review above has shown how researchers have sought to understand digital innovation ecosystems through systems theory. This is however dominated by approaches focused on systemic complexity (e.g. Complex Adaptive Systems theory (Holland, 1995)) with a functionalist bent (drawing upon its origins in natural sciences) (Jackson, 2019). Such North American systems’ perspectives towards ecosystems therefore focus on a hard systems (realist) ontology. Ontology, or the metaphysics of “being” is more akin to a spectrum of dimensions than a taxonomy – extreme positions are unrealistic<sup>2</sup>. Hard systems ontology see social reality as existent from naturalistic causes (e.g atoms, biology), and causality derivable through natural science methods (e.g. Lawson (2012) or Searle’s ontological positions (2006) seek repeatability) and even extending as far as causal relations or features. Ecosystems then are modelled as constituted from isolated agents who adapt based on their interactions leading to evolution. In contrast, we were influenced by Jackson’s assertion that the complexity theory upon which such studies are based: “*sees structure as micro-emergent but as possessing no independent reality and causal powers of its own. By contrast many sociologists tend to see humans as born into*

---

<sup>2</sup> As Sokal<sup>2</sup> famously argued, anyone wholly believing in subjectivity should experiment by throwing themselves off a tall building, however we could also add that extreme objectivity involves throwing human subjectivity, religion, and social critique of science off a building too. Metaphysics remains a complex domain of debate.

*social structures which constrain life opportunities and socialise individuals in ways that make it more likely that their agency will reproduce rather than change existing arrangements*” (Jackson, 2019, p.127). Our ontological position then is more relational and performative whereby stability is only enacted through ongoing accomplishment of actors within the digital ecosystem, whereby the subjective actions of those within the ecosystem performatively create the ecosystems. This ontological position led us back to the origins of systems thinking and more British softer perspectives based on an arguably more constructionist ontology.

Cybernetics emerged in the 1940s as a science of “control and communication in the animal and machine” (Wiener, 2019, originally 1948), and was pushed forward, by what Jackson (2019, p.95) terms, “British Cybernetics” - a particular branch that reflected a “performative idiom” (Pickering, 2002) in which systems are subjective and perform an active role embedded within contexts. For this branch, systems focus upon that which they broadly encounter in relation to its impact upon them. Ashby’s famous law of Requisite Variety (Ashby, 1956) is thus described as understandable in terms of a relationship between a biological system and its whole environment rather than mechanistic system actors’ interaction *per se* (Espejo, 1993). Indeed Stafford Beer, in his “management cybernetics”, sought to employ cybernetic concerns (feedback, black-box, and Ashby’s management of variety) to model an organisations response to all possible external stimulus (Beer, 1984). Beer’s Viable Systems Model (VSM) thus sought a dynamic equilibrium between the organisation and its whole environment (just as a biological ecosystem– say a pond<sup>3</sup> - might).

The digital ecosystems we are seeking to examine then are an entanglement of social, material and symbolic factors (Benbya et al., 2020). Yet existing CAS based approaches fail to account successfully in the social and symbolic aspect of this entanglement. British Cybernetics, with its links to the Tavistock institute<sup>4</sup> and its long history of sociotechnical studies (Mumford, 2000), offers an alternative lens that emphasises the social, and in which systems thinking is moved from an attempt to

---

<sup>3</sup> Indeed, interestingly, these cyberneticians sought to experiment whether a pond might be connected such that it could act as the control for the production of a factory!

<sup>4</sup> [Tavistock Institute of Human Relations \(tavistockinstitute.org\)](http://tavistockinstitute.org)

formally model agents interaction within an ecosystem, to being an epistemic lens to carve out elements of the sociotechnical mess (including a broad environment and subjective elements) as though it were systemic (Checkland, 1981; Checkland, 1997) –that is to “*bring rigor to the subjective*” (Checkland, 2002). Yet such pragmatic and interpretive approaches as SSM can fail to account for the role of complexity of material agency within such modern socio-technical contexts where various digital infrastructures (Tilson et al., 2010), cloud-based services, and AI systems which learn and act autonomously (Berente et al., 2021) are increasingly involved. We thus need a new approach.

Before designing our approach though it is useful to have an example context in which to discuss the emerging theory. For this we turn to digital innovation ecosystem of car insurance – a relevant site for the study of digital ecosystems in flux as it involves a range of different digital actors working to define what the future of transportation will be – and working to create that future.

#### **4.0 The car insurance digital innovation ecosystem in flux**

The future of car insurance is highly uncertain. Rapid technological developments, changing regulatory environments and a competitive environment with threat of new entrants continues to challenge traditional ways of working and operating business models.

Once a mechanical artefact, a car is now mostly equipped with and connected to a wide range of digital technologies, providing digitally-enhanced functionalities and services, such as driving support or automation, real-time road updates and route recommendations, car maintenance support, infotainment, etc. (Lenfle and Midler 2009; MarketsandMarkets 2021; McKinsey & Company 2014; Seiberth and Gruendinger 2018). The development of smart and (semi-)autonomous cars relies on a complex range of interconnected technologies such as cameras, sensors, computers and Artificial Intelligence-based systems to monitor the environment and support or automate driving decisions (Tu et al. 2022). These systems involve hardware, software and immense amounts of data, making the automotive one of the industries that generates most volume of data (Seiberth and Gruendinger 2018, p. 11).

At its core, insurance consists of risk transfer- i.e. pooling risks and sharing those through commercialisation of that pooling. It essentially involves risk analysis based on actuarial science (Rana et al., 2022). This defines its unique characteristic of being a historically data informed business model to assess and distribute risks. With the incorporation and deployment of advanced technology and IoT devices, these sources of data have increased manifold resulting in opportunities for innovative business models and services. Traditional approaches to calculating risk and assigning liability need to be revised. Partly-autonomous safety features are reducing accidents, so some insurers offer discounts to drivers who use these (Wiggers 2021). Furthermore, risks models for (semi-)autonomous cars cannot be based only on driver attributes but need to consider “technology viability, sensor shelf life, and the impact of local road conditions and infrastructure” (Deloitte US n.d.). In addition, risks related to cybersecurity or the malfunctioning of systems become significant (Deloitte US n.d.; Tu et al. 2022). Insurers will need to develop new skills and expertise to develop adequate risk assessments (Huckstep n.d.). These new risks might also require new kinds of protection, for instance, against identity theft.

New entrants, particularly big tech companies, InsurTech start-ups, and OEMs at the forefront of car automation such as Tesla, are disrupting the insurance market through data-driven models and competing with their own insurance offerings (Deloitte US n.d.; Quantalyse Belgium and Schönenberger Advisory Services 2019). Insurance companies are also adapting to data-driven business models. Leveraging data analytics, insurers can offer personalised rates, usage-based insurance premiums, improve their risk assessments (Nicley et al. 2020; Quantalyse Belgium and Schönenberger Advisory Services 2019), and overall can be “a potential game-changer for underwriting, pricing, claims, and business-line shifts.” (Deloitte US n.d.) However, data-driven business models present challenges in terms of controlling access to or acquiring these data (this may involve partnering with OEMs), integrating and ensuring the quality of data from a range of sources, and developing the necessary analytical skills and capabilities (Huckstep n.d.; Karp and Kim 2017; Nicley et al. 2020). They also raise ethical, and legal concerns around privacy, and customers attitudes also need to be considered (Huckstep n.d.).



Although the conundrum remains regarding the extent to which these new entrants are a threat to the traditional insurance firm - given the high variable costs, easy access to capital and lack of customer awareness (Lekkerkerk, 2023; Palmié et al., 2020; Ralph, 2023) – our empirical observations suggest that incumbents are addressing these changes in the environment and expected disruption by developing new capabilities, and new business models. Yet, in the midst of an important digital transformation taking place, the sense of flux is strongly felt by the incumbent financial firm we are studying, as their future seems uncertain. In trying to make strategic decisions, our firm tries to make sense of the emerging ecosystem and based on such imagined future, acts upon it – and thus changes it.

## 5.0 Our soft digital ecosystems methodology

The flux of our case study suggests the need for a temporal perspective on ecosystems. For this we define an **ecosystem-as-is** (today's insurance ecosystem – always in flux) and an **ecosystem-to-be** (the future ecosystem as currently imagined). Each of these are highly subjective notions (depending upon who we interview for example) and based on an individual or groups **Weltanschauung** (the “*stocks of images in our heads put there by our origins, upbringing and experience of the world which we use to make sense of the world and which normally go unquestioned*” (Checkland, 2013)). **Transformation** (also taken from Checkland) is, (based on the Weltanschauung adopted by the relevant actor) an ecosystem participant's action within the ecosystem, or action to change the ecosystem. Ecosystems thus evolve over time by the interrelationships of Transformations (undertaken by actors with intended purpose) – moving from ecosystem-as-is towards one specific ecosystem-to-be. Ecosystems are thus **emergent** and influenced by the power (broadly defined) within actors' transformations.

Each of these ecosystems are **open** (in that all are influenced by ‘outside’ things which actors might not consider part of the ecosystem such as culture, or globalisation or deglobalisation (Nambisan & Luo, 2022)). Actors within the ecosystem-as-is are involved in **sociotechnical self-organizing actions** that seek to bring *their anticipated* ecosystem-to-be into existence. They thus both learn from the current environment and **imagine** into existence the future environment (including through AI or predictive analytics). The relationship between cause and effects are thus highly

non-linear with significant feedback. In addition, the past influences the future – there is an ecosystem-as-was (the historic insurance market) which imposes norms, structures and roles (Checkland, 1999) (e.g. legislation, roles such as actuaries, norms such as expecting yearly insurance policies), but also an installed base of technology, upon the current and future plans.

An equilibrium is achieved by *autopoietic actors*– that is by actors that organise themselves to recreate themselves in relation to their ecosystem, and which is self-referential – rearranging them itself in the face of new knowledge to continue to maintain existence (Demetis & Lee, 2016) see also (Von Krogh et al., 1994). Actors within the insurance ecosystem are thus attempting (through their imagining and learning) to recreate themselves in relation to what others are doing and how they imagine the emergence of the ecosystem-to-be. New actors are also entering the ecosystem and putting forward their own images of ecosystem-to-be.

Within the ecosystem actors therefore **knowledge is emergent**, “it is always in the making, emerging from interactions between systems and forever leaving new things to be discovered” (Jackson, 2019, p.95) – but also agential as knowledge changes the ecosystem-to-be and thus the ecosystem-as-is. *“The interactions of a living system with its environment are cognitive interactions, and the process of living itself is a process of cognition. To live is to know”* (Maturana & Varela, 1980) – and as the actors within the ecosystem are live (companies have humans within them) so they are cognitive and imagining. Emergent unanticipated behaviour will thus occur through process of interaction between imagining, knowing, and acting leading to an inherent unknowability becoming of the ecosystem (Pickering, 2002). Learning and action are entwined as actors not only learn and imagine but also create (for example by setting up new insurance services to test new ideas or collaborating on legislative changes to influence other actors to share data). These imaginings thus become and the ecosystem-as-is moves forward. See table below for summary.

<p><b>Our soft ecosystems concepts</b></p>	<p><b>Our interpretation for digital innovation ecosystems</b></p>
--	--

<b>Actor</b>	An element of the ecosystem which is autopoietic – usually a company offering a service within the ecosystem who has humans who learn and imagine. AI/ML models, robots are also considered as actors in this system.
<b>Transformation</b>	The perceived intentional action undertaken by an actor (or group of actors) which changes some element of the ecosystem. Transformation is strongly influenced by power within the ecosystem. Transformation can be conceptualised in systems terms as input transformed into output – but remains subjective since, in SSM terms, we are using systems theory as an epistemic device rather than realist model.
<b>Ecosystem-as-is and emergence</b>	The contemporary ecosystem as it is understood by a human actor within it. This is highly subjective as the extent of the ecosystem can never be known and as actors hold differing Weltanschauung. Ecosystems-as-is however emerge over time as various transformations impact upon it moving it into a future (which may or may not reflect imagined ecosystems-to-be).
<b>Ecosystem-to-be</b>	The future ecosystem as it is imagined by a human actor within it. This is highly subjective and based upon actors' Weltanschauung.
<b>Ecosystem-as-was</b>	The <i>roles, norms and structures</i> (Checkland 1999) of the past ecosystem continue to influence the contemporary ecosystem-as-is. Similarly, the installed base of technology (e.g. cars). It is thus necessary to consider history in the analysis.
<b>Weltanschauung/Worldview</b>	The held beliefs of individuals and actors within the ecosystem.
<b>Imagining</b>	The act of making sense of the present and building mental models of the future. The process of imagining is human though it may be collective and sociotechnical (e.g. within a company and using predictive analytics or modelling tools and simulations).
<b>Knowledge</b>	Knowledge can be both tacit and explicit. Explicit knowledge can be digitised and non-digitised. This also includes AI/ML models since knowledge can also be based on predictions.
<b>Feedback</b>	In contrast to the CAS, for us feedback is both action and imagining – it can be the subjective view of what actors think others will do, and the inertia of previous actions combined. Views on feedback's positivity or negativity are somewhat subjective (though companies' failures can obviously be observed).

**Table 1. Soft ecosystems perspective of digital innovation**

## 6.0 Future research plans

Having developed this emerging theory we are working with an insurance provider to undertake a qualitative study of how they are learning, and acting, within the insurance ecosystem-as-is and how this may influence the future of that ecosystem

(the insurance-ecosystem-to-be). Thus far we have undertaken 5 initial exploratory interviews and a workshop. Our aim is to use this set of ideas within our analysis of the ecosystem's emergence and flux. We also intend to work with actors/organisations beyond the insurance firm, such as car manufacturers and data aggregators, to understand the ecosystem as a whole.

## References

- Ashby, W. R. (1956). *An introduction to cybernetics*. Methuen & Co Ltd.
- Badinelli, Ralph, Barile, Sergio, Ng, Irene, Polese, Francesco, Saviano, Marialuisa, & Di Nauta, Primiano. (2012). Viable service systems and decision making in service management. *Journal of Service Management*, 23(4), 498–526.
- Barnes, Marian, Matka, Elizabeth, & Sullivan, Helen. (2003). Evidence, Understanding and Complexity: Evaluation in Non-Linear Systems. *Evaluation*, 9(3), 265–284.
- Beer, S. (1984). The Viable System Model: Its provenance, development, methodology and pathology. *Journal of the Operational Research Society*, 35, 7-36.
- Benbya, Hind, & McKelvey, Bill. (2006). Toward a complexity theory of information systems development. *Information Technology and People*, 19(1), 12–34.
- Benbya, H., Nan, N., Tanriverdi, H., & Yoo, Y. (2020). Complexity and information systems research in the emerging digital world. *MIS Quarterly*, 44(1), 1-17.
- Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing Artificial Intelligence. *MIS Quarterly*, 45(3).
- Bhardwaj, Rohit, Srivastava, Saurabh, Bindra, Sunali, & Sangwan, Sumit. (2023). An ecosystem view of social entrepreneurship through the perspective of systems thinking. *Systems Research and Behavioral Science*, 40(1), 250–265.
- Brous, Paul, Janssen, Marijn, & Herder, Paulien. (2019). Next Generation Data Infrastructures: Towards an Extendable Model of the Asset Management Data Infrastructure as Complex Adaptive System. *Complexity*, 2019, 5415828.
- Checkland, P. (1981). *Systems Thinking, Systems Practice*. Wiley.
- Checkland, P. (1997). *A basic introduction to soft systems methodology*.
- Checkland, P. (2002). Launch Speech of Phronesis: The centre for Appreciative Systems Research. In U. O. Salford (Ed.). Salford.
- Checkland, P. (2013). Soft systems methodology. *Encyclopedia of operations research and management science*, 1430-1436.
- Choi, Thomas Y., Dooley, Kevin J., & Rungtusanatham, Manus. (2001). Supply networks and complex adaptive systems: control versus emergence. *Journal of Operations Management*, 19(3), 351–366.
- Deloitte US. (n.d.). *Autonomous vehicles and the insurance industry*. *Deloitte United States*.
- Demetis, D. S., & Lee, A. S. (2016). Crafting theory to satisfy the requirements of systems science. *Information and Organization*, 26(4), 116-126.
- Espejo, R. (1993). Giving Requisite variety to strategy and Information systems. In F. Stowell (Ed.), *Systems Science* (Stowell,FAet all ed.). Plenum Press.

- Gawer, Annabelle. (2021). Digital Platforms and Ecosystems: Remarks on the Dominant Organizational Forms of the Digital Age. *Innovation: Organization & Management*. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3900105](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3900105)
- Gomes, Leonardo Augusto de Vasconcelos, Facin, Ana Lucia Figueiredo, & Salerno, Mario Sergio. (2021). Managing uncertainty propagation in innovation ecosystems. *Technological Forecasting and Social Change*, 171, 120945.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75-105. <https://doi.org/10.2307/25148625>
- Holland, J. (1995). *Hidden Order: How adaptation builds complexity*. Basic Books.
- Hou, Hong, & Shi, Yongjiang. (2021). Ecosystem-as-structure and ecosystem-as-coevolution: A constructive examination. *Technovation*, 100, 102193.
- Huckstep, R. (n.d.). Four Ways Autonomous Vehicles will Change Auto Insurance. *The Digital Insurer*.
- Jackson, M. C. (2019). *Critical systems thinking and the management of complexity*. John Wiley & Sons.
- Karp, L., & Kim, R. (2017). *Insuring Autonomous Vehicles. An \$81 Billion Opportunity Between Now and 2025*. Accenture & Stevens Institute of Technology, The Innovation University.
- Lawson, T. (2012). Ontology and the study of social reality: emergence, organisation, community, power, social relations, corporations, artefacts and money. *Cambridge Journal of Economics*, 36(2), 345-385. <http://www.jstor.org.gate3.library.lse.ac.uk/stable/24232451>
- Lekkerkerk, Pieter. (2023). Letter: Another reason insurance has proved technology-shy. *Financial Times*. <https://www.ft.com/content/c4c9dbac-67cd-4ef2-b6e1-769416bb6245>
- Lenfle, S., & Midler, C. (2009). The launch of innovative product-related services: Lessons from automotive telematics. *Research Policy*, 38(1), 156–169.
- Lichtenstein, Benyamin B., & Plowman, Donde Ashmos. (2009). The leadership of emergence: A complex systems leadership theory of emergence at successive organizational levels. *Leadership Quarterly*, 20(4), 617–630.
- Lissack, Michael, & Letiche, Hugo. (2002). Complexity, Emergence, Resilience, and Coherence. *Emergence: Complexity and Organization*, 4(3), 72–94. <https://eco.emergentpublications.com/Article/e8e4594d-e1fd-41c8-b310-d77bd91d8039/jats>
- MarketsandMarkets. (2021). *Connected Car Market* (No. AT 5974).
- Maturana, H., & Varela, F. (1980). Autopoiesis: the organisation of the living. In H. Maturana & F. Varela (Eds.), *Autopoiesis and Cognition*. D.Reidel.
- McKinsey & Company. (2014). *Connected car, automotive value chain unbound*. McKinsey & Company.
- Meynhardt, Timo, Chandler, Jennifer D., & Strathoff, Pepe. (2016). Systemic principles of value co-creation: Synergetics of value and service ecosystems. *Journal of Business Research*, 69(8), 2981–2989.
- Moore, James F. (1993). *Predators and Prey: A New Ecology of Competition Harvard Business Review*.
- Morrison, Keith. (2008). Educational philosophy and the challenge of complexity theory. *Educational Philosophy and Theory*, 40(1), 19–34.
- Mumford, E. (2000). A socio-technical approach to systems design. *Requirements Engineering*, 5, 125-133.

- Nambisan, S., & Luo, Y. (2022). *The digital multinational: Navigating the new normal in global business*. MIT Press.
- Nicley, D. L. D., Lazaros, E. J., Truell, A. D., Zhao, J. J., & Davison, C. B. (2020). The Connected Car: A Glimpse into the Future of Transportation. *Issues in Information Systems*, 21(2), 49–56.
- Palmié, Maximilian, Wincent, Joakim, Parida, Vinit, & Caglar, Umur. (2020). The evolution of the financial technology ecosystem: An introduction and agenda for future research on disruptive innovations in ecosystems. *Technological Forecasting and Social Change*, 151, 119779.
- Parker, G., Van Alstyne, M., & Jiang, X. (2017). Platform Ecosystems: How Developers Invert the Firm. *MIS Quarterly*, 41(1), 255.  
<https://doi.org/10.25300/MISQ/2017/41.1.13>
- Phillips, Mark A., & Ritala, Paavo. (2019). A complex adaptive systems agenda for ecosystem research methodology. *Technological Forecasting and Social Change*, 148.
- Pickering, A. (2002). Cybernetics and the Mangle: Ashby, Beer and Pask. *Social Studies of Science*, 32(3), 413-437.
- Pujadas, R., Valderrama, E., & Venters, W. (2024). The value and structuring role of web APIs in digital innovation ecosystems: The case of the online travel ecosystem. *Research Policy*, 53(2), 104931.  
<https://doi.org/https://doi.org/10.1016/j.respol.2023.104931>
- Qiu, Y., Gopal, A., & Hann, I.-H. (2017). Logic Pluralism in Mobile Platform Ecosystems: A Study of Indie App Developers on the iOS App Store. *Information Systems Research*, 28(2), 225-249.  
<https://doi.org/10.1287/isre.2016.0664>
- Quantalyse Belgium, & Schönenberger Advisory Services. (2019). *The automotive digital transformation and the economic impacts of existing data access models* (Technical Report).
- Ralph, Oliver. (2023). Why technology has failed to disrupt insurance. *Financial Times*. <https://www.ft.com/content/d2dff24-8a14-4832-8ea0-1937268849f4>
- Rana, Aradhana, Bansal, Rajni, & Gupta, Monica. (2022). Big Data: A Disruptive Innovation in the Insurance Sector. *Big Data Analytics in the Insurance Market*, 165–183.
- Ritala, Paavo, & Gustafsson, Robin. (2018). Q&A. Innovation and Entrepreneurial Ecosystem Research: Where Are We Now and How Do We Move Forward? *Technology Innovation Management Review*, 8(7), 52–57.
- Roundy, Philip T, Bradshaw, Mike, & Brockman, Beverly K. (2018). The emergence of entrepreneurial ecosystems: A complex adaptive systems approach. *Journal of Business Research*, 86, 1–10.
- Sapir, Jonathan. (2019). Thriving at the Edge of Chaos. In *Thriving at the Edge of Chaos*. <https://doi.org/10.4324/9780429356582>
- Searle, J. R. (2006). Social ontology: Some basic principles. *Anthropological Theory*, 6(1), 12-29. <https://doi.org/10.1177/1463499606061731>
- Seiberth, G., & Gruendinger, W. (2018). *Data-Driven Business Models in Connected Cars, Mobility Services & Beyond* (BVDW Research No. 01/18) (p. 58). BVDW & Accenture.
- Scaringella, Laurent, & Radziwon, Agnieszka. (2018). Innovation, entrepreneurial, knowledge, and business ecosystems: Old wine in new bottles? *Technological Forecasting and Social Change*, 136, 59–87.
- Schatzki, T. R. (2000). The social bearing of nature. *Inquiry*, 43(1), 21-37.

- Schreieck, M., Wiesche, M., & Kremer, H. (2021). Capabilities for value co-creation and value capture in emergent platform ecosystems: A longitudinal case study of SAP's cloud platform. *Journal of Information Technology*, 36(4), 365-390. <https://doi.org/10.1177/02683962211023780>
- Schütz, Florian, Heidingsfelder, Marie Lena, & Schraudner, Martina. (2019). Co-shaping the Future in Quadruple Helix Innovation Systems: Uncovering Public Preferences toward Participatory Research and Innovation. *She Ji*, 5(2), 128–146.
- Searle, J. R. (2006). Social ontology: Some basic principles. *Anthropological Theory*, 6(1), 12-29. <https://doi.org/10.1177/1463499606061731>
- Tilson, D., Lyytinen, K., & Sørensen, C. (2010). Digital Infrastructures: The Missing IS Research Agenda. *Information Systems Research*, 21(4), 748-759.
- Tu, Y.-J. (Tony), Shang, S. S., & Wu, J. (2022). Is your autonomous vehicle as smart as you expected? *Communications of the ACM*, 65(2), 31–34.
- Von Krogh, G., Roos, J., & Slocum, K. (1994). An Essay on Corporate Epistemology. *Strategic Management Journal*, 15(Special Issue), 53-71.
- Wang, P. (2021). Connecting the Parts with the Whole: Toward an Information Ecology Theory of Digital Innovation Ecosystems. *MIS Quarterly*, 45(1), 397. <https://doi.org/https://doi.org/10.25300/MISQ/2021/15864>
- Werder, Karl, & Maedche, Alexander. (2018). Explaining the emergence of team agility: a complex adaptive systems perspective. *Information Technology and People*, 31(3), 819–844.
- Wiener, N. (2019). *Cybernetics or Control and Communication in the Animal and the Machine*. The MIT Press. <https://doi.org/10.7551/mitpress/11810.001.0001>
- Wiggers, K. (2021, August 23). Autonomous car insurance drives new opportunities. *VentureBeat*.
- Wollmann, Dewey, & Steiner, Maria Teresinha Arns. (2017). The strategic decision-making as a complex adaptive system: A conceptual scientific model. *Complexity*, 2017. <http://www.jstor.org.gate3.library.lse.ac.uk/stable/24232451>

# INTERNALIZATION OF DIGITAL TECHNOLOGIES: ADAPTING TO ORGANIZATIONAL INERTIA

**Chuying Ding**

*Durham University, Durham, United Kingdom*

*chuying.ding@durham.ac.uk*

**Spyros Angelopoulos**

*Durham University, Durham, United Kingdom*

*spyros.angelopoulos@durham.ac.uk*

## **Abstract**

*Digital transformation endeavours can become a facade due to inertia, and thus, they are increasingly perceived as failed, giving rise to concerns for practitioners as well as ongoing debates in the extant literature, which highlights the need to better understand the phenomenon. To address that lacuna, we explore the technology adoption process for enabling fundamental change in existing organizational systems in terms of internalization and inertia. In doing so, we conduct an empirical study at a stated-owned engineering organization to derive a process model. Our work contributes to the extant Information Systems literature by elucidating the internalisation process of digital technologies, further providing a systematic framework incorporating perspectives of inertia and internalisation. We discuss the anticipated findings as well as the implications of our work for theory and practice, and we delineate an agenda for future research on the topic.*

**Keywords:** Digital Transformation, Digital Technologies, Inertia, Structure, Process



## 1 Introduction

Digital transformation (DT) endeavours are increasingly treated as strategic imperatives (Angelopoulos et al., 2023; Struijk et al., 2023). The organisational design that should be coupled with such transformations, however, might become stagnant due to organisational inertia (Angelopoulos et al., 2023; Kaganer et al., 2023; Vial, 2021). Such endeavours might not be difficult to be implemented from a technical perspective (Arvidsson et al., 2014; Struijk et al., 2023) but they are likely to become a façade due to inertia (Markus & Robey, 2004; Vial, 2021) and authority (Lino et al., 2022). Inertia is inextricable with business and management system (Sydow et al., 2009). To avoid DT becoming a façade, it is vital for organisations to effectively deal with inertia and embed novel digital technologies in existing business systems (Baiyere et al., 2020; Silva & Hirschheim, 2007; Wessel et al., 2021). We refer here to the process of embedding digital technologies into daily business and management activities while changing existing ones as internalisation. The research question we address here, thus, is:

*How can organisations internalise digital technologies in the face of inertia?*

To explore our research question, we follow the cumulated body of literature conceptualizing DT as an ongoing process (Angelopoulos et al., 2023), where digital technologies become an intrinsic part of the organisational system. In doing so, we bring forward two key theoretical contributions: i) we elucidate the process of internalisation of digital technology, and ii) we provide a systematic framework incorporating perspectives of inertia and internalisation.

## 2 Theoretical Background

The literature has highlighted the importance of incorporating digital technologies into existing organisational pathways for technology adoption (e.g., Rajagopal, 2002). Research on technology adoption has long been dominated by the Technology Acceptance Model (Davis, 1989), Innovation Diffusion Theory (Rogers, 1995), and Theory of Planned Behaviour (Ajzen, 1991). This line of research, however, usually addresses technology adoption from the consumer perspective (e.g., Straub, 2009), and refers to evaluation research (e.g., Lai, 2017). To explore technology adoption within and around organizational settings (Salahshour Rad et al., 2018), we introduce the concept of internalisation and develop a mid-range process theory.

Despite that the concept of internalisation has been insinuated in prior studies, it generally refers to specific activities or practices (e.g., Noesgaard et al., 2023). Some studies, for

instance, elaborate different stages of information system (IS) implementation and relevant activities (e.g., Rajagopal, 2002), whereas the implementation stages merely involve linear relationship, which might fail to present the complex and complicated DT empirical reality, as well as its dynamic nature. We suggest, thus, that there is a need to move beyond treating internalisation as a static point regarding specific activities or practices and incorporate a process theory perspective to elucidate its generative mechanisms. By doing so, we can unpack the nature of the digital technology adoption in different stages and can have an overview of involved interlacing forces regarding how DT is impelled, encouraged, or resisted from the perspective of various actors.

Further, the concepts of ‘actual adoption’ and ‘intention to adopt’ can be ambiguous in IS studies (Salahshour Rad et al., 2018), whereas the concept of internalisation helps distinguish the two. Prior studies have proposed concepts relating to ‘actual adoption’ (e.g., Salahshour Rad et al., 2018) referring to improvements in organisational performances by changing business system, and reconfiguring processes and culture (Hammer & Champy, 2009), or redefining organisational identity and value proposition (Wessel et al., 2021). On the other hand, studies capturing concepts concerning ‘intention to adopt’ refer to forward-looking statements regarding the intent of organisations (e.g., Arvidsson et al., 2014), expecting that DT can bring changes to business system and enhances organisational performance. Such a conceptualization, however, can be ambiguous (Arvidsson et al., 2014; Salahshour Rad et al., 2018), which might bring forward the risk of rendering the findings of such research as opaque.

Concurrently, prior studies consider inertia as a key barrier to DT, especially when it contrasts with existing organisational design and other institutional elements (e.g., Cooper, 1994; Kaganer et al., 2023). Consistently, organisational inertia reflects the degree of stickiness of organisational structure meanwhile transformation signifies overcoming organisational inertia to realign the organisation with a new environment (Besson & Rowe, 2012). It is, thus, vital to embed digital technologies in business and managerial activities by dealing with inertia, which acts as a shield between internalisation and DT becoming a façade.

### **3 Methodology**

We conducted an in-depth longitudinal case study at CEG, a Chinese SOE with over 10,000 employees. CEG is upgrading its original integrated information technology (IT) services to a new generation of integrated intelligent ones. The current DT stage for CEG is to develop an

integrated digital platform with all the operational databases from the various Functional and Operational departments. The platform design has incorporated input from the Functional and Operational departments based on their needs. The DT started in October 2021 and is ongoing. Since August 2022, CEG has established the digital platform technically. In doing so, managers and employees can have an overview of the information from all departments and can easily extract or synthesize the information they need. Currently, the DT endeavour is mainly conducted by the Functional departments, while the Operational departments are catching up.

One of the authors was at CEG on average three days a week for three months, and had access to more than 1000 organisational documents, participated in various meetings, and had access to the platform covering all operational and managerial data. Our collected data include field notes, meeting minutes, interviews, strategic directives, news releases, photos documentation reviews and participant observations. We have conducted 39 semi-structured and unstructured interviews (average of 56 minutes). The participants had a variety of backgrounds and different positions. We applied a story-telling approach to the interviews, allowing participants to share their working experiences freely (Czarniawska, 2004). We coded the data in NVIVO and specifically focused our analysis on the internalisation process of digital technology and inertia.

## **4 Discussion**

Our findings indicate that internalisation entails systematically embedding digital technologies into routines to help with daily business or management activities. Examples of internalisation refers to analysing the data on a digital platform to trace and solve operational problems or simplifying a business process. Generally, antecedents of internalisation practices include fragmentation, inculcation and reinforcement, familiarization and habituation, routinization. Internalisation involves mechanisms of routinization and institutionalization. Noticeably, antecedents of internalisation not only involve aforementioned practices, but refer to a continuous and recursive process during the implementation and adoption of digital technologies. At the initial stage, the internalisation is enabled by the top leader's instruction or authority. Then the sub-leaders divide the top management instruction into small pieces (fragmentation) so that employees can execute the tasks regardless of understanding what they are doing. Meanwhile, the top management holds regular meetings to inculcate the imperative of DT and pushes employees to report their DT progress (inculcation). The enforced everyday use of digital technology reinforces a digital mindset and relevant behaviours (reinforcement)

during the execution stage and enables employees to gradually get used to the technology and overcome inertia (familiarization and habituation).

Going through the aforementioned process recursively, digital technologies internalise via the mechanisms of institutionalization and routinization. Institutionalization embodies in a top-down process, in which the top management collects feedback from the employees to modify existing regulations and institutions, and then go back to manage the employees. Inversely, formed based on the operational needs of employees, routinization arises from a bottom-up process, in which the employees modify their daily activities in terms of the changes brought by the digital technologies, and the changes gradually become a routine. Two kinds of routinization occur in the internalisation process: proactive and reactive. Proactive routinization refers to the employees spontaneously adopting a digital technology, during which they usually actively propose their needs regarding DT to the top management. Proactive routinization occurs when the employees find that a novel digital technology can help them be more efficient. We find that inertia is weak in the proactive routinization, as the employees have a propensity to change their behaviours. Conversely, reactive routinization occurs when the top management requires employees to execute DT related instructions, or the employees have to change their routines owing to the restriction of IT requirement. Pushed by daily operations and authority, reactive routinization usually occurs during strong inertia, as employees are pushed to break out of their comfort zone. As a result, despite reactive routinization bringing changes to the existing system, it gives rise to passive resistance from the employees. However, reactive routinization usually acts as an enabler for proactive routinization.

In conclusion, internalisation process entails systematically embedding digital technologies into routines to help with daily business or management activities. Antecedents of internalisation practices include fragmentation, inculcation and reinforcement, familiarization and habituation, routinization, and the generative mechanisms of internalisation process involves routinization and institutionalization. Internalisation appears to be impelled by highly centralized authority from the top management, whereas the substantial internalisation process in which technology is adopted occurs based on employee's operational needs. Noticeably, instead of being a linear process, internalisation is recursive as the top management impels the technology to employees, while the employees adopt the technology and embed the technology into management systems restricting managerial activities recursively. Further, responding to

calls for multilevel DT research (e.g., Appio et al., 2021), internalisation refers to a multilevel process, as the generative mechanisms behind the internalisation-inertia tension occur from the individual level covering the top management and employees, and then it is dealt with by organisational level activities involving regulations and institutions.

Our research brings forward two key theoretical contributions: i) we elucidate the process of internalisation of digital technologies, and ii) provide a systematic framework incorporating perspectives of inertia and internalisation. Our study is in line with the recent IS research agenda (Struijk et al., 2022), and can provide sorely needed insights for effectively internalising digital technologies. By coining the concept of internalisation, we present a novel perspective for appreciating technology adoption from a managerial perspective and contribute to the delineation of the concepts regarding ‘actual adoption’ and ‘intention to adopt’ in IS research. In addition, we unearth the full-scale technology adoption process in the face of inertia, covering its antecedents, barriers, and practices.

From the perspective of practitioner, we pave the way toward elucidating how the transformation strategy can be devised to internalise digital technologies into daily business activities. Moreover, our work can enable managers to prevent their DT endeavours from becoming a façade via coping with inertia impeding the internalisation process. Despite authority might result in inertia, it can also be used to deal with inertia by tactfully fragmenting instruction, and impelling changes brought by digital technology into operational routines.

Although we provide novel insight on internalisation in the face of inertia, some limitations need to be acknowledged that concurrently can pave the way for future studies. The first limitation naturally emerges from the research context. CEG is a Chinese SOE, which presents a different organisational setting compared to most western organizations. Compared to western organizations, a Chinese SOE entails more political constraints (Shen et al., 2020), which is likely to lead to alternative inertia and resistance especially when the business requirement contrasts with social responsibility. Moreover, the Chinese cultural background of harmony (Zhang & Zhang, 2013) might also display a different picture of inertia. In addition, the common features of centralized authority and paternalistic leadership (Zhu et al., 2014) imply that managers in Chinese SOEs might have stronger power over employees than managers in western organizations, which signifies that managers in a Chinese SOE might play a more important role in the internalisation process, as well as in generating and solving inertia. We encourage future studies to further examine how does the internalisation process unfold in

other organisational settings, as well as its differences between western and non-western organisations. Second, inferring from the first limitation, other agents such as institutional elements might impose impacts on the generative mechanism of internalisation process and the inertia. Our research, however, mainly focuses on the effects of actors rather the broader range of agents. We, thus, encourage future research to investigate what roles do other aspects such as institutional agents play in the internalisation process, and their relationship with inertia.

## References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2), 179-211.
- Angelopoulos, S., Bendoly, E., Fransoo, J. C., Hoberg, K., Ou, C. X., & Tenhiälä, A. (2023). Digital transformation in operations management: Fundamental change through agency reversal. *Journal of operations management*, Forthcoming.
- Appio, F. P., Frattini, F., Petruzzelli, A. M., & Neirotti, P. (2021). Digital transformation and innovation management: A synthesis of existing research and an agenda for future studies. *Journal of Product Innovation Management*, 38(1), 4-20.
- Arvidsson, V., Holmström, J., & Lyytinen, K. (2014). Information systems use as strategy practice: A multi-dimensional view of strategic information system implementation and use. *The Journal of Strategic Information Systems*, 23(1), 45-61.
- Baiyere, A., Salmela, H., & Tapanainen, T. (2020). Digital transformation and the new logics of business process management. *European Journal of Information Systems*, 29(3), 238-259.
- Besson, P., & Rowe, F. (2012). Strategizing information systems-enabled organisational transformation: A transdisciplinary review and new directions. *The Journal of Strategic Information Systems*, 21(2), 103-124.
- Cooper, R. B. (1994). The inertial impact of culture on IT implementation. *Information & management*, 27(1), 17-31.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

- Hammer, M., & Champy, J. (2009). *Reengineering the corporation: Manifesto for business revolution*, a. Zondervan.
- Kaganer, E., Gregory, R. W., & Sarker, S. (2023). A Process for Managing Digital Transformation: An organisational Inertia Perspective. *Journal of the Association for Information Systems*, 24(4), 1005-1030.
- Lai, P. C. (2017). The literature review of technology adoption models and theories for the novelty technology. *JISTEM-Journal of Information Systems and Technology Management*, 14, 21-38.
- Lino, A. F., Aquino, A. C. B. d., & Neves, F. R. (2022). Accountants' postures under compulsory digital transformation imposed by government oversight authorities. *Financial Accountability & Management*, 38(2), 202-222.
- Markus, M., & Robey, D. (2004). Why stuff happens: Explaining the unintended consequences of using IT. *The past and future of information systems*, 61-93.
- Noesgaard, M. S., Nielsen, J. A., Jensen, T. B., & Mathiassen, L. (2023). Same but different: Variations in reactions to digital transformation within an organizational field. *Journal of the Association of Information Systems (JAIS)*.
- Rajagopal, P. (2002). An innovation—diffusion view of implementation of enterprise resource planning (ERP) systems and development of a research model. *Information & management*, 40(2), 87-114.
- Rogers, E. M. (1995). Lessons for guidelines from the diffusion of innovations. *The Joint Commission journal on quality improvement*, 21(7), 324-328.
- Salahshour Rad, M., Nilashi, M., & Mohamed Dahlan, H. (2018). Information technology adoption: a review of the literature and classification. *Universal Access in the Information Society*, 17, 361-390.
- Shen, J. H., Zhang, J., Lee, C.-C., & Li, W. (2020). Toward a theory of internal governance structure of China's large SOEs. *Journal of Asian Economics*, 70, 101236.
- Silva, L., & Hirschheim, R. (2007). Fighting against windmills: Strategic information systems and organisational deep structures. *MIS quarterly*, 327-354.

- Straub, E. T. (2009). Understanding technology adoption: Theory and future directions for informal learning. *Review of educational research*, 79(2), 625-649.
- Struijk, M., Angelopoulos, S., Ou, C. X., & Davison, R. M. (2023). Navigating Digital Transformation Through an Information Quality Strategy: Evidence From a Military organisation. *Information Systems Journal*, 33(4).
- Struijk, M., Ou, C. X., Davison, R. M., & Angelopoulos, S. (2022). Putting the IS back into IS research. *Information Systems Journal*, 32(3).
- Sydow, J., Schreyögg, G., & Koch, J. (2009). Organizational path dependence: Opening the black box. *Academy of Management Review*, 34(4), 689-709.
- Vial, G. (2021). Understanding digital transformation: A review and a research agenda. *Managing Digital Transformation*, 13-66.
- Wessel, L., Baiyere, A., Ologeanu-Taddei, R., Cha, J., & Blegind-Jensen, T. (2021). Unpacking the difference between digital transformation and IT-enabled organisational transformation. *Journal of the Association for Information Systems*, 22(1), 102-129.
- Zhang, Z., & Zhang, M. (2013). Guanxi, communication, power, and conflict in industrial buyer-seller relationships: Mitigations against the cultural background of harmony in China. *Journal of Business-to-business Marketing*, 20(2), 99-117.
- Zhu, C. J., Zhang, M., & Shen, J. (2014). Paternalistic and transactional HRM: The nature and transformation of HRM in contemporary China. In *Whither Chinese HRM?* (pp. 22-40). Routledge.



# THE DELEGATION PARADOX: A STUDY ON HOW UNCERTAINTY IMPACTS EXPERTS' WILLINGNESS TO DELEGATE DECISIONS TO ALGORITHMIC SOLUTIONS

**Stefanie H. Gante**

*Durham University, United Kingdom  
Stefanie.H.Gante@durham.ac.uk*

**Spyros Angelopoulos**

*Durham University, United Kingdom  
Spyros.Angelopoulos@durham.ac.uk*

## **Abstract**

*The increased adoption of algorithmic solutions into existing pathways provides opportunities for freeing up experts' capacities and improving outcomes by delegating decision-making to such solutions. The literature, however, paradoxically suggests that experts are reluctant to delegate low-uncertainty decisions to algorithmic solutions, while their willingness to delegate increases for high-uncertainty ones. We refer to this phenomenon as the 'delegation paradox' and we attempt to further unpack it by drawing upon the theories of algorithm aversion and algorithm appreciation in the frame of a two-phase study, incorporating interviews and experiments. The anticipated findings of our study will contribute to the literature on algorithm aversion and algorithm appreciation and bear sorely needed insights for practice on human-algorithm interactions, as well as on decision delegation in the face of uncertainty.*

**Keywords:** *human-algorithm interactions, decision delegation, algorithm aversion*

## Introduction

The proliferation of algorithmic solutions<sup>1</sup> has enabled the automation of decision-making, particularly on clearly defined, low-uncertainty decisions. Such decisions, however, are the ones that experts are reluctant to delegate to algorithmic solutions due to i) an overestimation of their abilities (Fügener et al., 2022), ii) reluctance to accept support (Logg et al., 2019), and iii) belief that expected outcomes are better when they make such decision themselves (DeStefano et al., 2022). Concurrently, experts are less reluctant to delegate decisions to algorithmic solutions when they experience high uncertainty (*ibid*). In this context, we define experts as individuals who are financially compensated for providing valuable insights, solving complex problems, and making educated decisions in their area of specialization. The willingness of experts to delegate decisions to algorithmic solutions, thus, seems to be increased for high-uncertainty over low-uncertainty ones, which stands in contrast with our theoretical understanding as well as rational intuition. Such a counterintuitive phenomenon, which we coin here as the *delegation paradox*, bears valuable implications for the extant Information Systems (IS) theory (Struijk et al., 2022) and practice (Davison, 2022). Failing to address this phenomenon, might lead to missed opportunities for improving expected outcomes in decision-making under uncertainty, as well as further elucidate how contextual factors can facilitate successful human-algorithm collaboration and alleviate issues related to algorithm aversion. To further unpack this phenomenon, we specifically focus on the underlying relationship between *decision-uncertainty* and the willingness of experts to delegate a decision to an algorithmic solution. In doing so, the research question that we address here, is:

*How does uncertainty affect experts' willingness to delegate decisions to algorithmic solutions?*

To provide an answer to our research question and contribute to the extant IS literature on the topic, we draw upon the theories of *algorithm aversion* and *algorithm appreciation* to provide a mid-range theory with direct applicability in the specific situation. In doing so, we further bring our focus onto the healthcare sector, where the decisions that clinicians can delegate to algorithmic solutions have a clear spectrum and distinction from low- to high-uncertainty ones, with varied implications for the quality of healthcare provision as well as the lived experience of patients. In the frame of a two-phase study, we first acquire insights on clinicians' willingness to delegate a decision to algorithmic solutions through interviews, and then confirm such insights through

---

<sup>1</sup> We define algorithmic solutions as computational applications that are used for the execution of well-defined tasks.

experiments. Understanding how clinicians delegate decisions to algorithmic solutions can enable us to further develop strategies for fostering the use of such solutions and identify improper use for decisions that may have undesired implications for medical decision-making. The expected findings of our study can contribute to the IS literature on algorithmic solution decision-making, as well as the future of work, and has important implications for research and practice on the topic.

The rest of this paper is organised as follows. First, we present the theoretical background on the topic, and we then introduce the context and methodological approach of our study. In the penultimate section, we discuss the implications of the anticipated findings for both theory and practice, while we conclude the paper by delineating an agenda for future research on the topic.

## **1 Theoretical Background**

The proliferation of algorithmic solutions across industries calls for a better understanding of the factors affecting attitudes towards such solutions (Gante & Angelopoulos, 2023; Mikalef et al., 2022), bringing to the fore the concept of *algorithm aversion*. Algorithm aversion is conceptualised as the preference to delegate decisions to a human rather than an algorithmic solution, despite the probability of obtaining better results when delegating to the latter (Dietvorst et al., 2015; 2018). Such aversion appears to be dependent on the type of decision that is being delegated to an algorithmic solution and is increased for decisions that require empathy, while it decreases for decisions that call mainly for cognitive skills (Castelo et al., 2019). In contrast, recent studies investigating *algorithm appreciation* suggest that individuals prefer delegating a decision to algorithmic solutions under specific conditions, however, their level of expertise on the topic of the decision can cause a generic aversion to delegate it altogether (Logg et al., 2019). Concurrently, individuals with increased expertise tend to overestimate their abilities and consequently do not delegate decisions adequately (Fügner et al., 2022). Counterintuitively, in the context of high decision-uncertainty, aversion to delegate to algorithmic solutions decreases significantly, even overriding previous concerns such as low explainability of decision-making (DeStefano et al., 2022), suggesting that experts rely on their intuitions unless they cannot decide with high certainty. These insights from the literature inform our theorization of the *delegation paradox*, which implies the preferred delegation of high-uncertainty decisions over low-uncertainty ones, contrary to our understanding of task-dependent algorithm aversion. Experts' willingness to delegate to an

algorithmic solution, therefore, appears to vary based on i) the expertise of the delegator, ii) the nature of the decision, and iii) the perceived uncertainty regarding the decision.

## **2 Methodology**

We incorporate a two-phase approach in a bifold data collection process, which includes semi-structured interviews in Phase 1 and experiments in Phase 2, recruiting clinicians as medical experts for both stages. The first phase will enable us to i) gain deeper insights into the delegation paradox, ii) explore the underlying relationships without preconceptions, and iii) provide a space for the participants of our study to further share their personal experiences. We believe clinicians are particularly suitable for our data collection, as the decisions they can delegate to algorithmic solutions have a clear spectrum and distinction from low- to high-uncertainty ones, with varied implications for the quality of healthcare provision. The insights from the first phase will feed into the design of the second phase, during which we will experimentally test our research model and respective hypotheses. Such an approach will also allow us to uncover potential inconsistencies between expressed intentions and the actual behaviour of experts. Given their inherently different healthcare systems that will provide more diverse insights, we target the UK and the US for both rounds of data collection. In both markets, we have established connections to healthcare providers, facilitating participant recruitment. More specifically, during the first phase, the semi-structured interviews will focus on generating deeper insights into clinician's decision-making processes and delegation strategies in a spectrum of low- to high-uncertainty. The interviews are expected to last 30 minutes on average and will be mainly conducted online. The interviews will be transcribed and analysed in NVivo 14, using thematic analysis. The participants will be recruited through an existing network of clinicians involved in the use of algorithmic solutions for medical decision-making. The participants of Phase 1 will be also invited for Phase 2, while snowball and convenience sampling will be incorporated to ensure adequate participation in our experiments. The experiments will provide scenarios where participants will be given the choice to either decide themselves or delegate the decision to an algorithmic solution. Our methodological approach and data collection will abide by the ethical guidelines of the concerned institutional review board.

## **3 Expected Implications**

Our anticipated findings can contribute to the contrasting literature streams on algorithm aversion and algorithm appreciation, span their boundaries, and address existing disparities. In doing so,

our work will provide insights into the topic of human-agent interaction (e.g., Baird & Maruping, 2021), which is a burgeoning line of research in the extant IS literature. Furthermore, we expect to shed light on a phenomenon we coin as *delegation paradox* and uncover potential inefficiencies as well as risks of inadequate use of algorithmic solutions in clinical settings. Based on the relationships we propose, we expect that experts are more willing to delegate decisions with high perceived decision-uncertainty, which complements our theoretical knowledge of human-algorithm interactions (e.g. Fügener et al, 2022; Logg et al., 2019) and carries relevant ethical and practical implications for practice. While the use of algorithmic solutions in high decision-uncertainty contexts may inspire innovative approaches for problem-solving, algorithmic solutions perform best for well-defined, low uncertainty tasks. Unbased high confidence in their performance may, therefore, cause inappropriate use, implying significant consequences for outcomes as well as for the long-term use of such solutions. Concurrently, not using algorithmic solutions in low decision-uncertainty contexts implies missed opportunities to increase performance outcomes, free up capacities and decrease costs. Our expected findings bear timely implications for practice in paving the way toward the successful integration of algorithmic solutions into clinical pathways. Due to the increased diffusion of algorithmic solutions in all industries, our study will generate important insights that go beyond the healthcare industry on how to push for successful collaboration between humans and algorithms in general.

#### **4 Limitations and Future Research**

Our research design implies some limitations, which offer opportunities for future research. First, our data collection targets the US and UK markets. We chose these geographical areas, as they have two very different healthcare systems, suggesting potential implications for findings. Nevertheless, we accept that our findings might not be generalisable for all countries that could have different access to resources, political structures, and cultural backgrounds. A replication of our study in varying settings, thus, would improve the robustness of our findings for other contexts. Additionally, the use of algorithmic solutions for medical decision-making brings forward significant societal implications, which can give rise to ethical debates. Consequently, there is a possibility that the anticipated findings of our study might be influenced by social-desirability bias (Chung & Monroe, 2003), which would incentivise clinicians to depict their opinions and behaviours in a socially desirable manner, rather than in a factually representative one.

Furthermore, our chosen methodological approach does not address actual, but hypothetical behavioural choices, which might suggest a limitation to the generalisability of our findings. Future research, therefore, could address such a limitation by negotiating access to a real setting, where clinicians are called to choose whether they will delegate a decision to an algorithmic solution.

## 5 Conclusion

Our study addresses a timely topic on human-algorithm interactions: the use of algorithmic solutions by experts in the face of decision-uncertainty. In the frame of a two-phase research design, we anticipate generating insights on experts' willingness to delegate decision-making to algorithmic solutions, and in doing so, shedding light on a counterintuitive phenomenon emerging from the literature that we refer to as the *delegation paradox* and remains yet to be investigated. Our expected findings entail important implications for both IS theory as well as practice.

## References

- Baird, A. and Maruping, L.M., 2021. The Next Generation of Research on IS Use: A Theoretical Framework of Delegation to and from Agentic IS Artifacts. *MIS Quarterly*, 45(1), 315-341.
- Castelo, N., Bos, M. W., & Lehmann, D. R., 2019. Task-Dependent Algorithm Aversion. *Journal of Marketing Research*, 56(5), 809–825.
- Chung, J., & Monroe, G. S. (2003). Exploring social desirability bias. *Journal of Business Ethics*, 44, 291–302.
- Davison, R. M. (2022). Impact and implications for practice. *Information Systems Journal*, 33(2), 187-191.
- DeStefano, T., Kellogg, K., Menietti, M. and Vendraminelli, L., 2022. Why providing humans with interpretable algorithms may, counterintuitively, lead to lower decision-making performance. *MIT Sloan Research Paper*, 6797.
- Dietvorst, B.J., Simmons, J.P. and Massey, C., 2015. Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 1–14.
- Dietvorst, B.J., Simmons, J.P. and Massey, C., 2018. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, 64(3), 1155–1170.
- Fügener, A., Grahl, J., Gupta, A. and Ketter, W., 2022. Cognitive challenges in human–artificial intelligence collaboration: Investigating the path toward productive delegation. *Information Systems Research*, 33(2), 678–696.
- Gante, S. H., & Angelopoulos, S., 2023. *Information Overload in the Age of Algorithmic Solutions: The Effect of Patients' Information Processing Capability on Their Perceived Quality of Healthcare*. UKAIS Conference Proceedings 2023, Kent, UK.

- Logg, J.M., Minson, J.A. and Moore, D.A., 2019. Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90–103.
- Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A., 2022. Thinking responsibly about responsible AI and ‘the dark side’ of AI. *European Journal of Information Systems*, 31(3), 257–268.
- Struijk, M., Ou, C. X., Davison, R. M., & Angelopoulos, S. (2022). Putting the IS back into IS research. *Information Systems Journal*, 32(3), 1–4.

# The impact of the digitalisation of care on older, unpaid carers

Author Name: redacted for anonymity/peer review

## Abstract

The digitalisation of care involves the integration of technology and digital tools into social care and healthcare services (Hamblin, 2020). The present paper examines the impact of digitalisation of care on unpaid older carers in the UK and explicates the preliminary findings of our research. We employed a focus group that took place in Liverpool, including 13 older, unpaid carers and 1 carers' centre worker. The data were analysed through a constructionist thematic analysis approach. Participants constructed digitalised care as alienating. They delved into interoperability challenges and expressed concerns regarding online safety. Carers raised the financial burdens of caregiving, underscoring the need for increased funding, the necessity of inclusive training and the impact of inadequate policies on disabled individuals' societal participation. Lastly, participants stressed the importance of co-production in research and policy. We discuss these findings within the socio-political context and provide a brief overview of the study's future implications and limitations.

**Keywords:** Care, Unpaid; Older; Carers; Digitalisation; UK

## 1.0 Introduction

The global ageing population raises concerns in healthcare and social care, with a growing number of individuals aged 65 and above facing unmet care needs (Wright & Hamblin, 2023). This has led to a rising demand for long-term care, including informal long-term care, involving unpaid caregivers, becoming crucial in addressing these demands (Cylus et al., 2019; Tur-Sinai, 2020). Unpaid care is a significant societal phenomenon, with 7.3 million caregivers in England providing care valued at over £100 billion annually (Wright & Hamblin, 2023). However, cutbacks in social care, austerity and neoliberal ideologies, which encourage cutting public expenditure and assigning the responsibility for the lack of a welfare state to individuals, are placing added burdens on unpaid caregivers (Pearson, 2019).



In response to these challenges, carers centres in the UK have emerged as crucial support structures, filling the void left by insufficient funding. These centres offer invaluable resources such as free advice, emotional support, training, and various therapies for carers (Smith et al., 2003).

The UK's approach to care diverges from that of other European countries. The Adult Social Care (ASC) policies, legislation, and funding distribution are devolved to the four nations within the UK. Implementation of ASC services is managed by local councils, numbering 152 in England, 22 in Wales, 32 in Scotland, and five health and social care trusts in Northern Ireland (Hamblin, 2020). This decentralised model reflects regional differences, which are often accentuated by the challenges imposed by austerity.

Against this backdrop, the increasing care demands and prevailing austerity coincide with rapid technological advancement aiming at cost reduction and efficiency (Hirvonen et al., 2022), which has led to the digitalisation of care. Digitalisation involves the increased use of digital interfaces and automated services, mediated by algorithms, in public administration interactions, with some instances replacing human contact. It spans from simple paper-to-digital transitions to artificial intelligence (Sætra & Fosch-Villaronga, 2021). Key technologies driving the digitalisation of care include e-health, service automation and monitoring, applications, gadgets and automated helplines (Pekkarinen & Melkas, 2017).

Past research has explored how smart technologies can ameliorate caregiving, enhancing safety and independence for older individuals (Damant et al., 2020; Parzen et al., 2021). Governmental and independent services promote integrating information and communication technology (ICT) into social care and healthcare. **However, such initiatives are hindered by implementation challenges such as low interoperability (Hamblin, 2020), which in the present context is defined as a disconnect in information sharing across various services (Kubicek & Cimander, 2009).** This is further pronounced by challenges such as lack of internet access, safety standards, data security, service safeguarding issues and more (Hamblin, 2020; Wright & Hamblin, 2022). For example, Parzen et al. (2021) found that informal caregivers use technology to facilitate their independence and socialisation but face barriers like lack of

knowledge and technology literacy. Similarly, Damant et al. (2020) raised concerns regarding device accessibility, support availability, emerging technologies, and difficulties accessing online healthcare appointments and in-person options, with such factors raising concerns regarding digital exclusion.

Digital exclusion poses significant challenges for vulnerable groups. This systemic issue, marked by cost, skills, and rural disparities, hinders technology adoption for carers (Damant et al., 2020). This issue becomes more pronounced for older adults living with sensory impairments, disabilities, and cognitive difficulties, especially during the cost-of-living crisis (Hirvonen et al., 2022). Digital exclusion limits access to services and goods that can enhance quality of life. Given that older individuals tend to face more digital exclusion and unpaid carers tend to be older than paid carers (Lee et al., 2022), there is growing policy and scholarly interest in the unpaid work of older individuals (Cylus et al., 2019), raising questions about their digital exclusion.

While some qualitative studies have examined telecare's role for unpaid caregivers, much of the literature focuses on carers of people with dementia (e.g., Lorenz et al., 2019; Sriram et al., 2019). Scholars call for research addressing the barriers unpaid caregivers face in using technology, emphasising the need for older people's involvement in the co-production of services that meet their needs (Spann et al., 2019). However, broader research on unpaid care and digitalisation is limited. Quantitative and mixed-method studies have explored carers' technology perceptions, but there is insufficient evidence regarding the impact of new technologies on unpaid care; that is especially the case for older, unpaid carers (Catalyst, 2016; Egan et al., 2022; Pickard, 2015; Shaw et al., 2020).

Indeed, the literature pronounces the need for richer, contextual insights, to understand how older individuals use digital technologies, especially in later life (Neves & Mead, 2021). This is particularly crucial for older, unpaid carers, who remain significantly under-researched. Understanding the impact of digitalisation on older unpaid carers has practical implications for avoiding digital exclusion. Policies should address the challenges of implementing technology in care for older unpaid caregivers. This highlights the importance of interdisciplinary research approaches and a deeper exploration of the digital agency of older individuals (Spann et al., 2019). In line with

Damant et al.'s (2020) research, carers have called for collaborative efforts between ICT developers and carers, emphasising the need for a holistic approach to the digitalisation of care.

To address current academic and policy concerns, this study aims to investigate how the digitalisation of care affects older unpaid carers. Our research question is "What is the impact of the digitalisation of care on older unpaid carers?". **Our research has been conducted in the UK, due to its unique care context and emerging care-related technologies.** In this paper, we present preliminary findings from our study. More specifically, we analyse and explicate a focus group conducted in a carers centre in Liverpool. In the following sections, we present the methodology we employed, further covering the use of constructionist thematic analysis, study design rationale, sample selection, and our analytical process. We then discuss our preliminary findings, highlighting six themes relevant to the digitalisation of care and how carers construct these themes. Finally, we explore the novelty of our findings by connecting carers' constructions to the current socio-political and ideological context.

## **2. Methodology**

### **2.1. Study context**

The study's geographical setting holds contextual relevance. The focus groups were conducted in carers centres located throughout the UK, as an attempt to examine the effects of digitalisation while considering the north-south divide. The divide is a concept that emerged in post-war Britain and characterises the bifurcation in contemporary English politics in relation to varying socio-economic conditions that disproportionately affect the northern regions (Baker & Billinge, 2004). **As this paper presents preliminary results, we will be explicating the findings of the first focus group we conducted, which took place in Liverpool.**

### **2.2. Study design**

For the present study, we conducted interactive group interviews (focus groups) that collaboratively construct meaning and shed light on how inequalities are shaped through discussions and "normative discourses" regarding a specific topic (Smithson, 2000). We followed the interviewing guidelines of Turner III (2010), employed open-ended questions and emphasised that participants could prioritise their lived experiences. We framed the questions in an unbiased manner and included follow-up questions. Due to the sensitivity of the topic and the potential emergence of identifying information, the three researchers took notes of what was said instead of audio recording devices.

Story completion, valuable in constructionist research, was employed following the guidelines of Clarke et al. (2019), balancing imagination stimulation and ambiguity. Authentic scenarios, drawn from research reports and literature case studies, highlighted digital inclusion barriers and facilitators including digital poverty, digitalisation policies, and online safety (Braun et al., 2018). Story stems included situations like caring for a neighbour with dementia while facing online service challenges due to affordability, an individual seeking to enhance their understanding of data safety, and a policy-maker aspiring to create initiatives tailored to the specific needs of carers. To maintain deliberate ambiguity, we used gender-neutral names and incorporated both first-person and third-person story stems, which initiated conversations; yet participants primarily reflecting on personal experiences (Braun et al., 2018; Clarke et al., 2019). The focus group in the present study lasted for 2 hours.

### **2.3. Participants, recruitment and sampling technique**

The recruitment was conducted in August-September 2023 and was facilitated by a research partner with expertise in care-related policy and research. Through their professional network, they established connections with carers centres, which informed potential participants. This outreach strategy involved displaying study posters in the carers centre and during their coffee days, creating awareness about the research opportunity.

The inclusion criteria were set to encompass individuals aged 50 and above with current or past caregiving experience who needed to use any form of technology in their caring

role. Notably, specific levels of digital literacy were intentionally left unspecified, reflecting an aim to capture a diverse range of digitalization experiences among participants.

The sample for the focus group consisted of 13 carers, aged 55 and above, and an accompanying carers centre worker who was required to be present for safeguarding reasons. It has been suggested that 12 participants constitute a substantial sample size for a focus group. To ensure varied perspectives, we aimed for 14 participants per focus group, considering potential cancellations (Lazar et al., 2017). Out of the 13 carers, 4 were male, and 9 were female. All participants were UK residents, currently providing unpaid care for one or more individuals or who had previously cared for a loved one.

#### **2.4. Ethical approval**

Prior to the data collection we had secured ethical approval from the University of Sheffield, ethical application number (redacted for anonymity) following the procedures suggested by institutional guidelines and UK legislation, taking a “belt-and-braces” approach due to the multiple layers of sensitivity presented by the topic.

#### **2.5. Analytic procedure**

For the present study, we employed a constructionist thematic analysis. Care, both as lived experiences and in the context of societal structures and apparatuses, is a complex topic, especially when studying its relationship with digitalisation (Braun and Clarke, 2006; Charitsis & Lehtiniemi, 2023). Latent approaches in this study focus on exploring the construction of ideological subjects, involving in-depth and rigorous analysis with discursive elements. This approach also allows for collaboration and participation among multiple researchers (Braun & Clarke, 2019; Majumdar, 2022).

Following Braun and Clarke's (2006) guidelines, the researchers immersed themselves in the data. Three researchers collected data at the carers centre by taking notes during the focus groups, noting the participants' suggestions and relevant topics discussed. Each researcher's notes were tailored to their respective emphases. The first researcher focused on constructionist and discursive elements. The second and third researchers

provided descriptive notes with incident examples. The first author synthesised these notes, identifying both similarities and differences in the note-taking approaches, which they then discussed with the research team. This synthesis aimed at fostering an enriching analysis grounded in the constructionist epistemology, which acknowledges multiple constructions rather than a singular reality (Reid et al., 1996).

They then engaged in active immersion by repeatedly reviewing the collected data to identify codes and themes, documenting their observations, and organising the data into groups. Given the nuanced nature of the analysis, manual coding was employed, involving an examination of discursive elements. After three rounds of analysis, sub-themes were organised, and researchers discussed the generated themes. In the second sub-stage, themes were re-evaluated within the entire dataset. The research analyst then named and refined the themes. Throughout the analysis and write-up, we considered questions such as the implications of the themes and potential justifications of participants' discursive constructions (Majumdar, 2022).

### **3. Results**

In this section, we present the themes we explicated in the dataset. We present the feelings arising from digital transition and carers' struggles with techno-institutional lack of interoperability. We then discuss their concerns regarding online safety and cost of technology. We tackle their legitimisation of the need for funding, as well as their concerns over inclusivity and ableism. We conclude by discussing policy suggestions as outlined by carers themselves, emphasising the need for training and co-production in research and policy.

#### **3.1. Feelings of exclusion, abandonment and alienation**

A prominent theme was the challenges that caregivers face regarding the digitalisation of care. While everyone had access to devices, the vast majority reported being unable to use them effectively for caregiving purposes (*“what do I do with all of those”*). Participants emphasised the difficulties they encountered when navigating websites related to care or healthcare, describing them as having complicated and inaccessible

interfaces. Moreover, they reported a lack of help, both in person and even in the form of telephone support, which they found easier to navigate compared to the Internet. This complexity was worked up as time-consuming, which was juxtaposed with the 24/7 demands of their caregiving roles. Participants reported being interpellated (Althusser, 2014) to dedicate even more of their time as part of their caregiving labour via their familiarisation with technology. Many carers suggested that there should be some form of assistance or guidance on how to navigate online spaces. Several participants mentioned the Covid-19 pandemic as an excuse for not returning to the way things were before.

Participants frequently mentioned that they did not understand how to navigate the online domain or perform actions within it; such lack of understanding was constructed through emotive language and was often framed as anxiety-provoking. For instance, one participant stated, "*As soon as someone tells me, 'I'll send you the link,' my heart sinks!*". This sentiment was reflected through the collective agreement of several participants.

The participants' constructions regarding the transition online entailed elements of lacking agency/self-determination, with a few employing variations of the word "push" ("*Feel like I am being pushed online*"). Carers often had help from their families to navigate cyberspace but reported that it is challenging to depend on family members as they are not always present. Participants further constructed being positioned/interpellated (Althusser, 2014) as active users, but being unable to exert much control over using technology e.g. "*I don't know how to download an app (...) people do it for me.*" Within the dataset, there was an agent-subject distinction, which further pronounced the lack of agency participants felt. Consider the excerpts below, where technology and vocabulary are constructed as personified:

*"Technology takes things away"*

*"Vocabulary is changing"*

*"This needs to be made easy for us"*

During the focus group, the prevailing sense was that carers no longer feel valued or supported, leading to constructions around exclusion, alienation and abandonment

(*“there’s nobody there”*). Some examples of such constructions can be observed in the excerpts below:

*“It makes you feel more isolated”*

*“All this is going on and you have no idea what is going on”*

*“I’m dyslexic and I panic. I read again and again, and I panic”*

*“Who is gonna care for us?”*

Carers suggested that, at best, they have to use the vast majority of digitalised services with *“no help at all”*, often discussing the lack of helplines, instructions, or one-to-one help. Overall, the dominant constructions of technology were that it is challenging and difficult to operate.

### **3.2. Techno-institutional lack of interoperability**

Another topic often discussed by carers was the issue of interoperability. Carers indicated that, in addition to the practical challenges of navigating the digital realm, there was a **lack of interoperability in healthcare and social care services, where information, such as carer status or healthcare records, is not shared among interacting services**. This concept of interoperability encompassed both technological and institutional aspects. For example, carers mentioned instances where doctors failed to communicate with each other, and in addition to this communication gap, there was a substantial absence of shared digital medical records.

Carers expressed frustration at the lack of access to one-on-one communication, often being passed on to multiple different individuals. They contrasted this with one-on-one communication, which they constructed as more effective and personalised. They suggested that the lack of communication between healthcare staff and lack of shared records lead to substantial waiting times (e.g., medical referrals), despite the people they care for facing life-threatening conditions. Participants suggested that shared documents, perhaps in the form of digital diaries that could contain all their information and medical records, can be particularly helpful. They repeatedly stated that a variety of interventions and services are needed and yet are not joined, which was constructed as further pronouncing the carers’ lack of agency (*“if the surgery can’t do it, what hope do you have”*).



Moreover, participants challenged discourses around individual responsibility, suggesting that with technology, the onus falls on them to find someone to help. They further employed a variety of negative signifiers to work up the lack of interoperability and lack of help, with the strongest one being the employment of the word “*nightmare*”.

### **3.3. Concerns over online safety**

Carers often constructed data safety as a substantial concern, particularly in case their records went missing due to mishandling of their data. Consider the excerpt below:

*“I get worried cause things get lost, all records can go missing”*

Moreover, carers reported the possibility of hacking as concerning (“*I worry about being hacked*”), especially in the scenario of their information being leaked online. Participants also raised that they lack awareness regarding the information they provide on online domains. They suggested that they often had to consent to providing data to services and sites, but they were not sure what this meant; as a result, their sense of consent was contrasted with the lack of knowledge to further showcase that their consent was uninformed and thus questionable. A particular discussion regarding consent was centred around the use of cookies; participants repeatedly stated they lacked knowledge regarding what happened, it was time-consuming to read it and did not understand it.

*“We don’t know what is happening to this information”*

Similarly, participants constructed this difference in “dialects” between the discourse of the sites and their own discourse as “*sinister*”. They constructed a script (Edwards, 1994) where this disparity was employed by the site owners on purpose. What was often implied here was that this lack of awareness was operationalised to further exploit their data; such juxtaposition between their awareness and the awareness of the key stakeholders was constructed as concerning.

*“You can’t understand”*

*“They know people can’t understand”*

*“People know where you live”* (referring to how easily services now find their address)

*“We are saying yes but we don’t really know what it means”* (referring to cookies)

*“I don’t know what cookies are, I just accept it”*

Additionally, participants raised concerns regarding being worried about scams, which they reported as becoming increasingly elaborate and panoptical; in a sense, the scammers were perceived as having more knowledge than the participants themselves. This can be reflected in the following quote through an extreme case formulation:

*“I got scammed last Friday (..) so easily (...) they know more about me than I do”*

### **3.4. “Carers save millions for the government”: the negotiation and legitimisation of the need for further governmental and policy funding**

Carers constructed and positioned themselves as the community members doing the “extra tasks”, intermediaries between the government and policy agents and the people they care for. Yet, despite their labour, they did not face any facilitation of their work or any rewards. Throughout the focus group, they repeatedly tried to negotiate that they are worthy of technological inclusivity and funding. The framing of their own value, as well as the value of their labour, was constructed in market terms, via the financial surplus they create through their labour.

*“Carers save millions for the government“*

They also suggested that digital inclusion saves time and offers them more flexibility:

*“In the end the more savvy we are the more money we save”*

*“We are worth investment”*

Participants also highlighted the benefits of training and supporting them by raising not only concerns about financial impact but also through constructions of cost-benefit analysis. An example of that was a discussion that further legitimised the need for investment by interpellating the government and policymakers, suggesting that lack of funding and welfare policies would impact the health of the participants and result in a financial deficit.

*“We save the government millions of pounds ...the stress we are under... anxiety”*

*“This needs to be made easy for us”*

*“Funding has to be brought somewhere otherwise everything will cost more because our health will be worse”*

Similar arguments were employed for the need to receive training, which was framed as having a trickling down effect, through which the initial carer facilitators would also teach other carers in the community. This construction was also framed via the trickle-down economics construction and as a good investment and money well spent. Despite the marketised discourse, the attempt here is to legitimise the request for more funding. However, this is done through mirroring the discursive terrain, the structure and governance of society due to the ideological hegemony of the ruling class (Gramsci, 2011).

During the discussion, participants mentioned that they would prefer if the carers centre had more availability. This interpellated one of the individuals working at the carers centre to discuss the issues of funding from their own scope. They suggested they could afford to buy laptops and invite an IT tutor but, as the carers centre is a charity, they lack the resources to offer this everywhere and to everyone. This was further pronounced by the demographic differences presented in their borough, which was characterised as significantly large.

### **3.5. Cost of technology**

Carers highlighted the cost of caring for an individual. The financial impact they faced was omnipresent; it was part of the discussions regarding the affordability of training, the need for every person to be looked after as an individual, and the cost of keeping someone well.

When it came to technology, participants had means of communication; however, they brought up the smaller, often invisible costs of technology e.g., getting the computer fixed and software-related expenses for which they relied on their families to pay for (*“I would find it very hard to pay for the internet without my son paying for that”*; *“My family got me my computer”*). They indicated that even if claiming benefits, there are so many other expenses to pay, such as private dentists. In cases of means-tested benefits, those on a lower income can access such initiatives, but individuals on the borderline of such evaluations get nothing, and thus such costs are even more detrimental to them. Moreover, carers discussed that even paid carers would require further funding and initiatives as with the rising inflation, their salaries are particularly

low and thus have no access to such digital equipment. Such issues are further inflated by the cost of living.

These concerns were further exacerbated by potential disabilities, with carers suggesting that the costs are even more enhanced. Despite costs, the accessibility of information and time management were deemed important and thus funding was framed as a significant aspect of digitalisation.

### **3.6. Training**

The need for training was the most prominent theme. As one participant suggested, she would find beneficial technologies like the NHS application, but it is “*messed up*”; yet she would be interested in attending entry-level training as “*there are so many things I would like to be shown*”. Carers suggested they are “*happy to embrace*” technology when it is well designed but they “*would just love someone to go through with me*”. Others suggested that training is “*tricky*” and raised that the time demands of their roles do not allow them multiple hours to familiarise themselves with technology. They highlighted the significance and need for local drop-in centres but suggested that dropping into these centres has substantial time demands they often could not adhere to.

Participants stressed the need for inclusive training, highlighting the challenges faced by those with limited mobility or confinement. They proposed an ideal solution of technicians conducting home visits in such cases, particularly considering disability-related issues. Participants often relied on family members to drive them to training due to their limited mobility. As one participant pointed out:

*“It is not like jumping in the car”*

Carers highlighted the importance of community involvement in training and expressed a wish to ensure that professionals in care-related roles e.g., individuals specialising in social prescribing receive technology-related training so they can, themselves, help carers. The carers centre worker recognised the need for training and devices among older people but suggested that offering local support closer to people's residences is

not always feasible. Carers also stressed the importance of training when devices are provided.

Several examples were provided to illustrate the importance of technological training, such as when transitioning a cared-for individual to a "virtual ward" after hospital discharge. This required carers to use electronic monitoring tools. They compared this need to more basic training like email setup or website access, emphasising that training is a highly personalised and time-consuming task for carers. Thus, carers called for specialised training modules and highlighted the importance of maintaining in-person, face-to-face training options.

### **3.7. Inclusivity and ableism**

A recurring theme was the inadequacy of current policies, tools and technologies in promoting digital inclusion. It was often implied that the existing measures are designed for able-bodied individuals yet many individuals do not fit this assumption. For instance, when consulting with a GP, participants highlighted the challenge of hearing problems, especially when doctors leave messages on patients' phones, as they may not always be able to hear them. Participants also mentioned that individuals with dyslexia may find emails anxiety-provoking.

The financial burden associated with disabilities was another concern, with one participant mentioning their struggle to manage things despite claiming benefits due to their disabilities. Participants also expressed worries about the future of technological development and inclusivity. For instance, they discussed the transition from copper lines to digital technology, highlighting the potentially life-threatening consequences of losing communication during an electricity outage for individuals reliant on medical equipment.

### **3.8. Co-production and co-design in research and policy**

Participants suggested that the current technologies were not designed according to their needs. As one participant stated "*It's like going to a hotel that claims to be disabled-friendly but isn't.*" They highlighted the need for co-production and

collaboration with carers in policy development. They emphasised using inclusive language to employ diverse perspectives within technological policies.

Carers advised policymakers to simplify access to digital medical platforms and pronounced the cost-effectiveness and benefits for both doctors and patients. Participants suggested pilot studies involving carers to explore effective initiatives and make inclusive training programs, such as those offered through GP practices.

Regarding the development of digital tools and safety measures, the participants urged the use of more accessible language and safety checks to prevent scams. They emphasised the value of engaging in discussions with the target group of technology users to inform technology according to their needs. They additionally suggested testing trials of technological initiatives with carers. Overall, participants emphasised the implementation of the involvement of carers throughout the entire process, both during planning and design, to enhance the applicability and benefits of emerging technologies.

#### **4. Discussion and novel contributions of our research**

Our study was the first qualitative UK study to explicate the constructions of unpaid, older carers in relation to digitalisation and contextual issues such as the need for funding, concerns over technology, and feelings of exclusion through a constructionist prism. Moreover, our research offers concrete policy recommendations and can contribute to future policy implementations and research gaps. However, our study also contributes to the current body of knowledge via the theorisation of its findings.

The first contribution we make is the theorisation of the feelings of alienation that participants interpellated. The first theme, "Feelings of Exclusion, Abandonment, and Alienation," underscores the challenges carers face in adapting to digitalised care. Participants discussed their struggle to navigate online platforms independently. The emerging body of research pronounces the challenges older individuals face regarding technologies (Hirvonen et al., 2022). Our study highlights how unpaid, older carers themselves construct these challenges, via the employment of a discursive terrain based

on their lack of agency in cyberspace. We showcase how older, unpaid carers' constructions of technology frame it as difficult to operate and how technological development and the lack of clear guidelines were constructed as alienating. As Carrey and Foster (2013) suggest, the ideological state apparatus reinforces hegemonic actions which position individuals as ideological subjects. Carers are positioned by apparatuses as active technology users despite their lack of familiarisation with cyberspace. They are interpellated (Althusser, 2014) to handle technological changes that they are not familiar with, resulting in both social alienation and alienation from their own labour. These findings reflect Shaw et al. (2020), who studied care organising technologies, highlighted the importance of the political context and criticised the assumption that informal support will supplement/substitute state support.

Our second contribution lies within the theorisation of the techno-institutional lack of interoperability. In the second theme, "Techno-Institutional Lack of Interoperability," participants noted the lack of compatibility between different institutions/services and their technologies. Similar concerns have been raised by Wright and Hamblin (2022) suggesting that digitalisation comes with low levels of data sharing. However, we also provide a novel insight by tackling the experiences of older carers and theorising how they fit in the wider socio-political context, thus showcasing how this lack of technological interoperability is intertwined with the overall structural lack of interoperability in relation to healthcare. Such constructions and participants' lived experiences, including substantial waiting times and dehumanising conditions, echo Weber's (1978) notion of bureaucracy. Weber (1978) suggested that the development of capitalist bureaucratic procedures leads to dehumanisation, with bureaucracy constituting an iron cage in which workers are forced to obey challenging and often counterproductive organisational rules. This leads to organisational infrastructures that eventually become dysfunctional, dehumanising and do not show any regard for those in need.

Moreover, our study illustrates older, unpaid carers' concerns over online safety. Within the "Concerns Over Online Safety" theme, participants expressed concerns about the informed consent data process for services and applications. Gramsci (2011) constructed consent in capitalism as superficial, highlighting the disparity between our thoughts and the lack of consciousness in our actions. This results in a sense of

passivity, not necessarily due to individuals perceiving the dominant ideology as their own, but due to a lack of knowledge on how to enact their disagreement. The discursive and material possibilities of resistance reflect the ideological hegemony of the dominant class (Maglaras, 2013). We showcase how carers indicate that constructions of technological apparatuses are becoming increasingly hegemonic and result in surveillance realism. Surveillance realism is a theoretical concept pronouncing the discursive and institutional hegemony of big data and techno-economies underpinned by them and their collection, constituting “common sense”, underpinned by neoliberalism and privatisation. As such, everyday practices showcase how this surveillance is now constructed as voluntary, not due to consent but as a form of external push (Dencik, 2018).

We further pronounce carers’ perception of funding and lack thereof. In the theme titled "Carers' Contribution to Government Savings: Advocating for Increased Government and Policy Funding," participants emphasised the importance of receiving funding from the government and policymakers, framing these requirements within market-oriented terms. This resonates with the dominance of technological capitalism, which can be apathetic to whether its workers can afford healthcare expenses (Waters, 2020). It also reflects the collapse of the welfare state, where volunteering and unpaid labour are normalised under neoliberalism as they serve the neoliberal state and private service delivery (Hawksley & Georgeou, 2019). This demonstrates how the discourse is shaped by those in positions of power, inadvertently reproducing oppressive structures even when attempting to resist them (Gramsci, 2011). In many ways, the participants positioned themselves as the idealised neoliberal subject, framing their existence and labour as an enterprise producing capital (Althusser, 2014; Houghton, 2019). This, however, was constructed as a way of resisting the idea that they should financially cope with the lack of infrastructure and funds by themselves, an attempt to legitimise the necessity of being digitally included. This, perhaps, is an act of resistance to neoliberal constructions that promote individual responsibility against the welfare state, market rigidity and investment in human capital (Houghton, 2019).

Concerns about the "Cost of Technology" were raised in another theme, where participants expressed worries about the hidden, cumulative expenses associated with technology. Carers discussed the cost of technology and its impact on care expenses,



highlighting how governmental benefits-related funding is inadequate. This further resonates with the societal discourse that idealises volunteering and unpaid labour in neoliberalism while social care is considered necessary but too expensive for the market. The shift towards hegemonic volunteering discourses, coupled with a lack of healthcare and the collapse of welfare policies, becomes societal common sense (Moll, 2022). Neoliberalism aims to transfer socio-economic control from the public to the private sector through market transactions (Moll, 2022). As a result, participants reported that they have to financially compensate for the lack of both technology and welfare, such as covering costs related to dental home visits.

In the theme "Inclusivity and ableism," participants stressed the need for inclusive training, especially for individuals with limited mobility or confinement. This is a novel finding in the literature, and our contribution lies in theorising ableism in the context of the digitalisation of care. Carers highlighted the inadequacy of current policies, tools, and technologies in terms of inclusivity, which perpetuates feelings of alienation and contributes to digital exclusion. Ableism, deeply rooted in neoliberalism, regulates who contributes to labour and facilitates dependence on hyper-capitalist constructs, limiting societal participation. It creates an apparatus that forces individuals to attend to what is perceived as the norm using their own resources, without material support (Charitsis & Lehtiniemi, 2023).

Besides its theoretical contributions, our study has the potential for practical implications. The theme "The need for training" underscores the importance of policy interventions, with participants requesting assistance to familiarise themselves with technology. This aligns with older individuals often seeking training to improve their media literacy, with a recent body of research emphasising the significance of accessible and personalised one-on-one training for older people (Neves & Mead, 2021). Finally, in the theme "Co-production and co-design in research and policy," participants stress the importance of involving carers in planning and designing, to enhance the applicability of emerging technologies. This highlights the policies they desire, allowing them agency in relation to interventions and constructions related to the digitalisation of care. Therefore, our research underscores the need for ongoing collaboration and co-production. These recommendations promote more effective, context-sensitive, and tailored solutions in healthcare and caregiving.

Our study does not come without limitations. The predominantly female sample does not provide insight into male or non-binary constructions of care or the ability to explicate gendered differences, further highlighting the gendered roles associated with caregiving. As the findings of our study are preliminary, they are limited to a specific geographical area, and thus, the results could reflect the socio-economic context of Liverpool. Finally, when discussing our findings, we adopted a critical theoretical underpinning influenced by Gramscian ideological hegemony and thus did not consider alternative explanations.

## References

- Althusser, L. (2014). *On the reproduction of capitalism: Ideology and ideological state apparatuses*. Verso Books.
- Baker, A. R., & Billinge, M. (Eds.). (2004). *Geographies of England: the North-South divide, material and imagined*. Cambridge University Press.
- Braun, V., Clarke, V., Hayfield, N., Moller, N., and Tischner, I. (2018). *Qualitative Story Completion: A Method with Exciting Promise*. In Liamputtong, Pranee ed. *Innovative Research Methods in Health Social Sciences*. Research Methods in Health Social Sciences (3) (Ed., Liamputtong, P.). Singapore: Springer, pp. 1–18.
- Carey, M., & Foster, V. (2013). *Social work, ideology, discourse and the limits of post-hegemony*. *Journal of Social Work*, 13(3), 248-266.
- Charitsis, V., & Lehtiniemi, T. (2023). *Data ableism: Ability expectations and marginalization in automated societies*. *Television & New Media*, 24(1), 3-18.
- Clarke, V., Braun, V., Frith, H., & Moller, N. (2019). *Editorial introduction to the special issue: Using story completion methods in qualitative research*. *Qualitative Research in Psychology*, 16(1), 1-20.
- Cylus, J., Figueras, J., & Normand, C. (2019). Will population ageing spell the end of the welfare state? A review of evidence and policy options. *European Observatory on Health Systems and Policies*, Copenhagen.
- Damant, J., Freddolino, P., King, D., Wittenberg, R., Dangoor, M., Lorenz, K., & Hu, B. (2020). *Digital needs of unpaid carers of older people: A research agenda*. *Gerontechnology*, 19, 151.
- Dencik, L. (2018). *Surveillance realism and the politics of imagination: Is there no alternative?*. *Krisis: Journal for Contemporary Philosophy*, 2018(1), 31-43.
- Egan, K. J., Clark, P., Deen, Z., Dutu, C. P., Wilson, G., McCann, L., ... & Maguire, R. (2022). *Understanding current needs and future expectations of informal caregivers for technology to support health and well-being: national survey study*. *JMIR aging*, 5(1), e15413.
- Edwards, D. (1994). *Script formulations: An analysis of event descriptions in conversation*. *Journal of Language and Social Psychology*, 13(3), 211-247. Link
- Gramsci, A. (2011). *Prison Notebooks, Volumes 1-3*. Columbia University Press.
- Hamblin, K. (2020). *Technology and social care in a digital world : challenges and opportunities in the UK*. *Journal of Enabling Technologies*, 14 (2), pp. 115-125.
- Hamblin, K. (2022). *Sustainable social care: the potential of mainstream “smart” technologies*. *Sustainability*, 14(5), 2754.
- Hawksley, C., & Georgeou, N. (2019). *Gramsci 'makes a difference': Volunteering, neoliberal 'common sense', and the sustainable development goals*. *Third Sector Review*, 25(2), 27-56.

- Hirvonen, H., Tammelin, M., Hänninen, R., & Wouters, E. J. (Eds.). (2022). *Digital transformations in care for older people: Critical perspectives*. Routledge.
- Hughes, G., Carolan, S., Greenhalgh, T., Hinder, S., & Shaw, S. E. (2020). *Care organizing technologies and the post-phenomenology of care: An ethnographic case study*. *Social Science & Medicine*, 255, 112984.
- Houghton, E. (2019). *Becoming a neoliberal subject*. *Ephemera*, 19(3), 615-626.
- Kubicek, H., & Cimander, R. (2009). *Three dimensions of organizational interoperability*. *European Journal of ePractice*, 6, 1-12.
- Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*. Morgan Kaufmann.
- Lee, S., Ory, M. G., Dahlke, D. V., & Smith, M. L. (2022). *Social support, sense of belonging, and communication technology use among paid and unpaid caregivers of middle-aged and older adults*. *Frontiers in Public Health*, 10, 898042.
- Lorenz, K., Freddolino, P. P., Comas-Herrera, A., Knapp, M., & Damant, J. (2019). *Technology-based tools and services for people with dementia and carers: Mapping technology onto the dementia care pathway*. *Dementia*, 18(2), 725-741.
- Maglaras, V. (2013). *Consent and submission: Aspects of Gramsci's theory of the political and civil society*. *Sage Open*, 3(1), 2158244012472347.
- Majumdar, A. (2022). *Thematic analysis in qualitative research*. In *Research anthology on innovative research methodologies and utilization across multiple disciplines* (pp. 604-622). IGI Global.
- Moll, I. (2022). *The fourth Industrial revolution: A new Ideology. tripleC: Communication, Capitalism & Critique*. *Open Access Journal for a Global Sustainable Information Society*, 20(1), 45-61.
- Neves, B. B., & Mead, G. (2021). *Digital technology and older people: Towards a sociological approach to technology adoption in later life*. *Sociology*, 55(5), 888-905.
- Parzen, M., O'Keefe-McCarthy, S., Salfi, J., & Taplay, K. (2021). *Perceptions of informal caregivers' use of smart technology in caring for an older adult*. *Journal of Nursing Occupational Health*, 2(3), 230-239.
- Pearson, R. (2019). *A feminist analysis of neoliberalism and austerity policies in the UK*. *Soundings*, 71(71), 28-39.
- Pekkarinen, S., & Melkas, H. (2017). *Digitalisation in health care and elderly care services: From potholes to innovation opportunities*. *International Journal of Information Systems and Social Change (IJISSC)*, 8(1), 24-45.
- Pickard, L. (2015). *A growing care gap? The supply of unpaid care for older people by their adult children in England to 2032*. *Ageing & Society*, 35(1), 96-123.

- Reid, J. A., Kamler, B., Simpson, A., & Maclean, R. (1996). "Do you see what I see?" *Reading a different classroom scene*. *International Journal of Qualitative Studies in Education*, 9, 87-108.
- Sætra, H. S., & Fosch-Villaronga, E. (2021). Healthcare digitalisation and the changing nature of work and society. *Healthcare*, 9(8), 1007.
- Shaw, S. E., Hughes, G., Hinder, S., Carolan, S., & Greenhalgh, T. (2020). *Care organising technologies and the post-phenomenology of care: An ethnographic case study*. *Social Science & Medicine*, 255, 112984.
- Smith, F., Francis, S. A., Gray, N., Denham, M., & Graffy, J. (2003). *A multi-centre survey among informal carers who manage medication for older care recipients: problems experienced and development of services*. *Health & social care in the community*, 11(2), 138-145.
- Smithson, J. (2000). *Using and analysing focus groups: limitations and possibilities*. *International journal of social research methodology*, 3(2), 103-119.
- Spann, A., Vicente, J., Allard, C., Hawley, M., Spreeuwenberg, M., & de Witte, L. (2020). Challenges of combining work and unpaid care, and solutions: A scoping review. *Health & Social Care in the Community*, 28(3), 699-715.
- Sriram, V., Jenkinson, C., & Peters, M. (2019). Informal carers' experience of assistive technology use in dementia care at home: a systematic review. *BMC geriatrics*, 19, 1-25.
- Turner III, D. W. (2010). *Qualitative interview design: A practical guide for novice investigators*. *The qualitative report*, 15(3), 754.
- Waters, A. (2020). *Will neoliberal capitalism survive the coronavirus crash or is this the beginning of techno-feudalism?*. *Journal of Australian Political Economy*, The, (86), 406-431.
- Wright, J., & Hamblin, K. (2023). *Technology and adult social care in England*. In *Care Technologies for Ageing Societies* (Eds., Hamblin, K. and Lariviere, M.) Policy Press, pp. 18-48).

# Understanding Artificial Intelligence for Data with A High Level of Abstraction: Beyond Pixel Importance

## Abstract

*Artificial Intelligence (AI) models using explainable AI (XAI) can generate knowledge through pattern recognition. When data is used with a high level of abstraction, typical feature attribution methods may not suffice. For this work-in-progress paper, different methods of feature manipulation have been studied to further investigate which patterns in the data may be important for AI prediction, using dose measurements in radiotherapy as a sample. Thus far, we have explored superpixels, and textures using Fourier transform methods. Future research includes the further enhancement of AI predictions using geometric shapes, and possibly additional semantic features.*

**Keywords:** Explainable Artificial Intelligence (XAI), High-level Abstract Data, Radiotherapy, Healthcare

*In progress paper*

## 1.0 Introduction

Deep learning (DL) algorithms have shown high performance for difficult tasks, such as medical image interpretation, in some cases outperforming clinicians (Shen et al., 2019; Topol, 2019). Due to the inherent black-box characteristics, using DL algorithms comes at the expense of artificial intelligence (AI) transparency, which is undesirable in a high-risk settings such as healthcare, where errors can have severe implications for patients' health. Explainable artificial intelligence (XAI) at single patient-level, as for example feature-attribution in the form of heatmaps, is intuitive for users, however, can falsely engender trust in the AI prediction whereas the methods may not be reliable or robust (Ghassemi et al., 2021). In recent literature, researchers established a propensity to over-trust AI, which can become problematic in a healthcare setting (Borenstein et al., 2018; Howard, 2020). XAI could, however, play a role in quality assurance and AI monitoring in healthcare. In contrast to understanding feature attribution of single predictions, understanding how the AI works is key to identifying potential model errors (Fügener et al., 2021; Poursabzi-Sangdeh et al., 2021).

The current majority of proposed XAI methods aim to understand which input features are of most importance for the AI prediction model (Barredo Arrieta et al., 2020). For example, using Shapley Additive exPlanations (SHAP) (Lundberg & Lee, 2017),

(2023) chronic illness, exhaustion, peer pressure, and alcohol intake are identified as the most important features for their lung cancer prediction model (Rikta et al., 2023). The feature's predictive capability, often referred to as "*feature importance*," which is also used in this study. For data with higher complexity, simple feature importance may not suffice. For example, in imaging data, pixel values alone have no meaning. The interaction among features is, therefore, crucial for comprehending the model's behaviour. In clinical practice, additional information is used to assess the medical image, such as shape, lines, contrast, density etc. While it is still unknown what information the AI uses, it is intuitive to assume that the AI uses the same information as humans. With the proposed study, we aim to get a better understanding of the AI decision making process to increase safe use for clinical practice.

In this work-in-progress paper, we present our direction of research towards understanding global feature importance for image data with a high level of abstraction. Our aim is to evaluate which features, besides pixels, can be extracted from the data and gain knowledge about how the importance of different types of features can be interpreted in AI models. While our work focusses on the application of dose measurements in radiotherapy, we foresee a contribution to other fields as well, both medical as non-medical. The aim of our study is to investigate what type of information the AI uses for its prediction-making. By doing so, we can better predict when the AI will malfunction or when users need to be cautious. These objectives lead us to the following research question: *Can the use of global feature attribution, in addition to pixel attribution, contribute to medical physicists' understanding of AI model behaviour?*

## **Background**

Issues around transparency, liability, accountability, justifiability, and verifiability of use of AI in healthcare have led both researcher and practitioners to explore XAI for medical applications.

Humans trust the AI better if they understand how the AI works and what information the AI uses (Jussupow et al., 2021). Although local explanations seem intuitive to users, many local XAI methods, such as Local Interpretable Model Explanations (LIME)

(Benda et al., 2021) heavily simplify the AI model by fitting a (linear) surrogate model to the local prediction of the AI. The simplification introduces an uncertainty to the model prediction which challenges the fidelity of the explanation with regard to the actual AI model (Benda et al., 2021). A problematic result might be over-trust and therefore over-reliance of AI in healthcare, especially for difficult tasks (Vasconcelos et al., 2022). While many studies focus on local explanations for medical applications (Panigutti et al., 2023), trust in AI by medical professionals is driven by both local and global information (Subramanian et al., 2024). By investigating the global level, we can obtain a better understanding of its boundaries and its implications (Fügener et al., 2021).

## **2.0 Materials and methods**

The initial phase of our study focuses on radiotherapy. Data from a case setting was collected to test our methods on. In future work, we will include other medical fields as well, such as radiology.

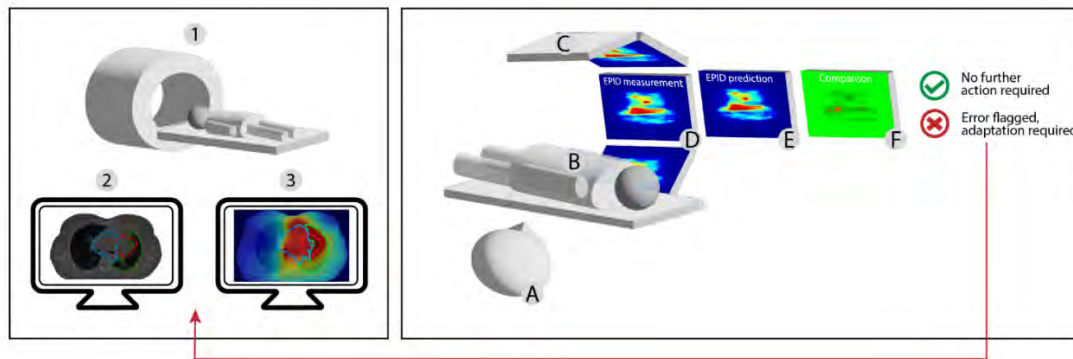
### **2.1 Case setting**

Radiation therapy is a treatment that utilizes radiation to treat cancer patients. The therapy has proven to be highly effective for a significant number of patients, with approximately one out of every two patients receiving some form of radiation therapy as part of their treatment (Baskar et al., 2012). The medical field is highly quantitative and technical in nature and requires a high accuracy since a small error (e.g., 2%) can be of clinical significance (Thwaites, 2013). In our clinical case, the patient is prescribed a dose over the span of one treatment, which consists of up to thirty fractions spread out over multiple days.

During treatment, the patient is positioned on a couch. The radiation source rotates around the patient to deliver the treatment. Opposite to the radiation source, a panel, i.e., the electronic portal imaging device (EPID), is located which captures the dose going through the patient (Figure 1). For each angle of the radiation beam, an EPID image is captured. Before treatment, by projecting the planned treatment on the computed tomography (CT) of the patient, clinicians are able to predict how the EPID image is expected to look like. Any disparities between the expected and actual EPID



images, measured using a gamma analysis, are considered treatment errors. Gamma analysis is a comparison method with predefined tolerance levels, allowing for small deviations. As patients come in multiple consecutive days, clinically significant errors can be addressed in the following fractions allowing for more accurate treatment and better treatment outcomes.



*Figure 1 Schematic representation of the radiotherapy workflow. The left frame illustrates the preparation steps of the radiotherapy workflow: 1. Imaging: each patient gets a CT scan, 2. Contouring: tumour(s) and critical organs are delineated on the CT image, 3. Treatment planning: irradiation plan and dose calculation are made based on CT image and contoured structures. The right frame illustrates the electronic portal imaging device (EPID) measurements and dose comparison-based flagging system. A: Radiation source rotating 360° around the patient, B: patient, C: EPID panel opposite to the radiation source, D: EPID measurement, E: EPID predictions derived from CT image and treatment plan, F: dose comparison between D and E, a predefined threshold determines whether the patient gets flagged for treatment adaptation.*

Clinical professionals in radiotherapy include radiation oncologists, radiotherapy technicians and clinical physicists. While many XAI for healthcare researchers target radiation oncologists, clinical physicists play a large role in quality assurance of systems used in clinical practice. As XAI serves as the communication system between the AI and the user, the XAI needs to be human-centred (Chromik & Butz, 2021; Haque et al., 2023). Due to the technical background of many clinical physicists, they typically have a inquisitive attitude towards technology such as AI.

## 2.2 Data

The dataset used in this study consists of synthetic gamma maps. For each error, gamma maps were synthetically developed by introducing an error in the planning CT or in the treatment plan. The dataset consists of 500 gamma maps, which were evenly distributed across the ten classes of treatment errors. For simplicity, abbreviations were used for the treatment errors (Table 1).

Class name	Class code	Class name	Class code
Anatomical changes: mediastinum shift	AMS	Linac mechanical errors: MLC shift systematic	MLCS
Anatomical changes: pleural effusion	APE	Linac mechanical errors: MU scaling random	MUR
Anatomical changes: tumor regression	ATR	Linac mechanical errors: MU scaling systematic	MUS
Anatomical changes: tumor shift	ATS	Patient positioning errors: rotation	PR
Linac mechanical errors: MLC shift random	MLCR	Patient positioning errors: translation	PT

**Table 1 Overview of class names & code for treatment errors**

### 2.3 AI architecture

For error detection, a convolutional neural network (CNN) based on a VGG-16 architecture was employed. The model's performance on the dataset is represented in the confusion matrix (Figure 2). It is noteworthy that the accuracy of the model is high for the majority of classes, with most of them achieving a precision rate of approximately 75%. However, it is important to highlight that the MUR and PT classes exhibit relatively lower accuracy levels, standing at 58% and 66%, respectively.

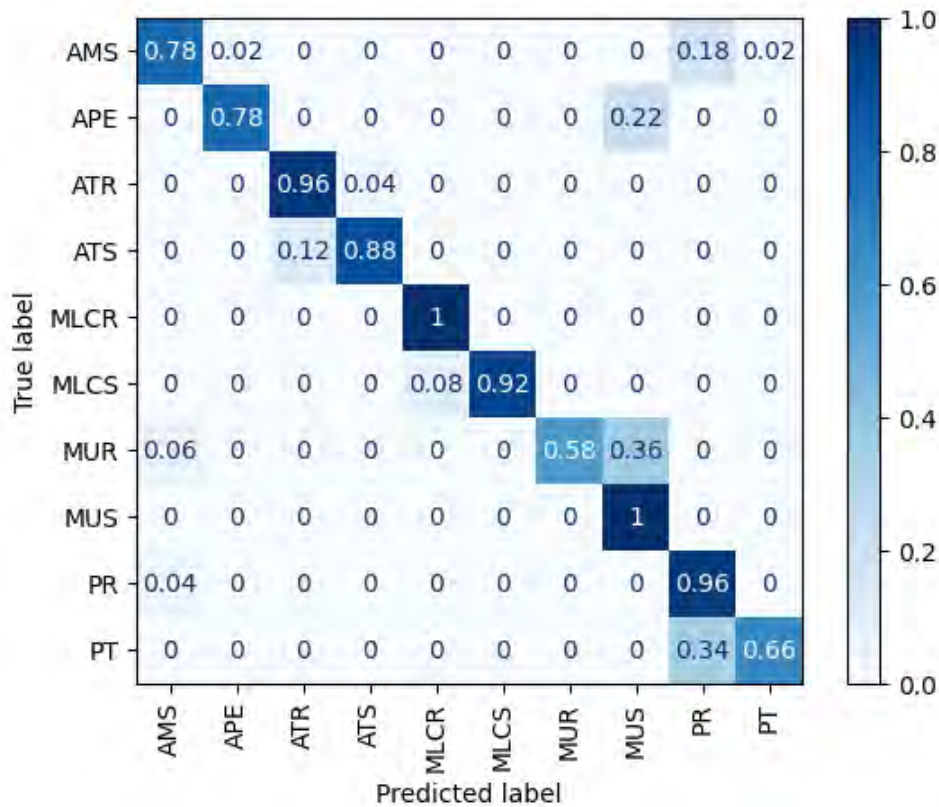


Figure 2 confusion matrix of the error classification model used for the XAI methods

## 2.4 Shapley additive explanations (SHAP)

SHapley Additive exPlanations (SHAP) (Lundberg & Lee, 2017) is an XAI technique rooted in game theory (Chen et al., 2023). SHAP aims to quantify the individual contributions of features to AI predictions, conceptualizing each feature as a participant in a predictive game. The SHAP is used to determine how each feature influences a model's predictions. Given that many prediction models cannot accommodate missing features, an imputation process becomes necessary to simulate the behaviour of non-participating features (Table 2). Imputation methods include substituting the missing value with the average, drawing a value from the feature's marginal distribution, or replacing it with zero.

---

Pseudo	
1	<b>for</b> i <b>in</b> length dataset:
2	data instance = the $i^{\text{th}}$ element of x
3	label = the $i^{\text{th}}$ element of y
4	<b>for</b> coalition <b>in</b> all_coalitions:
5	join_state = imputation(data instance, coalition)
6	no_join_state = imputation(data instance, coalition + feature)
7	y_pred_join = predict(join_state)
8	y_pred_no_join = predict(no_join_state)
9	performance_join = performance_join + performance metric(y_pred_join, label)
10	performance_no_join = performance_no_join + performance metric(y_pred_no_join, label)
11	feature_importance = (performance_join – performance_no_join) / number of data instances

---

**Table 2. Pseudo code for the calculation of feature importance values with the Shapley method.**

## 2.5 Feature extraction

Two different types of features were extracted for the purpose of the first stage of the study, firstly super pixels and secondly texture features. In further stages of the study, we plan to include different kind of features, such as for example, geometric shapes and absolute dose values.

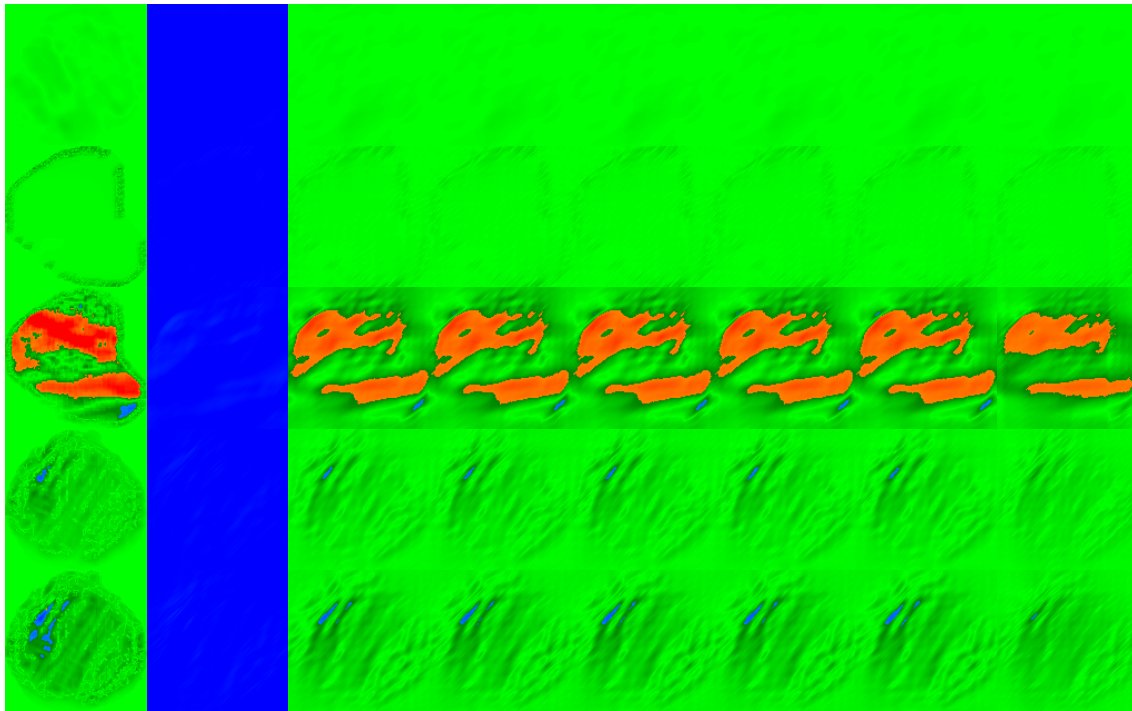
### 2.5.1 Super pixels

Superpixels are a collection of related pixels, for instance, a square of twelve-by-twelve pixels. Superpixels can decrease the number of features extensively and make the interpretation to humans more interpretable. In the process of calculating the feature importance of superpixels, one or more superpixels are imputed after each prediction. This imputation method involves substituting the designated superpixel(s) with those from a donor image. The underlying concept is that the feature's predictive capability, often referred to as "*feature importance*," can be assessed by measuring the difference in loss before and after imputation.

### 2.5.2 Texture features

By manipulating the input images with a Fourier transforms function, patterns in the image data can be distilled. Low frequencies in images relate to larger shapes and the background, whereas high frequencies typically represent fine details and texture (Zhou et al., 2001). Our XAI analysis focussed on understanding which frequency range was

the most important for the AI prediction. For example, higher frequency ranges could relate to boundaries of the measured dose, whereas lower frequencies could relate to systematic machine errors by ununiform dose delivery. In our analyses, the frequency ranged from 0 to 128 which we split up into seven buckets: 0-20, 21-40, 41-60, 61-80, 80-100, 100-114, 115-128. Figure 3 illustrates how the image looks when certain frequency ranges were filtered out of the image.



**Figure 3. Examples of images where certain frequencies are filtered out. Every row is from 1 image. Rows 1 & 2 are from the MLCS class, 3 is from the PT class, and 4 & 5 are from the APE class. The right picture is the original. From left to right the pictures have the following frequencies filtered out: 0-20, , 21-40, 41-60, 61-80, 81-100, 100-114, 115-128.**

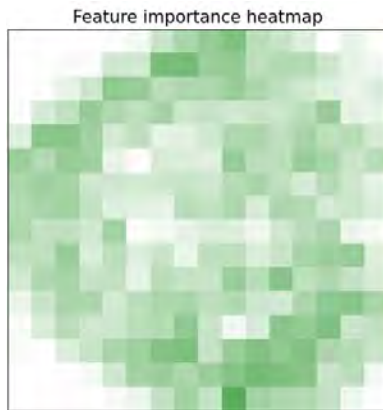
### 3.0 Preliminary results

Preliminary results of analyses with superpixels and Fourier frequencies are presented.

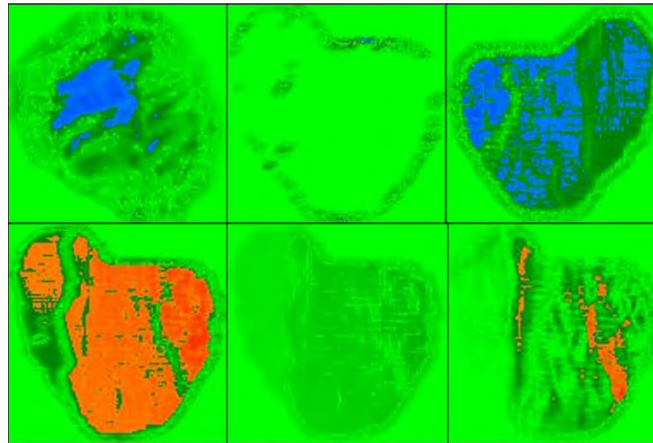
#### 3.1 Superpixels

An overall feature importance map is shown in Figure 4. No red superpixels are present in the overall heatmap. However, a closer examination demonstrates that the heatmap does contain some small negative values, with a minimum value at -0.003 (Table 2). Nonetheless, it's worth noting that the majority of superpixels contribute positively or not at all to the model's performance. On average, each superpixel holds a feature importance value of 0.0047. This signifies that substituting the pixels of a single

superpixel with random values results in decrease in accuracy of 0.47%. The single-class feature importance differs considerably from the overall values. The first difference is the presence of feature importance values that are far in the negative. This means that the performance of the model was higher when the values in that area are replaced with random values.



**Figure 4. Overall feature importance heatmap (based on full dataset).**



**Figure 5. Some typical gamma maps with edges in high feature importance areas. Class codes first row: APE, MLCR, MUS. Class codes second row: MUR, MUS, PR**

Class name	Mean	Standard deviation	Min	Max
Full dataset	0.0047	0.0027	-0.0003	0.0129
AMS	0.0054	0.0056	-0.0120	0.0221
APE	0.0053	0.0042	-0.0027	0.0197
ATR	0.0069	0.0062	-0.0125	0.0266
ATS	0.0063	0.0062	-0.0067	0.0259
MLCR	0.0004	0.0036	-0.0084	0.0171
MLCS	0.0057	0.0058	-0.0067	0.0263
MUR	0.0041	0.0046	-0.0071	0.0232
MUS	0.0061	0.0062	-0.0114	0.0242
PR	0.0048	0.0049	-0.0080	0.0217
PT	0.0016	0.0059	-0.0205	0.0158

**Table 2 Statistical summary of the feature importance values**

### 3.2 Textures

The feature contribution is generally further from zero for the frequency ranges 0-20 and 115-128. While the feature contribution values of the 115-128 range are generally closer to zero. This means that the overall colours and large shapes have more impact on the performance of the model than small objects. In some cases, there is also a slight

deviation from zero for the frequency ranges 21-40 and 101-114. Nevertheless, these deviations are only small compared to the 0-20 and 114-128 ranges. The feature contribution of the remaining frequency ranges is centred around zero. So, the prediction of the model does not, or only slightly, change when these frequency ranges are filtered out. The fact that the 0-20 range has the largest impact is not surprising given that the images changed the most after filtering out that range. The gamma maps turned completely blue when the 0-20 range was omitted.

## **5.0 Discussion**

Our preliminary findings indicate that using manipulated features to examine AI behaviour can shed different light on the inner workings of the AI than typical XAI methods. Using the feature importance map of superpixels, clinical physicists could generate a similar map for a period where the AI was used to evaluate whether the shape is consistent. More research is needed to finetune how to use these global findings in clinical practice. Moreover, larger shapes and areas were more important for the AI's prediction than fine lines and details. This could also relate to the shape of the error detected instead of, for example, specific lines.

Our next research agenda includes the evaluation of semantic features and geometric features. To evaluate the contribution of semantic features on the AI prediction, various techniques were proposed (de Mijolla et al., 2021). For example, using image-to-image translation, semantic features can be added. By modifying the semantic features, they were able to create a joint and not joint state. In our context, semantic features with regard to dose characteristics could be added.

## **5.2 Implications**

Using features attribution methods that go beyond pixel attribution presents promising results in exploring what information AI uses. Further exploring the information that AI uses for its predictions contributes to a better understanding of the boundaries of the AI models, perhaps even generate knowledge on how AI models can more easily be generalized towards other settings.

## References

- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Baskar, R., Lee, K. A., Yeo, R., & Yeoh, K.-W. (2012). Cancer and Radiation Therapy: Current Advances and Future Directions. *International Journal of Medical Sciences*, 9(3), 193–199. <https://doi.org/10.7150/ijms.3635>
- Benda, N. C., Novak, L. L., Reale, C., & Ancker, J. S. (2021). Trust in AI: Why we should be designing for APPROPRIATE reliance. *Journal of the American Medical Informatics Association: JAMIA*, 29(1), 207–212. <https://doi.org/10.1093/jamia/ocab238>
- Borenstein, J., Wagner, A. R., & Howard, A. (2018). Overtrust of Pediatric Health-Care Robots: A Preliminary Survey of Parent Perspectives. *IEEE Robotics & Automation Magazine*, 25(1), 46–54. <https://doi.org/10.1109/MRA.2017.2778743>
- Chen, H., Covert, I. C., Lundberg, S. M., & Lee, S.-I. (2023). Algorithms to estimate Shapley value feature attributions. *Nature Machine Intelligence*, 5(6), Article 6. <https://doi.org/10.1038/s42256-023-00657-x>
- Chromik, M., & Butz, A. (2021). Human-XAI Interaction: A Review and Design Principles for Explanation User Interfaces. In C. Ardito, R. Lanzilotti, A. Malizia, H. Petrie, A. Piccinno, G. Desolda, & K. Inkpen (Eds.), *Human-Computer Interaction – INTERACT 2021* (pp. 619–640). Springer International Publishing. [https://doi.org/10.1007/978-3-030-85616-8\\_36](https://doi.org/10.1007/978-3-030-85616-8_36)



de Mijolla, D., Frye, C., Kunesch, M., Mansir, J., & Feige, I. (2021). *Human-interpretable model explainability on high-dimensional data* (arXiv:2010.07384). arXiv. <http://arxiv.org/abs/2010.07384>

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2021). *Will Humans-in-The-Loop Become Borgs? Merits and Pitfalls of Working with AI* (SSRN Scholarly Paper 3879937). <https://papers.ssrn.com/abstract=3879937>

Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. *The Lancet. Digital Health*, 3(11), e745–e750. [https://doi.org/10.1016/S2589-7500\(21\)00208-9](https://doi.org/10.1016/S2589-7500(21)00208-9)

Haque, A. B., Islam, A. K. M. N., & Mikalef, P. (2023). Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. *Technological Forecasting and Social Change*, 186, 122120. <https://doi.org/10.1016/j.techfore.2022.122120>

Howard, A. (2020). Are We Trusting AI Too Much? Examining Human-Robot Interactions in the Real World. *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 1. <https://doi.org/10.1145/3319502.3374842>

Jussupow, E., Spohrer, K., Heinzl, A., & Gawlitza, J. (2021). Augmenting Medical Diagnosis Decisions? An Investigation into Physicians' Decision-Making Process with Artificial Intelligence. *Information Systems Research*, 32(3), 713–735. <https://doi.org/10.1287/isre.2020.0980>

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 30. <https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html>

Panigutti, C., Beretta, A., Fadda, D., Giannotti, F., Pedreschi, D., Perotti, A., & Rinzivillo, S. (2023). Co-design of Human-centered, Explainable AI for Clinical Decision Support. *ACM Transactions on Interactive Intelligent Systems*, 13(4), 21:1-21:35. <https://doi.org/10.1145/3587271>

Poursabzi-Sangdeh, F., Goldstein, D. G., Hofman, J. M., Wortman Vaughan, J. W., & Wallach, H. (2021). Manipulating and Measuring Model Interpretability. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–52. <https://doi.org/10.1145/3411764.3445315>

Rikta, S. T., Uddin, K. M. M., Biswas, N., Mostafiz, R., Sharmin, F., & Dey, S. K. (2023). XML-GBM lung: An explainable machine learning-based application for the diagnosis of lung cancer. *Journal of Pathology Informatics*, 14, 100307. <https://doi.org/10.1016/j.jpi.2023.100307>

Shen, J., Zhang, C. J. P., Jiang, B., Chen, J., Song, J., Liu, Z., He, Z., Wong, S. Y., Fang, P.-H., & Ming, W.-K. (2019). Artificial Intelligence Versus Clinicians in Disease Diagnosis: Systematic Review. *JMIR Medical Informatics*, 7(3), e10010. <https://doi.org/10.2196/10010>

Subramanian, H. V., Canfield, C., & Shank, D. B. (2024). Designing explainable AI to improve human-AI team performance: A medical stakeholder-driven scoping review. *Artificial Intelligence in Medicine*, 102780. <https://doi.org/10.1016/j.artmed.2024.102780>

Thwaites, D. (2013). Accuracy required and achievable in radiotherapy dosimetry: Have modern technology and techniques changed our views? *Journal of Physics: Conference Series*, 444(1), 012006. <https://doi.org/10.1088/1742-6596/444/1/012006>

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), Article 1. <https://doi.org/10.1038/s41591-018-0300-7>

Vasconcelos, H., Jörke, M., Grunde-McLaughlin, M., Krishna, R., Gerstenberg, T., & Bernstein, M. S. (2022). When do XAI methods work? A cost-benefit approach to human-AI collaboration. *CHI Workshop on Trust and Reliance in AI-Human Teams (New Orleans, USA)*. ACM. [https://chi-trait.github.io/Papers/CHI\\_TRAIT\\_2022\\_Paper\\_44.Pdf](https://chi-trait.github.io/Papers/CHI_TRAIT_2022_Paper_44.Pdf).

Zhou, F., Feng, J. F., & Shi, Q. Y. (2001). Texture feature based on local Fourier transform. *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, 2, 610–613 vol.2. <https://doi.org/10.1109/ICIP.2001.958567>

# MIND THE GAP: USING THRESHOLD CONCEPT THEORY TO ADVANCE BLOCKCHAIN EDUCATION

## **Abstract**

*Students entering tertiary education embark on a transformative learning path. In this setting, learning is characterised by grasping essential (threshold) concepts to fully engage with the course's academic material. Their educational journey is shaped by experiences both within the tertiary institution and during work placements. The challenging knowledge linked with threshold concepts poses notable obstacles to transformative learning. Blockchain is a decentralized and distributed digital ledger technology that records transactions across multiple computers in a way that ensures the security and transparency of data. As this technology becomes more prevalent in tertiary education, there's a growing need for research to identify threshold concepts in blockchain education. Understanding these pivotal concepts is crucial for effective teaching and comprehensive learning about the technology. The aim of this research is to review both the transformative learning processes in tertiary education and the foundational principles of blockchain technology. By delving deep into these areas, we hope to present a comprehensive framework for identifying threshold concepts when teaching blockchain technology.*

**Keywords:** Threshold Concepts, Blockchain Technology, Learning Barriers, Transformative Learning, Third Level Teaching and Learning

## **1.0 Introduction**

This research addresses an established educational framework that is emerging in the Blockchain technology research field that asserts threshold concepts as mediators of learning outcomes. At its core, a threshold concept pertains to specific knowledge in higher education that, when grasped by students, reshapes their perception or comprehension of the topic (Land and Meyer, 2010, Kiley and Wisker, 2009). The threshold concept offers a lens through which educators can gauge the progression of students as they navigate and master subjects that are inherently complex and challenging. Such concepts act as pivotal points or "gateways" in the learning journey, marking the transition from a superficial understanding to a deeper, more integrated knowledge of a subject (Meyer, and Land, 2006, Lucas and Mladenovic, 2007). By focusing on these core concepts, educators can tailor their teaching methods to ensure that students not only understand the basics but also appreciate the nuances and

intricacies of the subject matter (Cousin, 2008, Zepke, 2013). Furthermore, over the past decade, there had been increased attention in the literature pertaining to the role of threshold concepts in teaching and learning and for curriculum (re)design (Clohessy and English, 2022).

The nuances of Blockchain technology, with its decentralized nature and cryptographic foundations, present unique challenges in the educational realm. Furthermore, the dynamic and rapidly evolving landscape of Blockchain technology underscores the need for adaptive and flexible curricula (Berger, 2023). Compared to other disciplines, there has been relatively little published in the education literature on Blockchain technology threshold concepts, despite calls from authors for more research into this area (see Schneider, B., and Azan, 2022). To foster deeper investigation and discussion in the realm of threshold concepts and Blockchain technology, and to offer findings that can guide teaching methodologies, this research aims to shed light on how Blockchain technology students confront and manage challenging knowledge while learning about threshold concepts integrated into third level education syllabus. Specifically, this research aims to elucidate the following two research questions:

***Research Question 1: What threshold concepts do third level students encounter when learning about Blockchain technology?***

***Research Question 2: What coping mechanisms do third level students use to manage their transition through the liminal space when they encounter threshold concepts?***

The results of this research will offer valuable perspectives on enhancing curriculum and evaluation methods, spanning from the detailed module level, through the broader program scope, up to the extensive discipline context in the professional setting.

This paper is structured as follows. An overview of Blockchain technology is presented in Section 2.0. Section 2.1 discusses the challenges presented by Blockchain technology from a learner perspective. Section 3.0 introduces threshold

concepts and outlines its role in higher education. Finally, the paper concludes in section 4.0 and outlines the next steps for this study.

## **2.0Blockchain Technology**

Blockchain technology has emerged as an innovative foundational technology with the potential to reshape various sectors of the economy and society. Blockchain is a technology used in various projects, from cryptocurrencies to IoT applications. It is characterized by its unique features of immutability (unchangeable records), transparency (open visibility of transactions), and anonymity (privacy without revealing user identity). These features set blockchain apart from other distributed ledger technologies (Ghiro et al., 2021). Its decentralized, transparent, and secure nature has garnered significant attention from academia, industry, and governments worldwide (Clohessy and Acton, 2019). Zheng et al. (2018) highlighted the myriad of advantages provided by blockchain technology, such as decentralization, persistency, and auditability, with applications spanning from cryptocurrency to risk management and the Internet of Things (IoT). Furthermore, the globalisation of supply chains has presented challenges in their management, and blockchain, with its inherent transparency and traceability, offers solutions. Clohessy and Heaslip (2022) examined the role of blockchain in enhancing the sustainability of supply chains by fostering a concept known as decentralized autonomous supply chains in the content of the agriculture sector. Additionally, the energy sector has shown considerable interest in blockchains, with Andoni et al. (2019) providing a systematic review of its applications in the industry, discussing opportunities and challenges in areas like peer-to-peer energy trading and electric vehicle charging.

More recently, Blockchain technology has been increasingly incorporated into third level educational programmes, catering to a range of learners from undergraduate to postgraduate students. Various universities and academic institutions globally have recognised the potential and significance of blockchain technology and have integrated blockchain-related courses and programs into their curriculum. For example, global third level institution such as NYU Law, Princeton, Stanford, and UC Berkeley, and in Europe, the University of Cumbria, IT University of Copenhagen, and University of Nicosia among others, now offering blockchain courses.

Furthermore, a Gartner poll in 2023 highlighted that, 18% of higher education institutions are considering implementing it in their curriculums (Berger, 2023). However, extant research indicates that teaching blockchain technology in third level higher education institutions can be challenging from a learner and educator perspective (Labouseur et al., 2019, Xu et al., 2021, Froehlich et al., 2023). For example, Labouseur et al., (2019) showcased how, “Having enjoyed (and endured) several student-centered, research-based experiences spent trying to understand blockchain and develop an elementary implementation of it” led to a number of teachable moments which are currently under researched in the literature. One of those teachable moments identified that educators must distil the complexity of blockchain technology into manageable and digestible concepts. Furthermore, Xu et al., (2021) present a tiered “Four- Level Guidance” blockchain practice teaching model which guides undergraduates progressively through advanced levels of blockchain concept understanding.

In the next section we delineate the challenges faced by learners in grasping the multifaceted and inherent complex nature of blockchain technology.

## **2.1 Challenges Faced by Learners in Grasping Blockchain Technology**

Blockchain technology, while possessing huge potential for innovation and digital transformation, presents a steep learning curve for many learners. Its intricate nature, combined with its relatively recent emergence, means that understanding it fully, can be a daunting task. A review of the literature has identified the main challenges which are nuanced to blockchain technology which make it a difficult concept for learners to grasp and include the following:

- **Complex Technical Concepts:** Blockchain combines cryptography, distributed systems, and consensus algorithms. For those unfamiliar with these areas, especially cryptography, the concepts can be abstract and challenging (Labouseur et al., 2019, Clohessy, 2022, Nguyen, D., et al. 2021).
- **Lack of Standardised Educational Material:** Due to its novelty, there isn't a universally accepted curriculum or standardised set of educational materials for blockchain. This lack of standardisation can lead to information

inconsistencies, making it hard for learners to discern reliable sources (Clohessy, Godfrey, and Houston, 2018, Alsamhi, S., et al. 2022).

- **Rapid Evolution of the Technology:** Blockchain technology is rapidly evolving. New consensus algorithms, platforms, and applications are emerging regularly. Keeping up with this pace can be overwhelming for learners, especially when foundational knowledge is still being solidified (Xu et al., 2021, Chen, H., et al. 2021).
- **Technology Legitimation:** Due to its association with cryptocurrencies, many learners' approach blockchain with preconceived notions or misconceptions. Decoupling the capabilities of blockchain beyond just cryptocurrencies can be a hurdle (Schneider, B., and Azan, 2022, Tian, Y., et al. 2022).
- **Practical Implementation Challenges:** While theoretical knowledge is essential, understanding blockchain deeply often requires hands-on experience. Setting up nodes, understanding smart contracts, or participating in a blockchain network might require technical expertise beyond basic tutorials or courses (Xu et al., 2021, Nguyen, D., et al. 2021).
- **Interdisciplinary Nature:** Blockchain intersects with various fields like law, finance, and supply chain. For a comprehensive understanding, learners often need to venture outside their primary field of study (Clohessy, 2022, Alsamhi, S., et al. 2022).
- **Dearth of Blockchain Educators:** As a nascent field, there's a shortage of experts who are also effective educators. Finding instructors or mentors who can provide clear, accurate, and up-to-date information on blockchain can be a significant challenge (Froehlich et al., 2023, Clohessy et al., 2018, Chen, H., et al. 2021).

Addressing these challenges requires a multi-faceted approach, combining standardised education, hands-on training, and continuous learning where threshold concept theory offers a promising approach for educators to overcome these learners' challenges (Clohessy and English, 2022).

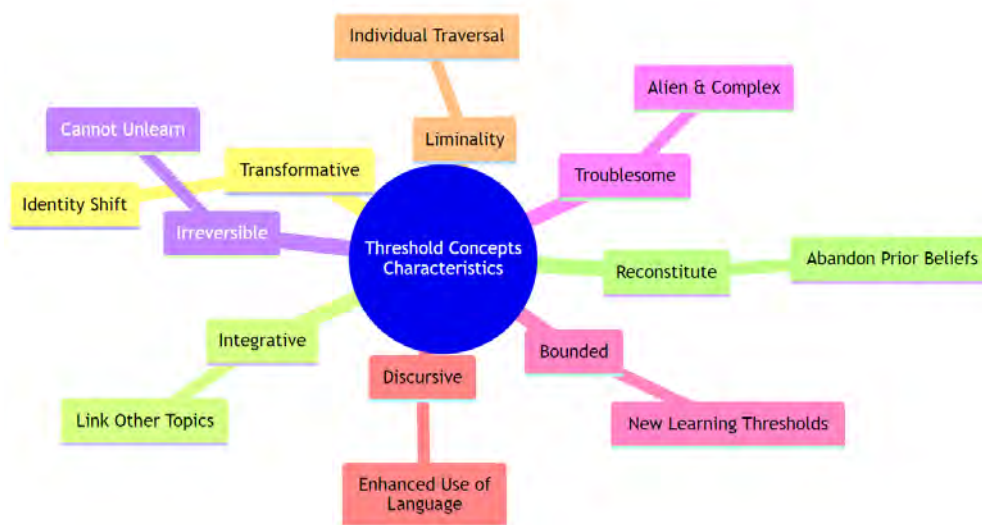


### **3.0 Threshold Concept Theory and its Role in Education**

Threshold concepts are conceptual doorways that, once crossed, offer learners a new viewpoint, revealing previously unseen aspects of a topic. These concepts enable students to approach a subject in a way they couldn't before. It's a transformative shift in a student's learning perspective. However, if students don't experience this transformation, their grasp of these pivotal concepts remains elusive, hindering their overall advancement in that subject. In essence, threshold concepts can either unlock or obstruct a student's learning journey (Land et al., 2010). Ultimately, threshold concepts serve as transformative gateways in various academic disciplines, reshaping a learner's understanding and perspective. For example, economics introduces foundational ideas such as "opportunity cost" and "marginal thinking," which shift the way decisions and benefits are evaluated (Davies and Mangan, 2007). Philosophy grapples with profound ideas like "existentialism" and "epistemology," questioning the nature of existence and knowledge (Vandenberg, 1993). These concepts, while initially troublesome, become foundational once grasped, paving the way for deeper exploration within each discipline. For example, let's delve deeper into the threshold concept of 'opportunity cost' from the field of economics. The concept of opportunity cost is more about influencing choices than being influenced by them. When students embrace this perspective, it profoundly alters their approach to decision-making. Not only does it reshape their understanding of their own decisions, but it also equips them to analyse the choices of others (Shanahan, Foster, & Meyer, 2006).

There are notable parallels between the threshold concept approach and existing studies on transformational learning. For example, foundational works by Mezirow (1978, 1990) introduced the idea of 'perspective transformation'. This concept is described as the act of critically examining and understanding the underlying assumptions that shape our worldview, then revising these assumptions to foster a more comprehensive and adaptive perspective, leading to actions based on this renewed understanding (Mezirow, 1990, p. 14). Kitchenham (2010) suggests that Mezirow's insights into transformative learning draw heavily from Kuhn's (1962) ideas on paradigms, Freire's (1970) studies on conscientisation, and Habermas's (1971, 1984) explorations into learning domains. Land and Meyer (2010, p.12) argue that Mezirow's concept of perspective transformation shares similarities with the

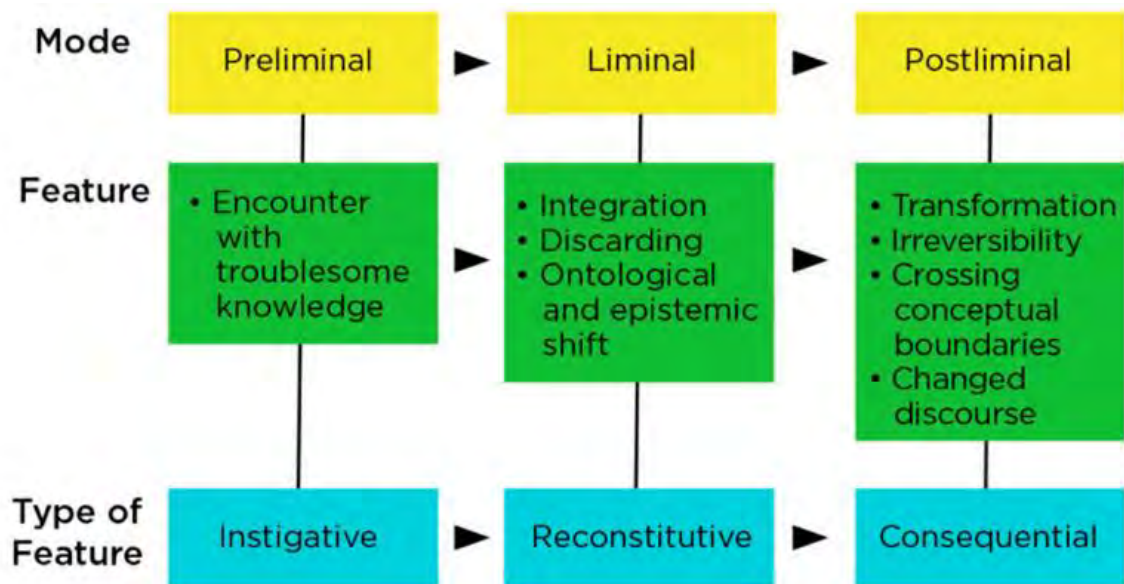
triggering effect of threshold concepts, the transitional phase in threshold theory, and the integrative process it involves. Now, we present a framework to outline and understand threshold concepts, drawing inspiration from the studies of Meyer and Land (2003, 2005, 2006). Their research offers insights into understanding specific educational elements that might be inherently difficult. The threshold concepts framework (Land and Meyer, 2010) is built on eight fundamental attributes highlighted in Figure 1.



**Figure 1.** Threshold Concepts Framework (Land and Meyer, 2010).

Central to the threshold concept, and a primary focus of this study, is the idea of students navigating a transitional phase, or liminal space, as they engage with a threshold concept (Land, Rattray, and Vivian, 2014). Figure 2 illustrates the stages of threshold concepts, encompassing preliminal, liminal, and postliminal phases (Land and Meyer, 2010). This view on transformational learning is influenced by Kegan's (1982) idea of 'consciousness shifts', Boyd and Meyers' (1988) model focusing on emotional processes, and Mezirow's (1978, 1990) theory of 'perspective transformation'. The latter describes a learner's transformative journey initiated by confronting a perplexing situation or concept (Mezirow, 1990). Mezirow's perspective transformation encompasses several transformative stages, akin to the 'characteristics' shown in Figure 1. According to Land and Meyer (2010), a student's path to grasping a threshold concept starts with encountering challenging troublesome knowledge in the preliminal phase. This knowledge acts as a trigger or instigative catalyst, leading

to a liminal state. In this liminal phase, students integrate this new knowledge, necessitating a reevaluation of their previous understanding and the abandonment of outdated views (Land and Meyer, 2010, p.11). This liminal phase is also denoted by students adopting specific coping mechanisms to assist them with the assimilation of new knowledge. This assimilation results in shifts in both being and knowing, seen as a transformative aspect of the threshold concept. This liminal mode is a phase where true learning transformation occurs leading to a reconstituted way of thinking about the subject matter. These transformative features help students traverse a conceptual boundary, leading to a postliminal stage and the subsequent new comprehension. In the postliminal stage, both the learning process and the learner undergo a transformation. This irreversible transformation is characterised by a change in discourse. The attributes of the postliminal phase are termed as the consequential characteristics of a threshold concept.



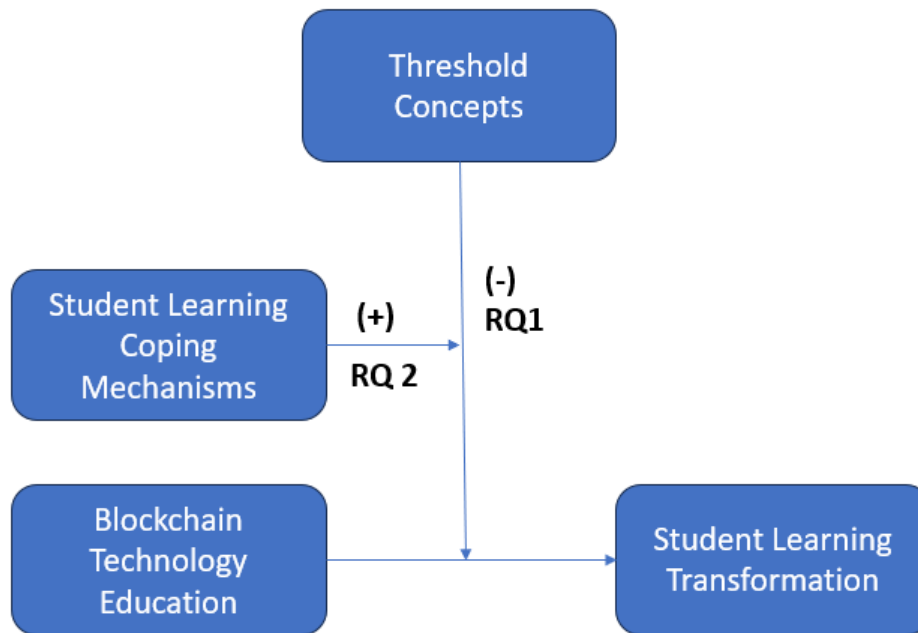
**Figure 2.** A Relational View of the Features of Threshold Concepts (Land and Meyer, 2010)

While this theory possesses cogent potential to identifying how students cope with troublesome knowledge it is important to acknowledge its limitations. First, the identification of specific threshold concepts can be subjective, particularly when a subject matter is multidisciplinary (Dawes, 2019). Second, not all students will engage in the same transformative learning process and may require additional time to

reach a level of understanding that may come quickly to other students (Cousin, 2006). Third, by focusing solely on threshold concepts, there is a risk that other aspects of teaching and learning may be neglected (Kiley and Wisker, 2009) (e.g., pedagogical strategies, assessment, and feedback). Finally, incorporating threshold concepts into extant curriculum can be challenging and require extensive restructuring and alignment with existing departmental/institutional goals and objectives (Meyer and Land, 2006). Despite these limitations, extant empirical evidence has outlined that threshold concept theory remains a valuable framework for enhancing the teaching and learning of challenging concepts such as blockchain technology which can ultimately lead to transformative learning experiences.

#### **4.0 Next Steps**

Currently, very little is known from a Blockchain technology perspective pertaining to what concepts can be constituted as threshold concepts. While some foundational ideas, such as distributed ledgers, consensus mechanisms, and cryptographic hashing, might be considered as potential threshold concepts, a comprehensive and systematic exploration is still required (Schneider, B., and Azan, 2022). In order to shed light on the two research questions presented in section 1.0 the next steps for this research will focus on both the preliminal instigative mode and liminal reconstitutive mode features (Figure 2) in the context of the teaching and learning case study of a third level higher education blockchain technology module which is due to commence in September 2024. Ultimately, studying these two modes using the threshold concept framework (Figure 1) will identify how learners arrive at the postliminal consequential mode. It will also identify core concepts inherent to Blockchain Technology which manifest as threshold concepts (research question 1) and also identify what coping mechanisms learners use to help them navigate threshold concepts when they encounter them (research question 2). Figure 3 highlights the research conceptual model for this study comprising both research questions.



**Figure 3.** Research Study Conceptual Model

This study will adopt the research methodology strategy depicted in Table 1. Similar, research strategies have been used to identify threshold concepts in other disciplines (see Clohessy and English, 2022).

Philosophical assumptions	Philosophical stance	Research strategy	Research methods
Epistemology and Ontology	Social Constructivism	Case Study <ul style="list-style-type: none"> <li>• Third Level Education Institutions Delivering Blockchain Modules.</li> </ul>	Mixed-Method <ul style="list-style-type: none"> <li>• Phase 1: Interview – Lecturers</li> <li>• Phase 2: Questionnaire – Students</li> <li>• Phase 3: Focus Group – Students</li> </ul>

**Table 1.** Research Methodology Strategy

Timmermans and Meyer (2019) emphasize that when examining threshold concepts in educational contexts, research should incorporate diverse insights from a range of stakeholders involved in the teaching and learning process. This study will encompass several research phases, tapping into the experiences of both educators and students. Phase 1 will involve the use of interviews with blockchain lecturers who possess diverse teaching experiences in teaching blockchain technologies. Interviews will be conducted until theoretical saturation is reached and no new themes can be identified (Strauss and Corbin, 1998). These thematic findings from phase 1 will form the basis for creating the study questionnaire which will be used in phase 2 and disseminated to blockchain students. This questionnaire will be sound boarded with these lecturers and recommendation sought prior to being issued to students. The final phase of the research will incorporate the use of focus groups session with students. Focus groups will continue until theoretical saturation is reached and will be carried out in line with recommendations by Morgan (1996).

The progression of each research phase will be systematic, with findings from one phase guiding the formulation of the research tool for the subsequent phase. The main research approach chosen for this investigation is a mixed-methods approach. A mixed-methods study, as described by Creswell and Clark (2017), integrates both quantitative and qualitative data gathering techniques to present results from a singular investigation. In this research, the questionnaire aims to identify trends and frequencies. On the other hand, the qualitative components (e.g., interviews and focus groups) will seek to capture and articulate the viewpoints of the research participants. Additionally, the qualitative aspect will facilitate a thorough exploration of any discrepancies that surfaced during the analysis of the questionnaire data. In terms of data analysis, the focus groups and interviews will be audio recorded with the express consent of the research participants. Quirkos will be used to find organise and insights from unstructured qualitative data (Harvey and Powell 2020). A two-stage coding process will then be carried out incorporating initial coding and axial coding in order to identify salient themes within the data (Saldaña, 2021). For the online questionnaire, data collection and analysis will be conducted using Microsoft Forms. This platform facilitates the extraction of meaningful insights from the gathered data through methods such as cross-distribution analysis, association rule analysis,

correlation study, and sentiment evaluation. From an ethical perspective, participant consent will be obtained, and a research study background information document provided to them to ensure that they are fully aware of the study's scope and purpose before involvement (Sim and Waterfield, 2019). This research phase will be thoroughly voluntary, with participants retaining the freedom to withdraw from the study at any point. Regarding the safeguarding of participant data, all identifiable information will be anonymised using pseudonyms, and access to this sensitive data will be strictly limited to researchers who are officially authorised to work with the data (Allmark et al., 2009).

Ultimately, it is hoped that this research will not only provide educators with a clear roadmap for curriculum (re)development but also ensure that third level students grasp the critical aspects of Blockchain technology facilitating a more profound and lasting understanding. Recognising and addressing these threshold concepts is paramount, as it can significantly enhance the quality of blockchain education and better prepare students for real-world applications of Blockchain technology.

## References

Allmark, P., Boote, J., Chambers, E., Clarke, A., McDonnell, A., Thompson, A., & Tod, A. M. (2009). Ethical issues in the use of in-depth interviews: literature review and discussion. *Research Ethics*, 5(2), 48-54.

Alsamhi, S. H., Almalki, F. A., Afghah, F., Hawbani, A., Shvetsov, A. V., Lee, B., & Song, H. (2021). Drones' edge intelligence over smart environments in B5G: Blockchain and federated learning synergy. *IEEE Transactions on Green Communications and Networking*, 6(1), 295-312.

Andoni, M., Robu, V., Flynn, D., Abram, S., Geach, D., Jenkins, D., ... & Peacock, A. (2019). Blockchain technology in the energy sector: A systematic review of challenges and opportunities. *Renewable and sustainable energy reviews*, 100, 143-174.

Berger, (2023). The nuance of blockchain learning in the education space. Available at:<https://www.forbes.com/sites/rodberger/2023/01/24/the-nuance-of-blockchain-learning-in-the-education-space/> [Accessed Date 12/08/2023].

Boyd, R. D., & Myers, J. G. (1988, October–December). Transformative Education. *International Journal of Lifelong Education* 7(4): 261–284.

Chen, H., Chen, Z., Lin, F., & Zhuang, P. (2021). Effective management for blockchain-based agri-food supply chains using deep reinforcement learning. *IEE Access*, 9, 36008-36018.

Clohessy, T., Acton, T., Godfrey, R., & Houston, M. (2018) ‘The Adoption of Blockchain in Ireland: Examining the Influence of Organisational Factors. A National University of Ireland Galway and Blockchain Association of Ireland Industry Report.’ <https://aran.library.nuigalway.ie/handle/10379/7352> [Accessed Date 12/08/2023].

Clohessy, T., & English, M. (2022). Bridging the gaps in information systems: a threshold concepts and troublesome knowledge perspective. *Irish Educational Studies*, 1-25.

Clohessy, T (2023). Assimilation of the Blockchain Exploring the Impact of Blockchain Technology on Supply Chain Management. *Blockchain in Supply Chain Transformation*. CRC Press.

Clohessy, T., & Heaslip, G. (2022) Exploring the Impact of Decentralized Autonomous Supply Chains (DASCs) in the Agricultural Sector: An Innovation Assimilation Perspective. *Blockchain IEEE Technical Report*.

Cousin, G. (2006). Threshold concepts, troublesome knowledge and emotional capital. *Overcoming barriers to student understanding: Threshold concepts and troublesome knowledge*, 134-147.



Cousin, G. (2008). Threshold concepts: Old wine in new bottles or a new form of transactional curriculum inquiry? In *Threshold concepts within the disciplines* (pp. 259-272). Brill.

Creswell, J. W., and V. L. Plano Clark. 2017. *Designing and Conducting Mixed Methods Research*. 3rd ed. Thousand Oaks, CA: SAGE Publications.

Dawes, L. (2019). Through faculty's eyes: Teaching threshold concepts and the framework. *portal: Libraries and the Academy*, 19(1), 127-153.

Davies, P., & Mangan, J. (2007). Threshold concepts and the integration of understanding in economics. *Studies in Higher Education*, 32(6), 711-726.

Freire, PG. (1970). *Pedagogy of the oppressed*. New York: Herter and Herter.

Froehlich, M., Vega, J., Pahl, A., Lotz, S., Alt, F., Schmidt, A., & Welp, I. (2022, September). Prototyping with Blockchain: A Case Study for Teaching Blockchain Application Development at University. In *International Conference on Interactive Collaborative Learning* (pp. 1005-1017). Cham: Springer International Publishing.

Ghiro, L., Restuccia, F., D'Oro, S., Basagni, S., Melodia, T., Maccari, L., & Cigno, R. L. (2021). A blockchain definition to clarify its role for the Internet of Things. In *2021 19th Mediterranean Communication and Computer Networking Conference (MedComNet)* (pp. 1-8). IEEE.

Habermas, J. (1971). *Knowledge of human interests*. Boston: Beacon.

Habermas, J. (1984). *The theory of communicative action*. Vol. 1: Reason and the rationalization of society (T. McCarthy, Trans.). Boston: Beacon.

Harvey, J., and J. Powell. 2020. "Factors Influencing the Implementation of Self-Management Solutions in Healthcare: An Interview Study with NHS Managers." *British Journal of Healthcare Management* 26 (3): 61–70.

Kegan, R. (1982). *The evolving self: Problem and process in human development*. Cambridge, MA: Harvard University Press.

- Kitchenham, A. (2008). The evolution of John Mezirow's transformative learning theory. *Journal of transformative education*, 6(2), 104-123.
- Kiley, M., & Wisker, G. (2009). Threshold concepts in research education and evidence of threshold crossing. *Higher Education Research & Development*, 28(4), 431-441.
- Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Labouseur, A. G., Johnson, M., & Magnusson, T. (2019). Demystifying blockchain by teaching it in computer science: adventures in essence, accidents, and data structures. *Journal of Computing Sciences in Colleges*, 34(6), 43-56.
- Land, R., Rattray, J., & Vivian, P. (2014). Learning in the liminal space: A semiotic approach to threshold concepts. *Higher education*, 67, 199-217.
- Land, R., & Meyer, J. H. (2010). Threshold concepts and troublesome knowledge (5): Dynamics of assessment. In *Threshold concepts and transformational learning* (61-79). Brill Sense.
- Lucas, U., & Mladenovic, R. (2007). The potential of threshold concepts: an emerging framework for educational research and practice. *London Review of Education*.
- Meyer, J., & Land, R. (2003). *Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising within the disciplines*. Edinburgh: University of Edinburgh.
- Meyer, J. H. F. & Land R. (2005). Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning. *Higher Education*, 49(3), 373-388.
- Meyer, J. H., & Land, R. (2006). Threshold concepts and troublesome knowledge: An introduction. In *Overcoming barriers to student understanding*. Routledge.
- Mezirow, J. (1981). A critical theory of adult learning and education. *Adult Education Quarterly*, 32, 3-24.

- Mezirow, J. (1978). Perspective transformation. *Adult Education*, 28 (2), 100 – 109.
- Mezirow, J. (1990). *Fostering critical reflection in adulthood*. San Francisco: Jossey-Bass.
- Morgan, D. L. (1996). Focus groups. *Annual review of sociology*, 22(1), 129-152.
- Nguyen, D. C., Ding, M., Pham, Q. V., Pathirana, P. N., Le, L. B., Seneviratne, A., ... & Poor, H. V. (2021). Federated learning meets blockchain in edge computing: Opportunities and challenges. *IEEE Internet of Things Journal*, 8(16), 12806-12825.
- Saldaña, J. 2021. *The Coding Manual for Qualitative Researchers*. Newcastle, UK: Sage.
- Schneider, B., & Azan, W. (2022). Perceptions and Misconceptions of Blockchain: The Potential of Applying Threshold Concept Theory. In *2022 IEEE 6th International Conference on Logistics Operations Management (GOL)* (pp. 1-6). IEEE.
- Shanahan, M., Foster, G., & Meyer, J. H. (2006). *Operationalising a threshold concept in economics: A pilot study using multiple choice questions on opportunity cost* (Doctoral dissertation, University of Bristol).
- Sim, J., & Waterfield, J. (2019). Focus group methodology: some ethical challenges. *Quality & quantity*, 53(6), 3003-3022.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research techniques*.
- Tian, Y., Li, T., Xiong, J., Bhuiyan, M. Z. A., Ma, J., & Peng, C. (2021). A blockchain-based machine learning framework for edge services in IIoT. *IEEE Transactions on Industrial Informatics*, 18(3), 1918-1929.
- Timmermans, J. A., and J. H. Meyer. 2019. “A Framework for Working with University Teachers to Create and Embed ‘Integrated Threshold Concept Knowledge’ (ITCK) in Their Practice.” *International Journal for Academic Development* 24 (4): 354–368.

Xu, K., Wang, Z., & Guo, F. (2021, April). The " Four-Level Guidance" Blockchain Practice Teaching Model for Undergraduate. In *Proceedings of the 6th International Conference on Information and Education Innovations* (pp. 30-34).

Vandenberg, B. (1991). Is epistemology enough? An existential consideration of development. *American Psychologist*, 46(12), 1278.

Zepke, N. (2013). Threshold concepts and student engagement: Revisiting pedagogical content knowledge. *Active Learning in Higher Education*, 14(2), 97-107.

Zheng, Z., Xie, S., Dai, H. N., Chen, X., & Wang, H. (2018). Blockchain challenges and opportunities: A survey. *International journal of web and grid services*, 14(4), 352-375.

# Teachers' Perceptions of Students' Use of Generative AI in Summative Assessments at Higher Education Institutions: An Exploratory Study

**Chekfoung Tan**

*UCL Centre for Systems Engineering, University College London*  
[Chekfoung.Tan@ucl.ac.uk](mailto:Chekfoung.Tan@ucl.ac.uk)

**Muna M. Alhammad**

*MIS Department, King Saud University*  
[malhammad@ksu.edu.sa](mailto:malhammad@ksu.edu.sa)

**Marta Stelmaszak**

*The School of Business, Portland State University*  
[s.marta@pdx.edu](mailto:s.marta@pdx.edu)

*Completed Research*

## **Abstract**

*The rise of generative AI (GAI) is affecting various sectors, including higher education (HE). In HE, educators are grappling with students' use of GAI, which might infringe upon academic integrity. Given the pervasiveness of this technology, particularly through free AI tools, it should be utilised for its benefits rather than merely blocked. Thus, responsible use of AI in higher education is essential. However, encouraging students to openly declare their use of AI in summative assessments, for transparency purposes, has led to fears of negative perception and potential marking down by teachers. There is currently limited research in this area. Hence, this study aims to explore the boundary of students using GAI in assessments by gathering views from teachers through an exploratory survey. It contributes theoretically to responsible AI literature, extending it to HE, and practically by developing guidelines for AI use in student assessments, potentially informing university policy.*

**Keywords:** Generative AI, Exploratory Research, Responsible AI, Teachers' Perception, Higher Education

**Acknowledgement:** This work is inspired by the outcome of the UCL AI ChangeMakers project back in the summer of 2023. We would like to thank UCL Arena Centre for the opportunity, as well as the student co-creators, Liam Perez and Kartik Gupta.

## 1.0 Introduction

The term “Artificial Intelligence” (AI) was officially coined in 1956 by a group of computer scientists during the Dartmouth Summer Research Project on Artificial Intelligence (DSRPAI) at Dartmouth College in New Hampshire (Haenlein & Kaplan, 2019). AI is seen as a system capable of interpreting external data accurately, learning from this data, and using these learnings to achieve specific goals and tasks through flexible adaptation (Haenlein & Kaplan, 2019).

The inception of AI use in teaching and learning can be traced back to 1924 when Sidney Pressey used a machine to assist students in finding correct answers to multiple-choice questions (Namatherdhala et al., 2022). Since then, AI has been adopted to personalise learning for students, tailoring to their learning goals and preferences based on performance, demographics, and behavioural information (Kaplan-Rakowski et al., 2023). However, OpenAI shocked various sectors in year 2022, including Higher Education (HE), with the release of ChatGPT, a chatbot driven by Generative AI (GAI) (OpenAI, 2022). GAI, trained on large language models (LLMs), can generate human-like text based on given prompts or contexts and is capable of performing natural language processing tasks such as text completion, conversation generation, and language translation (Baidoo-Anu & Ansah, 2023). The immediate concern within HE is that students might use this tool for writing assignments, coursework, or even answering open-book exams, potentially breaching academic integrity (Michel-Villarreal et al., 2023). However, a recent report from UNESCO (2023), suggests that the use of GAI tools is not entirely negative if used responsibly.

Nevertheless, since the emergence of GAI, there has been growing research on its adoption and how teachers or students perceive it in HE, as seen in Chan and Hu (2023) and Amani et al. (2023). However, limited research exists on teachers' perceptions of students using GAI in their summative assessments, such as written coursework, essays, online open-book exams, presentations, and media productions (e.g., videos, films, animations). While universities have issued some guidance on declaring the use of GAI tools, students might be reluctant to disclose their use of GAI if they fear negative perceptions from teachers. This research aims to address this research question - *How do teachers perceive students' use of GAI in summative assessments at HE institutions?* This paper is organised as follows. Section 2 covers related work on the application of GAI in HE, incorporating a responsible lens. Section 3 illustrates the research

methodology, guided by an exploratory survey, and Section 4 presents the results. The paper concludes with discussions, research implications and future work in Section 5.

## **2.0 Related Work**

### **2.1 GAI in Higher Education**

The rise of GAI has been met with intense interest in HE. Universities were quick to react and develop guidelines for the use of GAI, initiated working groups, and met in cross-organisational fora to discuss the impact of this technology. Multiple practitioner publications highlighted the opportunities and challenges of GAI for the sector (e.g., Hodges & Ocak, 2023; Schroeder, 2023).

Research soon followed. Existing papers focus on understanding the potential uses of GAI in HE and conceptualising the challenges. For example, Michel-Villarreal et al. (2023) outline that GAI can be used to generate answers to questions and ideas for essays, provide feedback, simulate a tutor. Research into students' perspectives shows a recognition of the potential of GAI in personalised learning support, writing, brainstorming, and research and analysis (Chan & Hu, 2023). From a teacher's perspective, it can create lesson plans, develop resources, and even assess written work (Baidoo-Anu & Ansah, 2023). For example, Kim et al. (2019) show that ChatGPT trained on human-graded essays could grade high school student essays with a correlation of 0.86 with human graders. The UNESCO (2023) report outlines ten roles illustrating how ChatGPT could be useful in the teaching and learning process. However, less research interest so far has been given to administrators' perspectives, and yet GAI has been found to be useful in administrative support with repetitive or tedious administrative tasks (Chan & Hu, 2023), creating manuals, or developing policy documents (Yeralan & Lee, 2023). Yet, the use of GAI in HE is a contentious topic, with multiple responsibility and ethics implications.

### **2.2 Responsible Use of AI in HE**

The growth in AI in general and GAI in particular has been accompanied by increased interest in responsible AI, that is a growing consensus that the use of AI should follow principles consistent with user expectations, organisational values, and societal laws and norms (Mikalef et al., 2022). Table 1 distils these dimensions into the principles of responsible AI in HE. Responsible AI is often discussed along the principles of fairness,

transparency, and accountability. For example, the principle of fairness as applied to HE suggests that the use of AI in this sector should enable inclusion and diversity and not lead to discriminatory outcomes between students. Transparency suggests that the use of AI in HE should be openly communicated and facilitate traceability. Accountability entails using AI in line with the set policies and regulations.

Responsible GAI comes into particular focus in HE in discussions around assessment. While GAI can be used by students in assessed work to help generate ideas, conduct research, or improve writing (Smolansky et al., 2023), it can also be deployed in ways that may go against the principles of responsible AI in HE. GAI can be potentially used by students to submit work prepared by GAI rather than themselves and not disclose it, which goes against the principle of transparent use. The use of GAI in assessment may reduce students' accountability for their own intellectual work and propagate plagiarism against university policies. Fairness may be at stake as well, as deploying GAI in assessment may lead to obtaining better marks. Both students and teachers admit that GAI has an impact on a range of assessment types, from short answer questions, through essays, to creative work and presentations (Smolansky et al., 2023).

By acknowledging assessments may particularly be impacted by GAI, attempts have been made to minimise this by implementing the responsible AI principles. For example, researchers investigate how to develop AI-proof forms of assessment (Rudolph et al., 2023), and how to develop tools capable of identifying the use of GAI (Lacey & Smith, 2023). Existing research suggests that teachers are concerned about the use of GAI in assessment and aim to work out alternatives that either preclude such a possibility or make it easier to detect it (Smolansky et al., 2023). This indicates that teachers may be against the GAI use by students, but little research has been conducted in this area.

### **3.0 Research Methodology**

Following Chan and Hu (2023) and Chan and Lee (2023), this study employed the exploratory study by using an online survey methodology to investigate teachers' perceptions on students' use of GAI tools in their summative assessment within HE. The survey encompassed a set of closed questions and two open-ended questions, with the goal of acquiring a holistic comprehension of participants' viewpoints.



<b>Responsible AI principles</b>	<b>AI Use in HE</b>	<b>Sources adapted</b>
Transparency	Must be disclosed, communicated, and understood by those involved, with clear oversight principles established.	(Clarke, 2019; Jobin et al., 2019; Mezgár & Váncza, 2022; Mhlanga, 2023; Noble & Dubljević, 2022)
Justice and fairness	Should be unbiased, promoting justice, diversity, and inclusion while addressing inequalities in access to education.	(Jobin et al., 2019; Mhlanga, 2023; Noble & Dubljević, 2022)
Non-maleficence	Must prioritise student safety and wellbeing, ensuring protection from potential threats.	(Clarke, 2019; Jobin et al., 2019; Mezgár & Váncza, 2022; Noble & Dubljević, 2022)
Responsibility / Accountability	Must follow accountability and liability rules, align with policies, and have clear regulations and consequences for non-compliance.	(Clarke, 2019; Jobin et al., 2019; Mhlanga, 2023; Noble & Dubljević, 2022)
Professional responsibility	Should be used purposefully, in collaboration with institutional staff, and maintain dialogue with relevant bodies.	(Noble & Dubljević, 2022)
Privacy / Consent	Must protect user privacy, provide consent options, and allow data use control.	(Jobin et al., 2019; Mhlanga, 2023; Noble & Dubljević, 2022)
Beneficence / Promotion of Human Values / Perceived Benefits / Education Values	Should prioritise wellbeing, the common good, and human values, aiming to benefit society, advance civilisation, and uphold human rights.	(Chan & Hu, 2023; Clarke, 2019; Jobin et al., 2019; Mezgár & Váncza, 2022; Noble & Dubljević, 2022)
Freedom and autonomy / Human control of technology / Dignity / Digital Literacy	Should be deployed to empower all stakeholders, remaining under human control and open to review. HE institutions must educate students on AI's functions, uses, limitations, and ethical considerations, enabling informed choices in their academic and future endeavours.	(Clarke, 2019; Jobin et al., 2019; Mezgár & Váncza, 2022; Mhlanga, 2023; Noble & Dubljević, 2022)
Trust	Stakeholders can trust AI to unlock potential and add significant value to pedagogy and teaching.	(Jobin et al., 2019; Mezgár & Váncza, 2022)
Sustainability and Inclusivity	HE institutions should embed AI sustainably without compromising core values, using it to foster sustainable societies and champion inclusive education, including tailored support for disabled students and top-tier education to students in remote learning.	(Jobin et al., 2019; Noble & Dubljević, 2022)

Solidarity	AI benefits in teaching and learning must be equitably shared among all stakeholders, ensuring no undue advantage for any group.	(Jobin et al., 2019)
Quality Assurance	AI deployment must be underpinned by rigorous regulations and standards, subject to regular review.	(Clarke, 2019)
Data Accuracy	It's vital that data from GAI tools is accurate, and both teachers and learners should critically assess information and cross-check with trustworthy sources.	(Mhlanga, 2023)
Robustness and Resilience	All stakeholders in AI in HE must ensure its robust and resilient, with responsibility proportional to benefits, data sensitivity, and potential risks in education.	(Clarke, 2019)
General Impact	Should evaluate the positive and negative consequences and implications of employing AI.	(Clarke, 2019)
Human-centred Design	AI systems should be customised for students and teachers, involving them in development, regularly gathering their feedback, and adapting to their academic experiences.	(Mezgár & Váncza, 2022)
Continuous Learning and Adaptability / Learning Skills	AI systems must continually learn and adapt, adjusting to student feedback, evolving educational standards, and AI advancements.	(Chan & Hu, 2023)
Ethics	AI research tools must uphold research ethics, including minimising harm, ensuring informed consent, and maintaining data confidentiality.	(Amani et al., 2023)

**Table 1. Responsible AI principles in HE**

A convenience sampling approach from Edgar and Manz (2017) was adopted to enlist participants for the study. The survey link was distributed to teachers or teaching academics in HE institutions across diverse international social media groups. Participants were presented with an informed consent form on the online platform, ensuring their awareness of the study's objectives and their rights as participants. The survey questions were adapted from the roles of how GAI tools could be applied in HE issued by UNESCO (2023) and responsible AI principles in Table 1. A five-point Likert scale (1-Strongly Disagree; 2-Disagree; 3-Neutral; 4-Agree; 5-Strongly Agree) was used for these questions. The collected data underwent rigorous descriptive analysis, examining the quantitative responses to closed questions. Thematic analysis was employed for analysing the data collected from the open-ended questions. Section 4 discusses the results.

## 4.0 Results

### 4.1 Demographics

Table 2 shows the full demographic information. Through the survey, 66 responses were collected, with most respondents falling into the 40-49 age group (39.4%) and being male (60.6%). Most respondents teach in the STEM field (56.1%) and have more than ten years of teaching experience (57.6%). Geographically, most respondents primarily teach in the United Kingdom (50.0%), followed by Saudi Arabia (16.7%). Other countries represented include the United States of America, the United Arab Emirates, Malaysia, China, Singapore, Ireland, and Portugal.

Characteristics	Count (n)	Percentage (%)
<b>Age Group</b>		
20 and below	1	1.5%
30-39	19	28.8%
40-49	26	39.4%
50-59	14	21.2%
60-69	5	7.6%
70 and older	1	1.5%
<b>Gender</b>		
Female	26	39.4%
Male	40	60.6%
<b>Teaching Domain</b>		
STEM	37	56.1%
Non-STEM	29	43.9%
<b>Level of teaching experience</b>		
Less than two years	2	3.0%
Two to five years	9	13.6%
Five to ten years	17	25.8%
More than ten years	38	57.6%
<b>Country where the participant primarily teaches</b>		
China	3	4.5%
Ireland	1	1.5%
Malaysia	4	6.1%
Portugal	1	1.5%
Saudi Arabia	11	16.7%
Singapore	2	3.0%
United Arab Emirates	6	9.1%
United Kingdom	33	50.0%
United States of America	5	7.6%

**Table 2. Demographic Information**

## 4.2 Familiarity with the GAI Tools and Frequency of Use

As shown in Table 3, most respondents have a moderate familiarity with GAI tools (42%), and 79% of them have created an account and used GAI tools for either personal or educational purposes. This indicates a high level of engagement with GAI tools among the respondents.

Characteristics	Count (n)	Percentage (%)
<b>Familiarity with GAI tools</b>		
Not familiar at all	1	1.5%
Slightly familiar	12	18.2%
Moderately familiar	28	42.4%
Very familiar	21	31.8%
Extremely familiar	4	6.1%
<b>GAI tools account creation for any purpose (either personal or educational)</b>		
Yes	55	83.3%
No	11	16.7%

**Table 3. Respondents' Familiarity and Engagement with General AI Tools**

Based on the 55 respondents who have signed up for a GAI tool account, Table 4 shows that the majority have used GAI tools for a period ranging from 1 to 6 months, with the highest percentages observed at 4 and 6 months, both at 14.5%. Adoption appears to decrease as the duration increases, with only 3.6% of respondents using GAI tools for more than 12 months. These findings suggest that most respondents are relatively new to using GAI tools, with a significant drop in usage beyond 6 months. This may be also because the most popular GAI tools have only emerged approximately 11 to 12 months ago, thereby providing a limited timeframe for the respondents to integrate these tools into their daily activities.

According to Table 5, ChatGPT is the most frequently used tool, with 40.7% of respondents using it once a week, 24.1% using it twice a week, 7.4% using it three times a week, and 25.9% using it more than three times a week. Only 1.9% of respondents never use ChatGPT. Google Bard, Bing Chat, and Microsoft 365 Copilot are used much less frequently, with more than 70% of respondents never using these tools. Snapchat AI also has a high percentage of respondents who never use it (88.6%), with small percentages using it at varying frequencies.

Months of using GAI tools	Count (n)	Percentage (%)
1 month	6	10.9%
2 months	5	9.1%
3 months	7	12.7%
4 months	8	14.5%
5 months	6	10.9%
6 months	8	14.5%
7 months	5	9.1%
8 months	1	1.8%
9 months	1	1.8%
10 months	5	9.1%
12 months	1	1.8%
11 months	0	0.0%
More than 12 months	2	3.6%

**Table 4. Months of Using GAI Tools**

The 'Others' category has a more even distribution across different frequencies, totalling 18.9%. The GAI tools included in this category are GrammarlyGo, Ernie Bot, Wordtune, Perplexity, Midjourney, Gamma, Notion, Hypotenuse AI, Writesonic, Invideo.io, and customised GAI tools utilising the GPT-4 API.

GAI tools		Frequency					Total
		Never	Once a week	Twice a week	Three times a week	More than three times a week	
ChatGPT	Count (n)	1	22	13	4	14	54
	Percentage (%)	1.9%	40.7%	24.1%	7.4%	25.9%	100.0%
Google Bard	Count (n)	34	8	2	2	2	48
	Percentage (%)	70.8%	16.7%	4.2%	4.2%	4.2%	100.0%
Bing Chat	Count (n)	34	6	3	2	3	48
	Percentage (%)	70.8%	12.5%	6.3%	4.2%	6.3%	100.0%
Microsoft 365 Copilot	Count (n)	41	2	1	0	1	45
	Percentage (%)	91.1%	4.4%	2.2%	0.0%	2.2%	100.0%
Snapchat AI	Count (n)	39	3	1	1	0	44
	Percentage (%)	88.6%	6.8%	2.3%	2.3%	0.0%	100.0%
Others	Count (n)	30	1	3	0	3	37
	Percentage (%)	81.1%	2.7%	8.1%	0.0%	8.1%	100.0%

**Table 5. Frequency of Usage of Various GAI Tools**

### 4.3 Teachers' Perceptions of Students' Use of Generative AI in Their Summative Assessments

#### 4.3.1 Roles of GAI Tools in Students' Summative Assessments

Based on the findings as shown in Table 6, the respondents generally have a positive perception of students using GAI tools in their summative assessments, as evidenced by the median scores of 4 for all statements. This suggests that most respondents are open to, or accepting of, the idea of students leveraging GAI tools for various aspects of their assessments. However, the mean scores reveal some differences in the level of acceptance across different uses of GAI tools.

No	Statement	Median	Mean	Standard Deviation
1	I can accept students using GAI tools to write queries and examine alternative responses for their assessments.	4.00	3.32	1.10
2	I can accept students entering prompts into various GAI tools, following the structure of a conversation or debate, to produce critical arguments for their assessments.	4.00	3.29	1.08
3	I can accept that working in groups, students use any GAI tools to find out information to complete tasks and assignments.	4.00	3.58	1.10
4	I can accept that students use various GAI tools to provide personalised feedback to them, based on information provided by students or teachers (e.g., formative feedback).	4.00	3.44	1.22
5	I can accept that students explain their current level of understanding related to the assessments to various GAI tools, and ask for ways to help them study the related material and prepare for other tasks in the assessment.	4.00	3.65	1.10
6	I can accept that students ask various GAI tools for ideas about how to extend their learning after receiving the scores for their assessment.	4.00	3.76	1.08
7	I can accept that students interact with various GAI tools in a tutorial-type dialogue, and then ask the tool to produce a summary of their current state of knowledge for their assessment.	4.00	3.56	1.12
8	I can accept that students ask various GAI tools to proofread the language of their assessment.	4.00	3.79	0.98

**Table 6. Roles of GAI Tools**

Respondents are most accepting of students using GAI tools for proofreading the language of their assessments (Statement 8, Mean: 3.79) and seeking ideas for extending their learning post-assessment (Statement 6, Mean: 3.76). This indicates a recognition of the value of GAI tools as supportive resources for enhancing learning experiences and improving work quality. However, respondents are less accepting of students using GAI tools to generate critical arguments for their assessments (Statement 2, Mean, 3.29). This suggests a preference for students to develop their critical thinking and produce arguments independently, rather than relying on GAI tools for such complex cognitive tasks. Statements 3 and 7 both have Mean scores above 3.5, but Standard Deviations above 1.10, indicating a moderate level of acceptance but with some variability in responses.

#### 4.3.2 Responsible Use of GAI Tools

Table 7 presents the results pertaining to students' responsible use of GAI tools in their summative assessments from various perspectives. Most respondents express a general acceptance towards the responsible utilisation of GAI tools by students in their summative assessments. This is reflected in the median values of 4.00 for Statements 1, 2, 3, 4, 6, 7, 8, 9, 11, 13, 14, 15, 17, 18, and 20, indicating agreement with the statements. The Mean values for these statements range from 3.45 to 3.97, further cementing this stance.

No	Statement	Median	Mean	Standard Deviation
1	I can accept students using GAI tools to support the production of their assessments, such as proofreading and idea creation, as long as they declare the use of it, but they must not use it to produce or write the entire assessment, such as content generation.	4.00	3.79	1.14
2	If I refrain from judging their use of GAI tools, my students will likely feel more at ease both using it and declaring its use.	4.00	3.71	0.91
3	If students are using GAI tools ethically, it will not affect the way I mark their assessment.	4.00	3.83	1.10
4	I can accept students' use of GAI tools in their assessments as long as they are aware of the tools are not substitute for human tutors.	4.00	3.68	1.15
5	I can trust my students to use GAI tools in their assessments responsibly and ethically.	2.00	2.64	1.05

6	I can accept students citing information from GAI tools in their assessment by justifying its relevance and confirming the accuracy of the information by comparing it with other sources or using personal judgement	4.00	3.52	1.11
7	I can accept that students use GAI tools, employing various plug-ins or functions, for idea creation when producing their assessments.	4.00	3.53	1.06
8	I believe that all students should be given access to the same GAI tools for their assessments to ensure inclusivity and equity.	4.00	3.59	1.20
9	I believe that students' use of GAI tools in their assessments will help them learn in an efficient way, as the tools can provide preliminary feedback.	4.00	3.45	1.10
10	I think GAI tools is a great tool for supporting students' assessment due to anonymity.	3.00	2.98	1.22
11	I believe that GAI tools could contribute to academic dishonesty behaviours.	4.00	3.70	1.16
12	I believe that using GAI tools to complete assignments undermines the value of university education.	3.00	3.00	1.15
13	GAI tools may reduce students' opportunities to interact with peers and socialise during coursework completion.	4.00	3.41	1.08
14	GAI tools may impede students' cultivation of generic or transferable skills, including teamwork, problem-solving, and leadership skills.	4.00	3.44	1.08
15	There's a risk that students might become overly dependent on GAI tools.	4.00	3.97	1.12
16	I believe that using GAI tools, to write essays or generate answers can enhance originality and creativity in students' work.	3.00	2.89	1.10
17	I believe that GAI tools can bolster students' digital competence.	4.00	3.62	1.00
18	I believe that GAI tools can help students save time.	4.00	3.67	1.09
19	I think that GAI tools can help students in becoming better writers.	3.00	3.23	1.17
20	In the long run, the integration of GAI tools in higher education is likely to have a positive impact on teaching and learning, influencing how students are taught and assessed.	4.00	3.68	1.03

**Table 7. Responsible Use of GAI Tools**



As per Statement 3, respondents generally concur that if students employ GAI tools responsibly, it will not impinge the way they mark assessments (Mean: 3.83, SD: 1.10). Statements 1 (Mean: 3.79; SD: 1.14), 4 (Mean: 3.68; SD: 1.15), 6 (Mean: 3.52; SD: 1.11), and 7 (Mean: 3.53; SD: 1.06) imply a conditional acceptance of students using GAI tools in their summative assessments for tasks such as proofreading, idea generation, acknowledging that GAI tools are not substitutes for human judgement, and critically evaluating the information generated by GAI tools.

A significant concern from the respondents is the lack of trust in students to use GAI tools responsibly, as shown in Statement 5 (Median: 2.00; Mean: 2.64), although the Standard Deviation indicates a moderate range of opinions among respondents on this issue. There are concerns about the potential for academic dishonesty, as indicated in Statement 11, although the Standard Deviation suggests moderate agreement among respondents on this matter (Mean: 3.76; SD: 1.16). Similarly, Statement 15 shows that respondents are concerned about students becoming overly dependent on GAI tools (Mean: 3.97; SD: 1.12). The role of anonymity in assessments, as in Statement 10, is a contentious issue (Median: 3.00; Mean: 2.98, SD: 1.22), and the highest SD indicates a diverse range of respondents' opinions.

Despite these concerns, respondents see potential benefits in GAI tools, such as enhancing originality and creativity, as indicated in Statement 16 (Median: 3.00, Mean: 2.89, SD: 1.10); bolstering digital competence, as in Statement 17 (Median: 4.00, Mean: 3.62, SD: 1.00); and saving time, as in Statement 18 (Median: 4.00, Mean: 3.67, SD: 1.09). In the long run, as indicated in Statement 20, respondents believe that the integration of GAI tools in higher education is likely to have a positive impact on teaching and learning (Median: 4.00; Mean: 3.68, SD: 1.03).

#### **4.4 Governance of the Use of GAI Tools in Higher Education**

Based on the open-ended questions where respondents were asked for additional comments about their perception of students' use of GAI tools in their summative assessments, they acknowledged the inevitability of students utilising GAI tools. Table 8 presents the results of a thematic analysis on the governance of GAI tools in HE institutions, which include the codes clustered into four main themes (*Guidelines for Responsible Use of GAI Tools, Regulations, Assessment and Academic Integrity, Training and Education*) along with the relevant sample quotes from the respondents.

Themes	Codes	Sample Quotes
Guidelines for Responsible Use of GAI Tools	HE Institutions Guideline	Respondent 52 - <i>"The HEI bodies should introduce, not only Russel group, which has produced a guidelines, but other at national or even international level should have the ethical guidelines, and acceptable use of GAI tools in HEIs."</i>
	Transparency	Respondent 12 - <i>"A full transcript of interaction between student and AI plus a summary of used inputs should be submitted with any thesis."</i> Respondent 19 - <i>"Student must provide a log, with timestamps when they use GAI for particular assignments"</i> Respondent 40 - <i>"Ensure transparency in AI systems and algorithms, including disclosing the sources of data, the methods used, and potential biases."</i>
	Accountability	Respondent 32 - <i>"...student takes ownership of their work and use GAI as a supplementary method to enhance to quality of their work."</i> Respondent 23 - <i>"Lecturers should be in control and aware of the sentence pattern created using GAI."</i> Respondent 40 - <i>"Hold individuals and departments accountable for their use of AI."</i>
	Bias Mitigation	Respondent 40 - <i>"Address bias in AI algorithms and data to ensure fairness and inclusivity."</i>
	Data Privacy and Security	Respondent 40 - <i>"Implement strong data privacy and security measures to protect sensitive information, ensuring compliance with relevant data protection laws and regulations such as GDPR."</i>
	Acceptable Use	Respondent 45 - <i>"The content generated should be treated as a framework/ broad guideline as against treating it as the final product."</i> Respondent 52 - <i>"Students should learn about prompt engineering to better use GAI tools"</i> Respondent 62 - <i>"Use it wisely, especially in the process of idea generation would be good."</i> Respondent 64 - <i>"Some students may find AI great for supporting their learning (e.g. proofreading or for systematising information)."</i>
Regulations	Academic Regulations	Respondent 21 - <i>"Academic regulations need to be updated."</i>
	Legal Compliance	Respondent 40 - <i>"Ensure compliance with all relevant laws and regulations related to AI, including intellectual property, data protection, and non-discrimination laws."</i>
	Review	Respondent 40 - <i>"Regularly review and update AI governance policies to adapt to evolving technologies, societal norms, and emerging ethical considerations."</i>

Assessment and Academic Integrity	Adjusting Assessments	Respondent 64 - <i>"Institutions need to start/continue debating the use of AI but I still think we should have one last point of assessment where students don't have access to any AI tool... students should be able to demonstrate part of their learning path without the help of these tools."</i>
	Plagiarism Prevention	Respondent 56 - <i>"It can be integrated with Turnitin to identify plagiarism."</i>
	AI Referencing	Respondent 32 - <i>"Made mandatory for the students to acknowledge the use of GAI in their assignments and highlight the aspects/elements of the assignments supported by GAI"</i>
	Formative Assessment	Respondent 63 - <i>"GAI should be widely used for formative assessment rather than summative assessment."</i>
	Summative Assessment	Respondent 26 - <i>"For some assessments, maybe it will be necessary to conduct them as invigilated exams again."</i>
Training and Education	Skill Development	Respondent 17 - <i>"...it also produce the wrong outcome, if they do not know how to ask. So, they at least need to understand the topic of what they are learning, and being able to assess if or not the generated contents are correct or not...to do that, it requires the deep knowledge of the topic and it is why the education is still important for us. "</i>
	Teacher's Role	Respondent 6 - <i>"Teachers must admit the change and do not force the students to lie. Faculties have to be more open for the change."</i>
	Staff Training	Respondent 21 - <i>"Universities need to provide clear leadership and guidance to staff, including training in the use of GAI."</i>
	Awareness	Respondent 40 - <i>"Educate stakeholders, including researchers, faculty, students, and administrators, about AI principles, potential risks, and ethical considerations to promote responsible AI use."</i>
	Digital Competence	Respondent 41 - <i>"Higher educational institutions should provide a curated GPT service to enhance the digital competence of staff and students."</i>
	Preparing Students for the Future Workplace	Respondent 26 - <i>"In HE, we will have to teach the use of GAI tools as preparation for the workplace."</i>

**Table 9. Thematic Analysis Results**

The *Guidelines for Responsible Use of GAI Tools* theme in HE encompasses several principles, as explained by the codes. *Transparency* in AI systems and algorithms is crucial to ensure *accountability* and ethical use. Students should be encouraged to use GAI as a supplementary method, taking ownership of their work, while faculty should maintain control and awareness of GAI-generated content. *Bias mitigation* in AI algorithms and data is essential to uphold fairness and inclusivity. Furthermore, robust *data privacy and security* measures must be in place to protect sensitive information. There should be a standardised *HE Institution Guideline* across the sector on the use of GAI tools. *Acceptable uses* generally include proofreading and idea generation.

Regarding the *Regulations* theme, it is imperative that *academic regulations* are updated and adapted to accommodate the use of GAI tools within HE institutions. This requires a comprehensive *review* of existing policies, alongside the introduction of new regulations specifically addressing the unique challenges and opportunities presented by GAI technology. Ensuring *legal compliance* with all relevant laws and regulations related to AI is a critical aspect of this process. Such measures will ensure that institutions can effectively manage and mitigate any risks associated with the use of GAI tools, while also maximising the potential benefits for both students and teachers. The *Assessment and Academic Integrity* theme reveals the importance of *adjusting assessment* methods to seamlessly integrate GAI tools. This adaptation is crucial to navigate the evolving landscape of academic integrity in the age of AI, particularly concerning *AI referencing* and citation practices. Incorporating AI text detectors could serve as a valuable component of *plagiarism prevention* strategies. GAI tools can be effectively employed for *formative assessments*, providing students with continuous and constructive feedback from tutors or lecturers. However, to uphold the academic integrity of *summative assessments*, it is advisable to conduct these evaluations without the reliance on AI tools, such as through invigilated exams. This approach ensures a comprehensive and fair assessment of a student's knowledge and capabilities.

Under the theme of *Training and Education*, as part of *skills development*, it is crucial to equip students with the necessary skills to effectively utilise GAI tools. This includes the ability to validate information obtained from these tools, which is a critical thinking skill in itself. Currently, there is a gap in students' proficiency in employing GAI tools effectively. *Teachers* play an integral *role* in guiding students on the appropriate use of GAI tools. In tandem, HE institutions should also provide *staff training* on the utilisation of GAI tools, as well as offering guidance on how to incorporate these tools

into teaching and learning processes. Moreover, it is important to raise *awareness* and educate all stakeholders about the principles of AI, potential risks, and ethical considerations associated with its use. Enhancing the digital competence of both staff and students is imperative, as is teaching the use of GAI tools as a means of *preparing students for the future workplace*.

Echoing the suggestions proposed by respondent 28 and 32, in governing the use of GAI tools in HE institutions, it might be useful to model the *'holistic life cycle of AI use in academic settings'*. HE institutions should play a proactive role in adapting academic practices by providing guidance to students on how to select a reliable tool, how to use them, how to apply critical thinking when analysing the AI output, and how AI has helped them in achieving the learning outcomes.

## **5.0 Discussions and Conclusion**

### **5.1 Research Implications**

This research explores the teachers' perceptions of students' use of GAI tools in their summative assessments. The findings indicate that teachers are inclined to accept the use of GAI tools in summative assessment by students, as long as that students utilise such tools responsibly. The findings from sections 4.3 and 4.4 significantly contribute to informing HE policy regarding the responsible use of GAI tools. Through these results, the research unveils two pivotal concepts – responsibility and trust – both of which are integral in shaping future HE policies.

This research recognises that, amid the current paucity of guidelines on the responsible use of GAI tools in summative assessments, it is challenging to clearly define responsible usage and subsequently hold students accountable. These two elements are essential to establish responsibility: without clear duties and obligations, agents (the students) cannot be held accountable for their actions (McGrath & Whitty, 2018). Similarly, without clear consequences for irresponsible actions, accountability is unattainable (Dunn et al., 2021). Therefore, a responsible use of GAI tools in HE requires the development of clear guidelines and policies that set standards for fairness, transparency, and acceptable use, as well as the implementation of penalties for irresponsible usage. However, as responsibility is an evolving and dynamic concept (Rowe et al., 2023), such ramifications cannot simply be imposed on students, teachers, or administrators. Rather, they should be co-created by all stakeholders. The findings

reveal a wide range of perceptions among teachers, and similarly, diverse views among other involved parties could be stipulated. Therefore, responsible use of GAI tools in HE is not a straightforward concept, but rather depends on a multitude of perspectives that need to be negotiated as stakeholders arrive at shared notions of responsibility.

Secondly, this study contributes to the discourse of trust. Even with clear guidelines and policies for the responsible use of GAI tools, the issue of trust remains. Namely, teachers in this study implied that they did not trust students to use GAI tools responsibly. This raises an important question regarding the role of trust in using GAI tools responsibly. As trust is usually founded on non-codified notions, such as norms, habits, culture, and expectations (Faulkner, 2010), building trust differs from developing and implementing rules and policies. With the novelty of GAI tools and its potential significant impact on HE, a lack of trust is evident. This lack of trust is further exacerbated by the fact that GAI tools evolve more rapidly than the ramifications for responsible usage. This situation presents opportunities for research and practice to develop sufficiently elastic and dynamic methods to foster the necessary trust to support responsibility in GAI tools usage. Such attempts require further investigation into teachers' perceptions and, specifically, into the ways in which trustworthiness can be increased.

## **5.2 Limitations and Future Work**

One limitation of this study is the limited number of respondents who participated in the survey, despite the exploratory nature of the research. In future research, more participants will be recruited to ensure a comprehensive representation of the demographic variations in teachers' perceptions of the responsible use of GAI tools in HE. This includes aspects such as familiarity with and proficiency in using these tools. Moreover, given the rapid advancements in GAI technology, it is crucial to consider the potential risks associated with its usage. The survey could be expanded to address psychological risks, for instance. One potential concern is that the use of GAI tools could lead to a detachment from reality, which may be challenging to detect and mitigate.

Future research will delve deeper into the concepts of responsibility and trust, particularly from a sociotechnical perspective. This could inform the development of hypotheses that can be tested to enrich understanding of these complex and multifaceted concepts. By increasing the sample size, more analyses can be conducted, comparing

teachers' perceptions across a range of demographics such as age, gender, and teaching domain. Moreover, from a cross-cultural perspective, further research could investigate how cultural differences affect teachers' expectations regarding students' learning and behaviours. This could shed light on the ways in which cultural norms and values shape teachers' perceptions of responsibility and trust in the context of GAI tool usage. Such insights would be invaluable in developing culturally sensitive guidelines and policies for the responsible use of GAI tools in HE.

## References

- Amani, S., White, L., Balart, T., Arora, L., Shryock, D. K. J., Brumbelow, D. K., & Watson, D. K. L. (2023). Generative AI Perceptions: A Survey to Measure the Perceptions of Faculty, Staff, and Students on Generative AI Tools in Academia. *arXiv preprint arXiv:2304.14415*.
- Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62.
- Chan, C. K. Y., & Hu, W. (2023). Students' Voices on Generative AI: Perceptions, Benefits, and Challenges in Higher Education. *arXiv preprint arXiv:2305.00290*.
- Chan, C. K. Y., & Lee, K. K. (2023). The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and Millennial Generation teachers? *arXiv preprint arXiv:2305.02878*.
- Clarke, R. (2019). Principles and business processes for responsible AI. *Computer Law & Security Review*, 35(4), 410-422.
- Dunn, B., Jensen, M. L., & Ralston, R. (2021). Attribution of responsibility after failures within platform ecosystems. *Journal of Management Information Systems*, 38(2), 546-570.
- Faulkner, P. (2010). *Norms of trust*. OUP Oxford.
- Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5-14.
- Hodges, C., & Ocak, C. (2023). *Integrating Generative AI into Higher Education: Considerations*. EDUCAUSE Review. Retrieved 1 September 2023 from <https://er.educause.edu/articles/2023/8/integrating-generative-ai-into-higher-education-considerations>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399.
- Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., & Papin, K. (2023). Generative AI and Teachers' Perspectives on Its Implementation in Education. *Journal of Interactive Learning Research*, 34(2), 313-338.
- Kim, S., Park, J., & Lee, H. (2019). Automated essay scoring using a deep learning model. *Journal of Educational Technology Development and Exchange*, 2(1), 1-17.
- Lacey, M. M., & Smith, D. P. (2023). Teaching and assessment of the future today: higher education and AI. *Microbiology Australia*, 44(3), 124-126.
- McGrath, S. K., & Whitty, S. J. (2018). Accountability and responsibility defined. *International Journal of Managing Projects in Business*, 11(3), 687-707.
- Mezgár, I., & Váncza, J. (2022). From ethics to standards—A path via responsible AI to cyber-physical production systems. *Annual Reviews in Control*.
- Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. *Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning (February 11, 2023)*.
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT. *Education Sciences*, 13(9), 856.



- Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and 'the dark side' of AI. In (Vol. 31, pp. 257-268): Taylor & Francis.
- Namatherdhala, B., Mazher, N., & Sriram, G. K. (2022). A comprehensive overview of artificial intelligence trends in education. *International Research Journal of Modernization in Engineering Technology and Science*, 4(7).
- Noble, S. M., & Dubljević, V. (2022). Ethics of AI in organizations. In *Human-Centered Artificial Intelligence* (pp. 221-239). Elsevier.
- OpenAI. (2022). *Introducing ChatGPT*. Retrieved 01 September 2023 from <https://openai.com/blog/chatgpt>
- Rowe, F., Jeanneret Medina, M., Journé, B., Coetard, E., & Myers, M. D. (2023). Understanding responsibility under uncertainty: A critical and scoping review of autonomous driving systems. *Journal of Information Technology*, 02683962231207108.
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1).
- Schroeder, R. (2023). *Generative AI in College and Departmental Administration*. Inside Higher Ed (IHE). Retrieved 1 September 2023 from <https://www.insidehighered.com/opinion/blogs/online-trending-now/2023/08/02/generative-ai-college-and-departmental-administration>
- Smolansky, A., Cram, A., Radulescu, C., Zeivots, S., Huber, E., & Kizilcec, R. F. (2023). Educator and Student Perspectives on the Impact of Generative AI on Assessments in Higher Education. Proceedings of the Tenth ACM Conference on Learning@ Scale,
- UNESCO. (2023). *ChatGPT and Artificial Intelligence in Higher Education*. [https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide\\_EN\\_FINAL.pdf](https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf)
- Yeralan, S., & Lee, L. A. (2023). Generative AI: Challenges to higher education. *Sustainable Engineering and Innovation*, 5(2), 107-116.

# Understanding Consumers' Reactance of Technology-Enabled Personalization: Perspectives on Stress and Fatigue

## Abstract

*A growing number of retail companies are implementing highly advanced in-store technologies to better serve their consumers by improving the shopping experience and assisting them as they proceed along the customer journey. This study looks into the connections between techno-stress, disengagement, privacy fatigue, and technology-enabled personalization. The research hypothesis is put forth based on previous literature, and it is projected that 502 in-store shoppers actually make use of the personalization technology. Second, the data for this study was gathered by a questionnaire survey using online survey through a panel data, and LISREL would then be used to verify the model's causality. Lastly, based on the research findings, this research provides insights into strategic implications and potential research areas for management and businesses in this domain, offering valuable guidance for improving the customer experience in retail settings.*

**Keywords:** technology-enabled personalization, techno-stress, privacy fatigue, disengagement

## 1.0 Introduction

*“It was not technology per se that served as the conceptual underpinnings of changes occurring in the digital marketplace. What was really driving changes in the marketplace was not just technology, but how technology-enabled interactions between the key marketplace entities—consumers and firms—were being transformed by technology.” (Yadav & Pavlou, 2020, p. 132)*

The customer-interfacing retail technologies, referring to front-end<sup>1</sup> offerings that inform the retail interface with current or potential customers, have been greatly revolutionizing today's retailing environment and customer experience (Roggeveen & Sethuraman, 2020; Wang et al., 2024). Especially during the widespread store closures and the turbulent times for brick-and-mortar retailers, managers aim to provide a competitive retail environment in which the in-store experience can merge the benefits of both digital and physical dimensions of retailing, using technology advances to meet the needs of digital native consumers who demand expansive customer experiences and convenience

---

<sup>1</sup> In contrast to front-end technology, back-end technologies facilitate retail operations, such as agile supply chains, employee management automation, vendor management, and assortment planning and so on (Roggeveen & Sethuraman, 2020). This current research focus on front-end technology.

(Choi & Lowry, 2024; Kahn et al., 2018; Roy et al., 2017; Riegger et al., 2021). Developments of retailing technologies involve expanded applications of artificial intelligence (AI), machine learning, virtual reality (VR), big data, and mobile apps (Guha et al., 2021; Roggeveen & Sethuraman, 2020), such as in-store technology (e.g., robots, smart displays, or augmented reality) for convenience or social presence (Grewal et al., 2020b), and AI for personalized customer engagement (Kumar et al., 2019; Huang & Rust, 2020). Dekimpe et al. (2020) delineate further applications of retailing technology used in offline shopping.

In addition to the technology development itself, marketing scholars argue that, increasingly, technologies not only have an existence in their own right, but also an existence that in part relies on human interaction, and even more, in part on interactions with other technologies (Morgan-Thomas et al., 2020; Novak & Hoffman, 2019; Yadav & Pavlou, 2020). Therefore, instead of focusing on the technology itself, managers are suggested to focus on the technology-enabled engagement process (Heller et al., 2021). Accordingly, this study suggests that the in-store customer experience can be viewed as part of the customer journey (Grewal & Roggeveen, 2020; Lemon & Verhoef, 2016; Roggeveen et al., 2020; Terblanche & Kidd, 2021); thus, the customer-interfacing retail technologies serve as the crucial touch points to customers.

To be competitive, customer in-store shopping experience has been added the amenities of online retailing, such as *data-driven personalization* with the aids of AI smart technologies (e.g., text mining, speech recognition, facial recognition) (Hoffman & Novak, 2018; Huang & Rust, 2020; Riegger et al., 2021; Roggeveen & Sethuraman, 2020; Salo et al., 2022). For example, the customers can benefit from its ability to recognize the *personalized* patterns from data, such as (1) personalized products based on customer preferences, (2) personalized prices based on customer willingness to pay, (3) personalized frontline interactions, and (4) customized promotional content for personal communication (Huang & Rust, 2020). However, the data capture process for personalization can also trigger customers' concerns on their information security and privacy risks (Corcoran et al., 2024; Grewal et al., 2020a; Lyu et al., 2023); for example, even when consumers intentionally share information, they are not aware of how this information is aggregated over time and across contexts (Puntoni et al., 2020). The *personalization-privacy paradox* (Aguirre et al., 2015; Cloarec et al., 2022; Grewal et al., 2017; Xu et al., 2011) reflects this situation, and firms are suggested to carefully use the personal information in a strategic manner to balance such paradox (Tong et al., 2019).

When most research on new technologies focus on the drivers of adoption, encouraging customers engage in the new retailing environment, relatively few studies focus on the disengagement in the in-store technology adaption. Until recently, scholars begin to study the barriers to adopt new retailing system (e.g., autonomous shopping systems) (de Bellis & Johar, 2020), and the consumers' psychological costs in their interactions with AI in retails (Mariani et al., 2022; Puntoni et al., 2020) or with smart retail technology (Riegger et al., 2021). However, those studies mostly focus on the drivers and outcomes of adopting new technologies, instead of why consumers adopt (or why they

do not adopt). To fulfill the research gap, the current study focuses on the reasons of consumers' psychological reactance, such as techno-stress, privacy fatigue, and disengagement toward technology-enabled personalization (TEP). Specifically, when consumers are aware of the data capture for personalization during the in-store shopping journey, and they may choose not to take advantage of the benefits of their personal data, resulted from the *techno-stress* which leads to *privacy fatigue* and *disengagement* in the in-store technologies. Based on the stressors-strain-outcome framework, this study proposes a research model to delineate the potential psychological mechanisms of consumers' reactions towards in-store personalization technologies.

In summary, by extending recent studies on technology-enabled personalization (Riegger et al., 2021; Riegger et al., 2022), technology-enabled interactions (Yadav & Pavlou, 2020), and technology-enabled engagement process (Heller et al., 2021), this timely research focuses on the importance of personalization technology applied in customers' in-store shopping journey, and provides a better understanding of the drivers of consumers' reactance towards new retailing technologies.

## **2.0 Theoretical Background and Hypotheses Development**

### **2.1. Technology-Enabled Personalization**

Technology in this study is considered to help integrate historic and real-time data without interference from store employees, which may combine cognitive technologies that collect, analyze, and react to customer data, as well as emotional technologies that enrich interactions and build customer relationships (Riegger et al., 2021). With the help of technology, retailers can integrate data-driven online personalization, such as recommendations from an algorithm (Aguirre et al., 2015), and face-to-face offline personalization, such as personal recommendations from store employees (Gwinner et al., 2005). Accordingly, in this study, *technology-enabled personalization* refers to the integration of offline and online personalization dimensions, in order to create an interactive, context-specific retailing environment and provide individual customers with relevant, context-specific information, based on the historic and real-time data in combination (Riegger et al., 2021; Riegger et al., 2022).

Such technology-enabled personalization in the in-store shopping journey may include recommendation services, or exclusive promotional offers, and such personalization is suggested to forecast user demand, enhance user engagement, and increase impulsive purchasing (Silva et al., 2023; Zhang & Sundar, 2019). However, it is worth-noting that marketers often use the term "personalization" and "customization" interchangeably, because both of them are aiming to achieve individualized information or service to customers. In this study, personalization refers to a firm-controlled process using previously collected customer data (such as personalized recommendations based on previous buying items), while customization is a customer-decided process when customers proactively specify the design and delivery of the offering (de Bellis et al., 2019; Kumar et al., 2019).

Therefore, during customers' in-store shopping journey, personalization may lead to a higher level of security concerns than customization, due to the uncertainties of how the firms collect and use their personal data for personalization. In other words, in order to achieve personalization, the firms will need to capture consumers' personal data (including behavioral data); so when consumers worry about how their data are collected and used, personalization may become a double-edged sword, both enhancing and diminishing consumer engagement with the firm (Tong et al., 2020). This study explores the unfavorable outcomes from consumers' perspectives, mainly based on the stressor-stress-strain model explained below.

## 2.2. Techno-Stress

Techno-stress is the stress that individuals experience from their inability to cope with the demands of using Information Technologies (Ayyagari et al., 2011). In essence, stress refers to a condition or event in the situation, the person's reaction to the situation, or the relationship between the person and situation (Hobfoll, 1989; Bliese et al., 2017), although it may have multiple meanings in previous psychological, organizational, marketing, and information systems studies (e.g. Bolino et al., 2015; Masood et al., 2021). Stress can be viewed as a combination of *stressors* (environmental events that act on an individual), and *strain* refers to the individual's response to those stressors (Griffin & Clarke, 2011). Thus, *stress* is a process whereby environmental stimuli trigger cognitive and physiological changes that culminate in *strain* outcomes, like exhaustion, depression (Gonzalez-Mulé & Cockburn, 2017; Gonzalez-Mulé et al., 2021). In the most general case, stressors increase strain (Tarafdar et al., 2010).

More recently, technology in general and Information and Communication Technologies (ICT) in particular have emerged as conditions for the cause of stress, and *techno-stress* represents an emerging area of scholarly investigation in the field of Information Systems (e.g., Ayyagari et al., 2011; Ragu-Nathan et al., 2008). Studies on *technostress* are mainly investigating how and why the use of IS causes various demands on the individual (Salo et al., 2022; Tarafdar et al., 2019). For example, Ragu-Nathan et al. (2008) suggest that *technostress creators* (factors that create stress from the use of ICTs) can be multi-dimensional, including techno-overload, techno-complexity, techno-insecurity, techno-uncertainty and techno-invasion. Tarafdar et al. (2010) describe the techno-stress in using information systems in the organizational contexts, including application multitasking, constant connectivity, information overload, frequent system upgrades and consequent uncertainty, continual relearning and consequent job-related insecurities, and technical problems associated with the organizational use of ICT. Pirkkalainen et al. (2019) investigate the proactive and reactive coping behaviors of how individuals address, manage, and deal with stressful situations caused by using information technologies. The authors suggest distress venting (e.g., expressing emotions during stressful situations) and distancing from IT (e.g., separating oneself from stressful situations) as two main reactive coping behaviors: IT users might vent their negative emotions regarding IT and distance themselves from the situation when they experience IT use-related stressors (Pirkkalainen et al., 2019).

In this study, techno-stress can also be viewed from the marketing and consumer behavior perspectives. Based on resources theory, stress results from the relationship between the person and an environment the individual considers taxing, in excess of his or her resources, or dangerous to his or her well-being (Chan & Wan, 2012; Schwarzer & Taubert, 2002). Stress can be inherent within the concept of consumer desire which is traditionally described as a motivation to reduce a felt discrepancy between a desired and current state (Dorsch et al., 2017). When consumers experience technology-enabled personalization, it is very likely the discrepancy between their desired and current states increases, so the level of techno-stress is increased and stressed consumers are motivated to resolve the discrepancy. In particular, under the context of technology-enabled personalization, consumers may receive large amount of information, even though personalized, such as the in-store promotion, recommended products, which lead to perceived information overload (Chen et al., 2019; Ragu-Nathan et al., 2008; Westermann et al., 2015). Consumers may also face stressful demands of using personalized technologies that create a threat to consumers who “feel incapable of completing impending tasks with the resources at hand” (Chan & Wan, 2012, p. 120). More importantly, when consumers receive highly-personalized information, they may also concern about their privacy which leads to perceived intrusiveness (Chen et al., 2019; Tarafdar et al., 2010). Therefore, this study suggests the following hypotheses:

**H1: Technology-enabled personalization is positively related to techno-stress.**

**H2: Techno-stress is positively related to disengagement.**

### **2.3 Privacy Fatigue**

Followed by the stress-coping theory, this study focus on consumers’ coping behaviors toward the threat of technostress in the context of technology-enabled personalization. Furthermore, comparing with proactive coping behavior (e.g., dealing with ongoing stressful situations through personal growth and resource accumulation) (Aspinwall & Taylor, 1997; Folkman & Moskowitz, 2004), consumers are more likely to present reactive coping behaviors, becoming passive and disengaged in adopting technologies. One of the main passive reactions is privacy fatigue, explained as follows.

Fatigue is the subjective feeling of weariness or exhaustion, either as a lasting condition that remains stable over weeks or months or as a state that may fluctuate with contexts (French & Allen, 2020). Privacy fatigue reflects a sense of weariness toward privacy issues (Choi et al., 2018), in which individuals feel pointless to attempt to protect their personal data (Keith et al., 2014; Lee et al., 2016; Zhang et al., 2016). It is suggested that privacy fatigue not only arises from the complexity of privacy assurance systems (Schermer et al., 2014) which make consumers feel stressful, but also frequent data breaches which make people feel a loss of control over personal information (Sen & Borle, 2015).

In this study, when individuals under stressful situations (e.g., figuring out how and why their

personal information is used for personalization), it has become difficult and burdensome to maintain and control their privacy. Based on the stressor-strain-outcome framework, when facing the personalization-privacy paradox, consumers may also experience a manifestation of a multi-dimensional psychological strain, characterized as cynicism, exhaustion, and reduced efficacy (Choi et al., 2018; Schaufeli et al., 1996). *Cynicism* is developed from a failure to produce the expected outcomes, usually accompanied by other negative emotions such as frustration, hopelessness, and disillusionment (Andersson, 1996; Johnson & O’Leary-Kelly, 2003), and suggested to be a core component of fatigue. *Emotional exhaustion* is defined as the draining of emotional resources, reflecting the stress generated by excessive emotional demands (Maslach et al., 2001). Finally, cynicism and emotional exhaustion caused under a stressful situation will ultimately decrease one’s sense of effectiveness, thus reducing efficacy in fulfilling a requirement (Schaufeli et al., 1996). It reflects feelings of *reduced efficacy* in one of the dimensions. Thus, this study suggests the following hypothesis:

**H3: Technology-enabled personalization is positively related to privacy fatigue, in terms of cynicism, emotional exhaustion, and reduced efficacy toward their privacy management.**

## 2.4 Disengagement

Disengagement is defined as “reducing one’s effort to deal with the stressor, even giving up the attempt to attain goals with which the stressor is interfering” (Carver et al., 1989, p. 269). As fatigue is observed as a reluctance for further effort, psychologists (e.g., Hopstaken et al., 2015) suggest that fatigue is related to a general disengagement and low vigor in contrast to the possibility of exploiting the benefits of a certain task (Boksem et al., 2006). In addition, fatigued people are more likely to exhibit disengagement due to their inability to handle decision-making (Levav et al., 2010). Fatigue can also lead to discontinuous usage for existing users of ICTs, as Zhang et al. (2016) empirically investigate and verify the relationship between social network fatigue and discontinuous use intentions.

In this study, disengagement refers to the extent to which consumers reduce their effort to deal with techno-stress, resulted from adopting personalization technologies, because they are tired of responding privacy threats. Moreover, disengagement in this study also represents consumers’ withdrawal of diverse coping behaviors due to privacy threats (Choi et al., 2018). Individuals, under a controllable situation, generally believe that further steps can be taken to achieve desirable outcomes (Weinstein, 1980), but they exhibit reluctance to put further efforts when feeling an irreversible loss of control (Hopstaken et al., 2015). Such consumers’ feelings of losing control, in terms of protecting their personal information and avoiding exposure to personalization content that may makes them stressful (Hinds et al., 2020), may make them disengage in personalization technologies. Therefore, individuals experiencing privacy fatigue due to information overload, feelings of losing control, and unmet expectations are likely to exhibit disengagement behavior in which they are no longer

motivated and willing to invest more efforts on adopting personalization technologies. Thus, this study suggests the following hypothesis:

**H4: Privacy fatigue is positively related to disengagement.**

Personalized offerings (including recommendation services, or exclusive promotional offers) usually aim to forecast user demand, enhance user engagement, and increase impulsive purchasing (Zhang & Sundar, 2019). From users' perspective, mobile personalization via in-store mobile phone use can provide more convenient and relevant services, and simultaneously reduce users' cognitive load in decision-making as well as improve overall user experience (Zhang & Sundar, 2019). In addition, personalization, shown to work in both online and offline environments, can be viewed as a process that interlinks customers and marketers and solidifies the relationship between them, which will lead to emotional bonding progress to a state of engagement, and finally influence customer engagement behaviors (Kumar et al., 2019). Specifically, shoppers need to place demands on their mental resources not only for shopping decision making (Inman et al., 2009) but also for processing the personalized promotion-related information via mobile phone. Both tasks require the use of visual and verbal processing resources (Sciandra et al., 2019), which makes shoppers engage in the decision-making process better. Thus, this study suggests the following hypothesis:

**H5: Technology-enabled personalization is negatively related to disengagement.**

### **3.0 Research Method**

#### **3.1. Research Contexts**

More and more retailing industries adopt truly advanced in-store technologies to serve their customers by facilitating the shopping process and helping customers move through each stage of their customer journey (Grewal et al., 2020b). Retailers aim to provide in-store shoppers with personalized offerings to improve a customer's experience and engagement (Kumar et al., 2017), sensing and monitoring consumers' in-store shopping behaviors through iBeacons, mobile POS, Near Field Communications (NFC), and the Internet of Things (IoT) (Inman & Nikolova, 2017). In our study, we focused on retail stores employing one or more of the following technologies: Product Experience Walls, Interactive Fitting Rooms, Smart and Social Mirrors, Smart Shelves, (Touch) Screens, Service Robots, and 360-Degree Cameras for Virtual or Augmented Reality (refer to Riegger et al., 2021 for further details). These technologies are increasingly being integrated into the apparel and cosmetics industries in Taiwan. Notably, most of our respondents reported their usage experiences in well-established Textile & Apparel stores (such as NET and Zara) and Cosmetics & Personal Care Products stores (like Watsons and Cosmed). In these settings, the



deployment of novel technologies often involves the collection of customer information for personalization purposes. This process may occur without explicit notification or prior consent from consumers (for example, without the need for downloading an app or registering as a member). This aspect raises important considerations regarding consumer privacy and the ethical use of personal data in technology-enhanced retail environments.

### **3.2. Research Design and Sample Collection Process**

To validate the research model, the data were collected from the in-store shoppers who actually use the personalization technologies. However, as this study looks into how individuals perceive techno-stress due to technology-enabled personalization, and how they deal with ongoing stressful situations, it is of high importance to retain the anonymity of the respondents (c.f., Pirkkalainen et al., 2019). Therefore, this study uses online survey through a panel data owned by a marketing survey company. It is a valid and relevant approach to ensure two criteria: The panel data included the companies which use personalization technologies in the in-store retailing environment. With the help of the marketing survey company, the questionnaires were sent to those who experienced personalization technologies.

To increase the response rate, an incentive of NTD 50 voucher were provided for each completed questionnaire. Consequently, 502 respondents joined the survey. After viewing their responses records, 398 usable responses were included for testing the model. Among them, 207 were male (52%) and 191 were female (48%), and most of them (96%, n=382) are having university degrees (or above). Furthermore, the measurement items used for each construct were mainly adapted from validated scales obtained from the literature, but with minor adjustments to fit the scenario in the study. All the variables were measured by participant responses to questions on a Likert-type scale, ranging from “strongly disagree” to “strongly agree.” In addition, the entire survey was translated from English into Chinese and then back-translated into English by two independent bilingual researchers to ensure equivalency of meaning (Brislin, 1980).

## **4.0 Results**

This study used a two-stage approach to test the models (Anderson & Gerbing, 1988). Firstly, this study conducted a confirmatory factor analysis to assess the quality of the measurement model. Then this study tested the theoretical model using the partial least squares (PLS) method, using SmartPLS 3.0, which is suitable for use when theory development is the focus (as opposed to LISREL, which would be preferred for confirmatory tests of the fit of a theoretical model to observed data; Chin, 1998).

### **4.1. Measurement Model Validation**

The adequacy of the measurement model was assessed by evaluating the reliability of the individual

items, the internal consistency between items designed to measure the same construct, and the discriminant validity between constructs. To determine the reliability of each item, the significance and magnitude of the item's loadings were examined. In this study, composite reliability ranged from 0.69 to 0.87, and the average variance extracted ranged from 0.62 to 0.76; all values were above the recommended cut-off levels of 0.60 and 0.50, respectively (Bagozzi & Yi, 2012). Therefore, the constructs exhibited good internal consistency. For discriminant validity, the diagonal elements representing the square roots of the average variance extracted for each construct were greater than the correlations of each specific construct with any of the other constructs in the model (i.e., the off-diagonal elements; Chin, 1998), ranged from 0.79 to 0.87, and were all above 0.50 (Fornell & Larcker, 1981). Next, constructed 95% confidence intervals were conducted for each correlation coefficient to determine whether the correlations among the latent variables were significantly less than 1 (Bagozzi & Yi, 2012). None of the confidence intervals included the value of 1. Thus, there was evidence of discriminant validity.

## 4.2. Structural Model Evaluation

The  $R^2$  value of disengagement is 0.21, demonstrating that the model explains a substantial amount of the variance in the outcome variables. The PLS results are listed as follows. Technology-enabled personalization is negatively and significantly related to techno-stress, with standardized  $\gamma$  coefficient of -0.17 ( $p < 0.01$ ). Therefore, H1 is not supported. For H2, as predicted, techno-stress is positively and significantly related to disengagement ( $\beta = 0.14$ ,  $p < 0.05$ ), thus supporting H2. Technology-enabled personalization is negatively and significantly related to privacy fatigue, with standardized  $\gamma$  coefficient of -0.15 ( $p < 0.01$ ), thus not supporting H3. Privacy fatigue is positively related to disengagement, although not significantly ( $\beta = 0.05$ ,  $p > 0.05$ ), thus H4 is not supported. Finally, technology-enabled personalization is negatively and significantly related to disengagement, with standardized  $\gamma$  coefficient of -0.40 ( $p < 0.001$ ). Regarding the control variables, gender and educational level were both not significantly related to engagement.

## 5.0 Discussions

Although this study is different from recent studies on users' behavior toward a specific technology, such as IoT (e.g., Novak & Hoffman, 2019), AR (e.g., Heller et al., 2021), AI (e.g., Puntoni et al., 2020) and robot (e.g., Huang & Rust, 2020), this study answers the call that marketing researchers will need to establish both the fundamental effects of various new technologies, from a multidisciplinary perspective, as well as customer insights into the underlying processes that explain the behavioral consequences of facing the in-store personalization technologies (Grewal et al., 2020a). Accordingly, this study integrates a series of studies in the fields of psychology, information systems (IS), and marketing, such as the concept of techno-stress (Ayyagari et al., 2011), privacy fatigue (e.g., Choi et al., 2018) to explore some of the implications related to technology-enabled personalization in physical shopping environments. The research makes the following contributions.

First, although personalization is suggested to be a double-edged sword, eliciting both favorable and unfavorable consumer outcomes (Riegger et al., 2021), such as well-known personalization-privacy paradox. However, surprisingly, those negative concerns of technology-enabled personalization did not appear in this study, and this study finds technology-enabled personalization negatively relates to disengagement. It is possible that previous studies on the negative consequences of personalization (i.e., techno-stress, privacy fatigue) are mainly focusing on adapting novel technologies, and using information systems in the organizational contexts (e.g., Ayyagari et al., 2011; Ragu-Nathan et al., 2008), using social network service (e.g., Zhang et al., 2016; Masood et al., 2021), instead of in the in-store shopping environment. The current study complements Riegger et al.'s (2021) qualitative research which provides a comprehensive overview of technology-enabled personalization, but lack of empirical examination. Thus, this study actually advances the qualitative insights gained from previous research on privacy fatigue (e.g., Hinds et al., 2020) and technology-enabled personalization (Riegger et al., 2021).

Secondly, based on the stressor-strain-outcome framework, this study provides an integrated research framework to examine the ongoing process of coping the techno-stress resulted from technology-enabled personalization. Based on the finding that techno-stress significantly leads to disengagement, this study suggests that customer disengagement should be viewed as a process, in the form of misfit of users' resources (i.e., time, skills, and knowledge) to handle the personalization technologies. Accordingly, this study contributes to literature on techno-stress coping (e.g., Chen et al., 2019; Pirkkalainen et al., 2019; Tarafdar et al., 2019) by proposing a process that propagating consumers' psychological reactance to disengagement. This process view helps managers understand both the dimensions of disengagement and the importance of stress and fatigue that trigger disengagement within the context of adopting technology-enabled personalization (Heller et al., 2021).

For practitioners, it is an obvious necessity to understand, from consumer perspectives, the benefits and risks of technology used in marketing contexts, and the ways of how consumers interact with the technologies. So far, little is known about how engagement with retailing technologies emerges (or diminishes) and whether the process results in expected marketing-relevant outcomes (Heller et al., 2021). This timely study provides insights into in-store shopper behaviors with technology-enabled personalization experience, which can help practitioners design more engaging marketing strategies and avoid negative impact of novel personalization technologies which may lead to disengagement.

## **6.0 Conclusion**

This study significantly broadens our understanding of new retailing technology adoption, while advancing research into consumer responses to technology-enabled personalization and the personalization-privacy paradox. Moreover, it addresses a notable gap in technology discontinuance research, particularly within the realm of retailing technology for personalization, by delving into the adverse aspects of technology use that can act as stressors leading to discontinuance behaviors (Chen et al., 2019).

From a practical standpoint, this research sheds light on the dual aspects of challenges and opportunities presented by technological personalization in contemporary retail, especially in the face of consumer skepticism towards AI-driven and data-driven personalization strategies. Importantly, the study offers valuable insights into striking a balance between the provision of personalized services and the management of consumer privacy concerns, suggesting a more nuanced approach to this pivotal issue.

Furthermore, this research serves as an essential guide for retailers, outlining strategies to effectively leverage AI and data analytics to positively impact customer purchasing behavior. This is particularly pertinent in the context of smart retailing and the integration of AI in the retail marketing mix, showcasing both preliminary findings and ongoing research that sets the stage for future explorations. These future studies could further investigate the underlying reasons for consumer hesitancy in embracing AI-generated outcomes, such as personalized offers.

In conclusion, this study acts as a foundational piece of research, seeking to develop, evaluate, and mitigate the negative implications of new personalized mobile technologies for both retailers and consumers. It stands as a pivotal contribution to the field, paving the way for further inquiry into the complex dynamics of retail technology, consumer behavior, and AI integration.

## References

- Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., & Wetzels, M. (2015). Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. *Journal of Retailing*, *91*(1), 34-49.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411-423.
- Andersson, L. M. (1996). Employee cynicism: An examination using a contract violation framework. *Human Relations*, *49*(11), 1395-1418.
- Aspinwall, L. G., & Taylor, S. E. (1997). A stitch in time: Self-regulation and proactive coping. *Psychol Bull*, *121*(3), 417-436.
- Ayyagari, R., Grover, V., & Purvis, R. (2011). Technostress: Technological antecedents and implications. *MIS Quarterly*, *35*(4), 831-858.
- Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science*, *40*(1), 8-34.
- Bliese, P., Edwards, J., & Sonnentag, S. (2017). Stress and well-being at work: A century of empirical trends reflecting theoretical and societal influences. *Journal of Applied Psychology*, *102*(3), 389-402.
- Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2006). Mental fatigue, motivation and action monitoring. *Biol Psychol*, *72*(2), 123-132.
- Bolino, M., Hsiung, H.-H., Harvey, J., & LePine, J. (2014). "Well, I'm tired of tryin'!" Organizational citizenship behavior and citizenship fatigue. *The Journal of Applied Psychology*, *100*(1), 56-74.
- Brislin, R. W. (1980). Translation and content analysis of oral and written material. In Triandis, H. C., & Berry, J. W. (eds.), *Handbook of Cross-Cultural Psychology*, *1*, 389-444.
- Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing coping strategies: A theoretically based approach. *J Pers Soc Psychol*, *56*(2), 267-283.
- Chan, K. W., & Wan, E. W. (2012). How can stressed employees deliver better customer service? The underlying self-regulation depletion mechanism. *Journal of Marketing*, *76*(1), 119-137.
- Chen, J. V., Tran, A., & Nguyen, T. (2019). Understanding the discontinuance behavior of mobile shoppers as a consequence of technostress: An application of the stress-coping theory. *Computers in Human Behavior*, *95*, 83-93.
- Chin, W. W. (1998). Commentary: Issues and opinion on structural equation modeling. *MIS Quarterly*, *22*(1), 7-16.
- Choi, D. D., & Lowry, P. B. (2024). Balancing the commitment to the common good and the protection of personal privacy: Consumer adoption of sustainable, smart connected cars. *Information & Management*, *61*(1), 103876.
- Choi, H., Park, J., & Jung, Y. (2018). The role of privacy fatigue in online privacy behavior. *Computers in Human Behavior*, *81*, 42-51.

- Cloarec, J., Meyer-Waarden, L., & Munzel, A. (2022). The personalization–privacy paradox at the nexus of social exchange and construal level theories. *Psychology and Marketing, 39*(3), 647-661.
- Corcoran, E., Clement, A., & Gabrielli, J. (2024). A measurement model of online privacy cognitions in a sample of US adolescents. *Behaviour and Information Technology*, Published online (22 Jan 2024), <https://doi.org/10.1080/0144929X.2024.2305286>.
- de Bellis, E., & Johar, G. V. (2020). Autonomous shopping systems: Identifying and overcoming barriers to consumer adoption. *Journal of Retailing, 96*(1), 74-87.
- de Bellis, E., Hildebrand, C., Ito, K., Herrmann, A., & Schmitt, B. (2019). Personalizing the customization experience: A matching theory of mass customization interfaces and cultural information processing. *Journal of Marketing Research, 56*(6), 1050-1065.
- Dekimpe, M. G., Geyskens, I., & Gielens, K. (2020). Using technology to bring online convenience to offline shopping. *Marketing Letters, 31*(1), 25-29.
- Dorsch, M. J., Törnblom, K. Y., & Kazemi, A. (2017). A review of resource theories and their implications for understanding consumer behavior. *Journal of the Association for Consumer Research, 2*(1), 5-25.
- Folkman, S., & Moskowitz, J. T. (2004). Coping: Pitfalls and promise. *Annual Review of Psychology, 55*, 745-774.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.
- French, K. A., & Allen, T. D. (2020). Episodic work-family conflict and strain: A dynamic perspective. *Journal of Applied Psychology, 105*(8), 863-888.
- Gonzalez-Mulé, E., & Cockburn, B. (2017). Worked to death: The relationships of job demands and job control with mortality. *Personnel Psychology, 70*(1), 73-112.
- Gonzalez-Mulé, E., Kim, M., & Ryu, J. W. (2021). A meta-analytic test of multiplicative and additive models of job demands, resources, and stress. *Journal of Applied Psychology, 106*(9), 1391-1411.
- Grewal, D., & Roggeveen, A. L. (2020). Understanding retail experiences and customer journey management. *Journal of Retailing, 96*(1), 3-8.
- Grewal, D., Hulland, J., Kopalle, P. K., & Karahanna, E. (2020a). The future of technology and marketing: a multidisciplinary perspective. *Journal of the Academy of Marketing Science, 48*(1), 1-8.
- Grewal, D., Noble, S. M., Roggeveen, A. L., & Nordfalt, J. (2020b). The future of in-store technology. *Journal of the Academy of Marketing Science, 48*(1), 96-113.
- Grewal, D., Roggeveen, A. L., & Nordfalt, J. (2017). The future of retailing. *Journal of Retailing, 93*(1), 1-6.
- Griffin, M. A., & Clarke, S. (2011). Stress and well-being at work. In *APA handbook of industrial and organizational psychology, Vol 3: Maintaining, expanding, and contracting the organization*. (pp. 359-397). American Psychological Association.
- Guha, A., Grewal, D., Kopalle, P. K., Haenlein, M., Schneider, M. J., Jung, H., Moustafa, R.,

- Hegde, D. R., & Hawkins, G. (2021). How artificial intelligence will affect the future of retailing. *Journal of Retailing*, 97(1), 28-41.
- Gwinner, K. P., Bitner, M. J., Brown, S. W., & Kumar, A. (2005). Service Customization Through Employee Adaptiveness. *Journal of Service Research*, 8(2), 131-148.
- Heller, J., Chylinski, M., de Ruyter, K., Keeling, D. I., Hilken, T., & Mahr, D. (2021). Tangible service automation: Decomposing the technology-enabled engagement process (TEEP) for augmented reality. *Journal of Service Research*, 24(1), 84-103.
- Hinds, J., Williams, E. J., & Joinson, A. N. (2020). "It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human-Computer Studies*, 143, Article 102498.
- Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, 44(3), 513-524.
- Hoffman, D. L., & Novak, T. P. (2018). Consumer and object experience in the internet of things: An assemblage theory approach. *Journal of Consumer Research*, 44(6), 1178-1204.
- Hopstaken, J. F., Van Der Linden, D., Bakker, A. B., & Kompier, M. A. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement. *Psychophysiology*, 52(3), 305-315.
- Huang, M. H., & Rust, R. T. (2020). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, 48, 1-8.
- Inman, J. J., Winer, R. S., & Ferraro, R. (2009). The interplay among category characteristics, customer characteristics, and customer activities on in-store decision making. *Journal of Marketing*, 73(5), 19-29.
- Inman, J. J., & Nikolova, H. (2017). Shopper-facing retail technology: A retailer adoption decision framework incorporating shopper attitudes and privacy concerns. *Journal of Retailing*, 93(1), 7-28.
- Johnson, J. L., & O'Leary-Kelly, A. M. (2003). The effects of psychological contract breach and organizational cynicism: not all social exchange violations are created equal. *Journal of Organizational Behavior*, 24(5), 627-647.
- Kahn, B. E., Inman, J. J., & Verhoef, P. C. (2018). Introduction to special issue: Consumer response to the evolving retailing landscape. *Journal of the Association for Consumer Research*, 3(3), 255-259.
- Keith, M. J., Maynes, C., Lowry, P. B., & Babb, J. (2014). Privacy fatigue: The effect of privacy control complexity on consumer electronic information disclosure. In *International Conference on Information Systems (ICIS 2014), Auckland, New Zealand, December* (pp. 14-17).
- Kumar, V., Anand, A., & Song, H. (2017). Future of retailer profitability: An organizing framework. *Journal of Retailing*, 93(1), 96-119.
- Kumar, V., Rajan, B., Venkatesan, R., & Lecinski, J. (2019). Understanding the role of artificial intelligence in personalized engagement marketing. *California Management Review*, 61(4), 135-155.

- Lee, A. R., Son, S.-M., & Kim, K. K. (2016). Information and communication technology overload and social networking service fatigue: A stress perspective. *Computers in Human Behavior*, 55, 51-61.
- Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, 80(6), 69-96.
- Levav, J., Heitmann, M., Herrmann, A., & Iyengar, Sheena S. (2010). Order in product customization decisions: Evidence from field experiments. *Journal of Political Economy*, 118(2), 274-299.
- Lyu, T., Guo, Y., & Chen, H. (2023). Understanding the privacy protection disengagement behaviour of contactless digital service users: the roles of privacy fatigue and privacy literacy. *Behaviour and Information Technology*, 1-17.
- Mariani, M. M., Perez-Vega, R., & Wirtz, J. (2022). AI in marketing, consumer research and psychology: A systematic literature review and research agenda. *Psychology and Marketing*, 39(4), 755-776.
- Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, 52(1), 397-422.
- Masood, A., Feng, Y., Rasheed, M. I., Ali, A., & Gong, M. (2021). Smartphone-based social networking sites and intention to quit: self-regulatory perspective. *Behaviour and Information Technology*, 40(11), 1055-1071.
- Morgan-Thomas, A., Dessart, L., & Veloutsou, C. (2020). Digital ecosystem and consumer engagement: A socio-technical perspective. *Journal of Business Research*, 121, 713-723.
- Novak, T. P., & Hoffman, D. L. (2019). Relationship journeys in the internet of things: a new framework for understanding interactions between consumers and smart objects. *Journal of the Academy of Marketing Science*, 47(2), 216-237.
- Pirkkalainen, H., Salo, M., Tarafdar, M., & Makkonen, M. (2019). Deliberate or instinctive? Proactive and reactive coping for technostress. *Journal of Management Information Systems*, 36(4), 1179-1212.
- Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2020). Consumers and artificial intelligence: An experiential perspective. *Journal of Marketing*, 85(1), 131-151.
- Ragu-Nathan, T. S., Tarafdar, M., Ragu-Nathan, B. S., & Tu, Q. (2008). The consequences of technostress for end users in organizations: Conceptual development and empirical validation. *Information Systems Research*, 19(4), 397-521.
- Riegger, A.-S., Klein, J. F., Merfeld, K., & Henkel, S. (2021). Technology-enabled personalization in retail stores: Understanding drivers and barriers. *Journal of Business Research*, 123, 140-155.
- Riegger, A. S., Merfeld, K., Klein, J. F., & Henkel, S. (2022). Technology-enabled personalization: Impact of smart technology choice on consumer shopping behavior. *Technological Forecasting and Social Change*, 181, 121752.
- Roggeveen, A. L., & Sethuraman, R. (2020). Customer-interfacing retail technologies in 2020 & beyond: An integrative framework and research directions. *Journal of Retailing*, 96(3), 299-



- Roggeveen, A. L., Grewal, D., & Schweiger, E. B. (2020). The DAST framework for retail atmospherics: The impact of in- and out-of-store retail journey touchpoints on the customer experience. *Journal of Retailing*, *96*(1), 128-137.
- Roy, S. K., Balaji, M. S., Sadeque, S., Nguyen, B., & Melewar, T. C. (2017). Constituents and consequences of smart customer experience in retailing. *Technological Forecasting and Social Change*, *124*, 257-270.
- Salo, M., Pirkkalainen, H., Chua, C. E. H., & Koskelainen, T. (2022). Formation and Mitigation of Technostress in the Personal Use of IT. *MIS Quarterly*, *46*(2), 1073-1108.
- Schaufeli, W. B., Dierendonck, D. V., & Gorp, K. V. (1996). Burnout and reciprocity: Towards a dual-level social exchange model. *Work & Stress*, *10*(3), 225-237.
- Schermer, B. W., Custers, B., & van der Hof, S. (2014). The crisis of consent: How stronger legal protection may lead to weaker consent in data protection. *Ethics and Information Technology*, *16*(2), 171-182.
- Schwarzer, R., & Taubert, S. (2002). Tenacious goal pursuits and striving toward personal growth: Proactive coping. In *Beyond coping: Meeting goals, visions, and challenges*. (p. 19-35). New York, NY, US: Oxford University Press.
- Sciandra, M. R., Inman, J. J., & Stephen, A. T. (2019). Smart phones, bad calls? The influence of consumer mobile phone use, distraction, and phone dependence on adherence to shopping plans. *Journal of the Academy of Marketing Science*, *47*(4), 574-594.
- Sen, R., & Borle, S. (2015). Estimating the contextual risk of data breach: An empirical approach. *Journal of Management Information Systems*, *32*(2), 314-341.
- Silva, S. C., Silva, F. P., & Dias, J. C. (2023). Exploring omnichannel strategies: a path to improve customer experiences. *International Journal of Retail and Distribution Management*, *52*(1), 62-88.
- Tarafdar, M., Cooper, C. L., & Stich, J.-F. (2019). The technostress trifecta - techno eustress, techno distress and design: Theoretical directions and an agenda for research. *Information Systems Journal*, *29*(1), 6-42.
- Tarafdar, M., Tu, Q., & Ragu-Nathan, T. S. (2010). Impact of technostress on end-user satisfaction and performance. *Journal of Management Information Systems*, *27*(3), 303-334.
- Terblanche, N. S., & Kidd, M. (2021). Exploring an in-store customer journey for customers shopping for outdoor apparel. *Journal of Retailing and Consumer Services*, *63*, 102722.
- Tong, S., Luo, X., & Xu, B. (2019). Personalized mobile marketing strategies. *Journal of the Academy of Marketing Science*, *48*(1), 64-78
- Wang, Y., Zhu, J., Liu, R., & Jiang, Y. (2024). Enhancing recommendation acceptance: Resolving the personalization–privacy paradox in recommender systems: A privacy calculus perspective. *International Journal of Information Management*, *76*, 102755.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, *39*(5), 806-820.
- Westermann, T., Möller, S., & Wechsung, I. (2015). *Assessing the Relationship between Technical*

*Affinity, Stress and Notifications on Smartphones*. Paper presented at the Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct.

- Xu, H., Luo, X. R., Carroll, J. M., & Rosson, M. B. (2011). The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. *Decision Support Systems*, 51(1), 42-52.
- Yadav, M. S., & Pavlou, P. A. (2020). Technology-enabled interactions in digital environments: a conceptual foundation for current and future research. *Journal of the Academy of Marketing Science*, 48(1), 132-136.
- Zhang, B., & Sundar, S. S. (2019). Proactive vs. reactive personalization: Can customization of privacy enhance user experience? *International Journal of Human-Computer Studies*, 128(8), 86-99.
- Zhang, S., Zhao, L., Lu, Y., & Yang, J. (2016). Do you get tired of socializing? An empirical explanation of discontinuous usage behaviour in social network services. *Information & Management*, 53(7), 904-914.

# **RESPONSIBLE DIGITAL: CO-CREATING SAFE, SECURE AND WISE DIGITAL INTERVENTIONS WITH VULNERABLE GROUPS**

**Author Name**

*Affiliation*

*Research In progress*

## **Abstract**

*Responsible digital innovation encompasses the inclusion of ethical and responsible design and use of digital technologies as well as digital well-being and environmental responsibility. Having the knowledge and skills to navigate the digital world effectively, safely and wisely becomes critical when digital literacy and access to technologies are limited and livelihood possibilities are precarious such as in the context of vulnerable migrants. We present two short examples of sensitive contexts where we are working with vulnerable groups to co-create digital interventions aimed at improving their lives. We use the Responsible Research and Innovation (RRI) framework in its operationalised version (AREA) as a lens to reflect on our research-practice and to evaluate the co-creation activities undertaken. The paper aims to contribute to knowledge by using the AREA framework in the context of South-South migration; to methodology by highlighting the procedures followed with vulnerable groups; and to practice through the promotion of responsible practices.*

**Keywords:** Responsible Digital, Digital Intervention, Vulnerability, Co-creation, Safe, Secure and Wise Use

## **1.0 Introduction**

It is crucial to ensure the safe and responsible use of digital technology, especially when working with vulnerable groups. While common online risks include cybersecurity threats, digital harassment, bullying, and fake news, vulnerable individuals may face additional risks, often due to limited digital literacy and internet

access (Mancini et al., 2019; Maclellan, 2019). These risks encompass online privacy, hate speech, child safety, scams, surveillance, and lack of awareness of social media settings, legal rights and reporting mechanisms. These results could be part of the ‘unintended impacts’ Owen et al. (p.27, 2013) refers to while counterbalancing the knowledge that innovation produces. The rise of new technologies such as generative artificial intelligence has led to renewed interest in responsible digital innovation that takes account of the social and ethical consequences of digital technologies (Sambasivan and Holbrook, 2018; Okolo et al., 2023). The literature is sparse in relation to responsible digital interventions when working with people with limited digital skills, resources, and access (Wakunuma et al., 2021). The examples presented here involve vulnerable populations in South Africa and Nepal (Authors, 2023). Their vulnerabilities are linked to their migration status alongside other intersectional inequalities as well as low levels of digital literacy.

In this paper we present two cases relating to our experiences of co-creating digital solutions with migrants and tech developers in two different contexts, South Africa and Nepal<sup>1</sup>, and then reflect on what responsible digital means in the context of vulnerability. In both cases, the researchers take the role of facilitators (Sanders and Stappers, 2008) adopting a local-centric, non-techno-deterministic approach that supports local values and aspirations (Dutta, 2021). The next section briefly discusses responsible digital innovation and presents the Responsible Research and Innovation (RRI) framework, and its operationalised version, the AREA framework (Stilgoe et al., 2013; Owen, 2014), as our interpretive lens. The paper then presents our research and intervention context as well as the methods we used for our research and practice. This is followed by the findings and analysis of our two cases using the AREA framework. We end with reflections on responsible digital for vulnerable groups and identify some implications for theory, methods and practice.

## **2.0 Literature Review**

Responsible Digital encompasses the study of ethical and responsible design and use of digital technologies (safety and privacy, cybersecurity, laws and regulation) as well

---

<sup>1</sup> <https://ict4d.org.uk/technology-inequality-and-migration/interventions/>

as digital well-being and environmental responsibility (Burr and Floridi, 2020). Responsible digital behaviours refer to using digital technologies in safe and socially responsible ways and considering the impact of one's digital actions on oneself, others, and society as a whole. Key principles of responsible digital behaviours include digital literacy, online safety and privacy, cybersecurity, respectful communication, digital well-being, critical thinking, environmental responsibility, online reputation management, and digital citizenship (Trier et al., 2023; Burgess-Wilkerson et al., 2019). Responsible digital behaviours are crucial for fostering a safe and positive online environment, requiring both individual responsibility and collective efforts to benefit everyone in the digital ecosystem.

Responsible digital innovation frameworks provide a means to frame both the development and evaluation of new digital interventions. They aim to engage all relevant stakeholders to ensure the contextual relevance of digital systems and reduce the value tensions that often arise from their deployment (Anand and Brass, 2021). However, there is a gap in the literature on digital risks and responsible digital innovation in the context of migration between countries in Africa, Asia and South America (Authors, 2023), characterised by some of the largest migrant flows in the world. Indeed, many well-meaning digital initiatives specifically aimed at supporting migrants often fail due to a lack of shared understanding of the desired social outcomes. Our research has identified digital risks faced by vulnerable migrants against a backdrop of poor digital skills that often lead them to severely underestimate the potential for harm (Authors, 2022; Authors, 2021).

## **2.1 The Responsible Research and Innovation Framework and its operationalisation (AREA)**

Responsible Research and Innovation (RRI) has been developed 'to include computer science, robotics, informatics, and ICT' (Jirotko et al., 2017, p.64). It aims to contribute towards research and innovation processes to ensure they serve the public interest and involve more inclusive and democratic decision-making by including stakeholders affected by new technologies. Proponents of RRI seek to transform existing processes, making research and innovation more acceptable, desirable, and risk aware. A widely used approach to operationalise RRI is the framework developed

by Stilgoe et al. (2013) and adapted into a more actionable framework by the UK's EPSRC (Owen, 2014) as the AREA framework which emphasizes the following key stages in relation to RRI: Anticipation, Reflection, Engagement, and Action. As digital interventions are complex and context-dependent, with many competing interests and uncertainties, the aim of RRI/AREA is to create a dialogue between difference stakeholders to make them mutually responsive to each other when undertaking digital innovation activities (Jirotko et al., 2017). Informed discourses should be at the base of responsible innovation where participatory approaches with relevant stakeholders create discursive processes that can represent an ethical solution. This framework was deemed well suited to support the analysis of our interventions due to the complex and sensitive research contexts and the involvement of multiple stakeholders.

### **3.0 Research-practice context and methods**

Our research and practice<sup>2</sup> within the MIDEQ project on South-South migration, inequality and development<sup>3</sup>, involves bringing together migrants, migrant and civil society organisations, tech developers and researchers to craft digital interventions together that can address migrant-defined inequalities and thereby improve migrants' lives. The MIDEQ project was based on six migration corridors consisting of twelve countries of departure and of destination for migrants. Our intervention focused on one country of departure, Nepal, and one destination country, South Africa. Following digital surveys and interviews with migrants, returnees and family members in both contexts, we facilitated focus groups in South Africa and a sandpit in Nepal to arrive at collective decisions on the digital interventions that migrants believed could address inequalities faced by them and their families. Two main ideas for interventions emerged from these activities:

- information sharing via videos in South Africa (informative videos made by migrants for migrants, covering various topics alongside workshops to improve

---

<sup>2</sup> <https://www.mideq.org/en/themes/digital-technologies-and-inequality/>

<sup>3</sup> <https://www.mideq.org/en/>

- digital skills, including communication, video editing, online safety, and more); and
- information sharing via a digital portal in Nepal that provides links to relevant, updated and verified information for migrants, returnees, and their families, as well as the need for migrants to develop better digital skills.

In Nepal, almost all the discussions during our initial sandpit in September 2022 related to the positive impact of digital tech on migrants, although we had also raised the negative impacts as well as the need for safe, wise and secure use of digital tech. Given our research findings on the limited digital literacy among migrants and risks such as online scams and surveillance, we also explored in a collaborative ideas generation workshop (online) with migrants and their organisations in December 2022 what might be included in a one-hour training session on digital risks and the importance of safe, secure and wise use of digital tech. This led to a collective understanding of the digital risks faced by migrants and their information needs. This subsequently formed a central plank of the co-design of the digital interventions in Nepal<sup>4</sup>, resulting in the creation involving multiple stakeholders of free and open resources (Creative Commons license BY-SA) on digital preparedness for migrants. These training resources, available in six Nepali languages along with guidance notes, can be used by anyone to train migrants, and it is our hope that they can form part of the regular training provided by the government for all migrants.

In South Africa, while the issue of digital safety emerged during the focus groups, it only became a priority for our interventions during the sessions on digital skills and risks including a collaborative digital body mapping exercises (Jager et al., 2016) with migrants from more than 10 sub-Saharan countries living in South Africa. Here, participants created visual representations of their personal online information highlighting concerns around privacy, child safety, misinformation, online risks and data security. Subsequently, participants translated their newly gained knowledge into videos<sup>5</sup> and also began teaching online safety in their communities. Support, including a "train the trainer" programme, was provided by members of our team as well as external consultants. Across all the workshops, the responsible use of digital tools to

---

<sup>4</sup> <https://ict4d.org.uk/technology-inequality-and-migration/nepal/>

<sup>5</sup> <https://www.youtube.com/@FusionAvenueOfficial/videos>

protect vulnerable individuals from online harassment was central, particularly when dealing with social media and use of images and videos.

In both contexts, all further activities with migrants, including training and dissemination, have featured the safe, secure, and wise use of digital tech as a key part of the interventions. We also planned for sustainability by bringing together diverse stakeholders within each context who can continue to support the interventions beyond our project timeline.

## 4.0 Findings and Discussion

In this section, we use the AREA framework to reflect on our cases and to evaluate the co-creation activities undertaken. To aid our analysis, we use Jirotko et al.'s (2017) guidance on the application of the AREA framework through a series of questions for further exploration (see table 1).

**Figure 1. The AREA Plus framework.**

	<b>Process</b> Rhythm of ICT	<b>Product</b> Logical malleability and Interpretive flexibility	<b>Purpose</b> Convergence and pervasiveness	<b>People</b> Problem of many hands
<b>Anticipate</b>	Is the planned research methodology acceptable?	To what extent are we able to anticipate the final product, future uses, and impacts? Will the product be socially desirable? How sustainable are the outcomes?	Why should we pursue this research?	Have the right stakeholders been included?
<b>Reflect</b>	What mechanisms are used to reflect on process? How might we do it differently?	How do we know what the consequences might be? What might be the potential use? What do we not know? How can we ensure social desirability? How might we do it differently?	Is the research controversial? How might we do it differently?	Who is affected? How might we do it differently?
<b>Engage</b>	How can we engage a wide group of stakeholders?	What are the viewpoints of a wide group of stakeholders?	Is the research agenda acceptable?	Who prioritizes research? For whom is the research being done?
<b>Act</b>	How can your research structure become flexible? What training is required? What infrastructure is required?	What needs to be done to ensure social desirability? What training is required? What infrastructure is required?	How might we ensure the implied future is desirable? What training is required? What infrastructure is required?	Who matters? What training is required? What infrastructure is required?

Table 1. The AREA Plus Framework. Source: Jirotko et al., 2017, p.67.

### 4.1 Anticipate



Our approach to our interventions was very much one of working ‘with’ the migrants and all relevant stakeholders in the context of use. There were two sides to these co-creation activities: the migrants prioritised information sharing activities while our facilitation focused on mitigating the unintended harmful consequences of digital interventions and collectively the process led us to integrate aspects relating to the safe, secure and wise use of digital tech. While we anticipated the importance of discussing digital risks and safety among vulnerable groups through our research findings, we also felt strongly that this key dimension needed to emerge through our activities so that migrants themselves recognised its value and owned the resulting interventions. Only then might such interventions have any chance of becoming sustainable.

## **4.2 Reflect**

Our approach was very much one of facilitation and co-creation which involved a continuous cycle of action and reflection and further discussions with all relevant stakeholders. We obtained feedback not only through traditional approaches such as feedback forms but also through video feedback and informal conversations. However, on further reflection, we should have developed our formal monitoring, evaluation and learning (MEL) processes earlier in our practice with the direct involvement of other relevant stakeholders. Although we did conduct some MEL processes in every phase with participants followed by detailed de-brief sessions by the project team, ideally, we should have done more. The reflections that we did have nevertheless allowed for new opportunities to emerge during the dissemination phase for each context as well as new opportunities in new contexts with new partners. During the ongoing final stages of the project, we are focused on dissemination and MEL.

## **4.3 Engage**

We chose to prioritise the relationships between local migrant organisations, tech developers and researchers so that the resulting interventions are not imposed from above but rather emerge from the context of use. We created a core team of migrants/migrant organisations and tech developers in each country to drive the

interventions locally while also facilitating wider networks of local and international organisations that could help amplify our collective efforts. Throughout the project's timeframe, it was essential for us to maintain a strong collaborative relationship with our partners and pertinent organizations to ensure that they are adequately prepared to support, build on and develop subsequent interventions to enhance the lives of migrants. This was done through periodic online calls and very frequent WhatsApp messages and discussions.

#### **4.4 Act**

Our co-design approach was inherently flexible as we wanted the interventions above all to be context-relevant and emerge from our collective activities to develop ownership and sustainability. This required us, the “researchers”,<sup>6</sup> to listen and to be ready to challenge assumptions based on the literature. At each stage, we chose to prioritise what the migrants wanted while at the same time considering potential harmful consequences and mitigation. Developing interventions with vulnerable groups in resource-constrained contexts also require researchers to think carefully about the practicalities of working with such groups and the resources that might be required to enable them to participate despite various work-life challenges.

### **5.0 Conclusion: implications for research-practice**

Our research and practice within the MIDEQ project has found that migrants use digital technologies in many different ways in different contexts; applications specifically designed for them are largely unknown to them or not used; and the knowledge of the potential of digital technologies, and specifically of the most frequently used device, the smart phone, is limited as are the risks related to it. This led us to our digital interventions in the two countries reported in this paper.

Our contribution to knowledge lies in the application of the RRI/AREA framework in a different and sensitive context, providing valuable insights into the dynamics of digital responsibility among vulnerable populations. While the RRI/AREA approach

---

<sup>6</sup> Although all of the migrants with whom we were working were themselves also, in a sense, researchers.

provides a template for the responsible design of digital interventions and a tool for reflection, we also note that its application in resource-constrained contexts with vulnerable groups requires researchers to take a sensitive approach that constantly emphasises the potential of the interventions to do more harm than good. In such contexts, researchers may need to be brave and call off digital interventions that may exacerbate or create digital risks and inequalities.

Our research emphasizes the need to prevent the harm that an intervention can generate, even unintentionally, and protect participants from the risks associated with digital tech use, particularly when working with vulnerable groups. Therefore, from a methodological perspective, we would recommend the introduction of safe, secure and wise use of digital tech as a necessary part of the toolkit for researchers and practitioners active in this field.

By anticipating, reflecting, engaging, and acting (AREA) on the challenges associated with digital technologies, our ongoing interventions try to mitigate digital risks, particularly affecting vulnerable categories who could benefit from the ability to leverage the opportunities presented by digital technologies.

## References

- Authors, (2023, October)  
Authors, (2023, November)  
Authors, 2022  
Authors, 2021  
Anand, N., and Brass, I. (2021). Responsible innovation for digital identity systems. *Data & Policy*, 3, e35. doi:10.1017/dap.2021.35  
Burgess-Wilkerson, B., Hamilton, C., Garrison, C., and Robbins, K. (2019). *Preparing Millennials as digital citizens and socially and environmentally responsible business professionals in a socially irresponsible climate*. arXiv preprint arXiv:1901.06609.  
Burr, C., and Floridi, L. (2020). *The ethics of digital well-being: A multidisciplinary perspective*. Ethics of digital well-being: A multidisciplinary approach, 1-29.  
Dutta, U. (2021). *Communication design and co-creation of information solutions for sustainable social change at the margins*. Chapter 19, Handbook of communication and development, edited by Melkote, Srinivas R., Singhal, Arvind, Edward Elgar Publishing  
Jager, A. de, Tewson, A., Ludlow, B., and Boydell, K., *Embodied ways of storying the self: A systematic review of body-mapping*. In Forum qualitative sozialforschung/forum: Qualitative social research, vol. 17, no. 2. 2016.

- Jirotko, M., Grimpe, B., Stahl, B., Eden, G. and Hartswood, M., *Responsible research and innovation in the digital age*. Communications of the ACM 60, no. 5 (2017): 62-68.
- Maclellan, S. (2019). "Fake news," dangerous speech and mass violence: challenges for social media in the developing world. In *Media and Mass Atrocity: The Rwanda Genocide and Beyond* edited by Allan Thompson, Centre for International Governance Innovation
- Mancini T, Sibilla F, Argiropoulos D, Rossi M, and Everri M (2019). *The opportunities and risks of mobile phones for refugees' experience: A scoping review*. PLoS ONE 14(12): e0225684.
- Okolo, C. T., Aruleba, K., and Obaido, G. (2023). *Responsible AI in Africa—Challenges and Opportunities*. In *Responsible AI in Africa: Challenges and Opportunities*. (Eds, Eke, D.O., Wakunuma, K. and Akintoye, S.) Palgrave Macmillan, pp. 35-64.
- Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., and Guston, D. (2013). *A framework for responsible innovation*. Ch 2 in *Responsible innovation: managing the responsible emergence of science and innovation in society*, 27-50.
- Owen, R. (2014). *The UK Engineering and Physical Sciences Research Council's commitment to a framework for responsible innovation*. *Journal of Responsible Innovation*, 1(1), 113-117.
- Sambasivan, N., and Holbrook, J. (2018). *Toward responsible AI for the next billion users*. *Interactions*, 26(1), 68-71.
- Sanders, E. B. N., and Stappers, P. J. (2008). *Co-creation and the new landscapes of design*. *Co-design*, 4(1), 5-18.
- Stilgoe, J., Owen, R., and Macnaghten, P. (2013). *Developing a framework for responsible innovation*. *Research policy*, 42(9), 1568-1580.
- Trier, M., Kundisch, D., Beverungen, D., Müller, O., Schryen, G., Mirbabaie, M., & Trang, S. (2023). *Digital Responsibility: A Multilevel Framework for Responsible Digitalization*. *Business & Information Systems Engineering*, 65(4), 463-474.
- Wakunuma, K., Castro, F. D., Jiya, T., Inigo, E. A., Blok, V., and Bryce, V. (2021). *Reconceptualising responsible research and innovation from a Global South perspective*. *Journal of Responsible Innovation*, 8(2), 267-291.

# ANOTHER TOOL FOR THE TOOLBOX? STRATEGIC ISSUES, SOCIAL MEDIA, AND STAKEHOLDER SENSEMAKING

Tiantian Qin [ml18tq@leeds.ac.uk](mailto:ml18tq@leeds.ac.uk)

Josh Morton [j.morton1@leeds.ac.uk](mailto:j.morton1@leeds.ac.uk)

## Abstract

*Social media has democratised strategic communication. Strategists are adopting social media in their (digital) work as they evolve from the being discreet decision-makers operating behind the scenes to becoming prominent communicators at the forefront of strategic discourse. With this practice change, strategists must possess language skills to help others make sense of strategic issues, subsequently influencing stakeholder sensemaking in a desired direction. This paper asks how strategists are adopting social media for sensemaking on strategic issues. Our netnography research is conducted using the social media posts of strategists in FTSE 100 companies. The preliminary finding shows linguistic features such as expressed emotion and rhetoric are prominent in our data. We conclude with some emerging contributions where this work-in-progress intends to extend prior research by advancing understanding of specific discursive forms of strategic sensemaking afforded by social media.*

**Key words:** Social media, Strategizing, Strategy practice, Sensemaking.

## 1.0 Introduction And Background

Social media is, we contend, becoming yet another ‘tool’ for strategists and their (digital) work. The adoption of digital tools by strategists, including social media, has spurred a burgeoning interest at the intersection of strategy and information systems (Morton et al., 2022). An important theme in digital strategizing is the use of digital technology for strategic transparency with internal and external actors (Tavakoli et al., 2017). Social media empowers stakeholders with diverse access to information within the organisation, with or without the help of 'official' sources. This influences strategic communication to be more transparent, more authentic, and more empathetic (Quigley, 2012). While strategic communication was traditionally the role of peripheral actors such as public relations departments (Heavey et al., 2020), often with a temporal lag, social media has democratised strategic communication as it affords direct

strategic communication from strategists (e.g., C-suite level executives) to both internal (e.g., employees) and external stakeholders (e.g., customers, investors) (Heavey et al., 2020). The responsibilities of strategists extend beyond strategy formulation and communication within the organization to encompass strategy communication beyond the organizational boundary (Heavey et al., 2020). Therefore, social media offers distinct opportunities, and also challenges, to the conventional roles and work of strategists (Heavey et al., 2020; Morton et al., 2020).

Whilst prior studies have made strides in unpacking how social media is being used by strategists, such as in extant behaviours and practices (Heavey et al., 2020), the discursive intricacies of strategists' use of social media – including the language strategists use – remains unclear. This is surprising given the focus on the linguistic nature of strategy in prior strategy work (Vaara and Fritsch, 2022). As we know, language is vital for formulating, making sense of, and framing strategic issues (Bjerregaard & Jeppesen, 2023). A potential consequence of discursive practices is that they embody managerial sensegiving, with the intention of shaping stakeholder interpretations (Whittle et al., 2023). This is notable, as sensegiving is a process through which strategists can interpret and explain strategy to influence stakeholders' understandings of strategic issues and secure stakeholder endorsement – their sensemaking. This has implications for strategy communication afforded by social media, such as the degree to which a strategy is likely to garner acceptance (Cornelissen & Werner, 2014).

Nonetheless, there remains a paucity of work comprehending how strategists actually employ social media for the dissemination and communication of strategic issues. Examining these intricate strategic framing practices on social media holds significance in order to gain a comprehensive understanding of strategists' use of social media for communicating strategic issues and, subsequently, how this digital work impacts stakeholder sensemaking (Kohtamäki et al., 2022; Whittle et al., 2023). This considered, we argue that the examination of sensemaking stands as a pivotal driver for the evolution of strategists' roles in the digital era. Furthermore, it is imperative to direct more explicit attention to the particular linguistic aspects of strategy relating to strategists' digital work. In response, we pose the following research question: 'How are strategists adopting social media for sensemaking on strategic issues?'

As this short paper is based on work-in-progress, our intention is to outline our emergent findings and, looking ahead, how we aim to leverage these initial insights as a foundation for extending our research and to craft more precise contributions.

## 2.0 Method

In the context of social media strategists pertain to those who are responsible for the overall strategic direction of organisations as part of their role, including CEOs, senior managers, and founders (Heavey et al., 2020). Our empirical work is based on a constructionist netnography method, Netnography is a research method derived from ethnography that explores, and makes sense of, naturalistic internet conversations, languages, and interactions, including non-verbal interactions (Costello et al., 2017). Use of netnography allows this study to promptly observe strategists and their stakeholder communities in their natural settings and record and analyse their digital work, specifically their organic communications, instantaneously (Morton et al., 2020). Therefore, our study is also informed by (digital) ethnographic research in the strategic management domain (Vaara and Fritsch, 2022) and aims to provide inspiration to strategy research that wishes to adopt netnography.

We study the (strategy focused) social media posts of strategists from FTSE 100 companies, specifically looking at LinkedIn. This setting ensures a rich and varied sample. The data collection commenced in June 2023 and concluded in October 2023. The first step was to acquire the latest FTSE 100 lists from the London Stock Exchange website, followed by locating and accessing each companies' official website. The majority of company websites incorporate a section titled 'Our Leadership' or a similar designation, which provides the information on strategists (e.g., top management team members). Where needed, we referred to the annual reports from the company's website to identify the executives. We then used a social media scraping tool to retrieve relevant LinkedIn posts. Overall, data from 96 of 100 FTSE companies were available to be collected. We visited the earliest posts from strategists, some of which dated to approximately 10 years ago. We checked 968 strategists with a LinkedIn presence and aggregated the comprehensive posts from 438 strategists with an 'active' presence. Being 'active' was contingent on having 20 posts and exhibiting an obvious strategic focus.

Considering the voluminous nature of social media data, we are using a data analysis software, Linguistic Inquiry and Word Count (LIWC), to begin the process of better understanding our data. LIWC is used for analysing language, such as to understand thoughts, feelings, personality, and the ways of connecting with others (Pan et al., 2018). To show how LIWC

works for our research, here is an illustrative input drawn from a social media post by HSBC's CEO.

*"It's been a big week for sustainability in the financial sector. On Wednesday we helped launch the Net Zero Banking Alliance, which will play a critical role in improving the consistency and raising the ambition of net zero commitments in the banking sector. On the same day, I took part in an International Institute for Finance webinar on the work of the Sustainable Markets Initiative, and our industry's role in providing innovative solutions to complex climate problems. It's clear the financial sector's momentum showed that the financial sector's momentum towards net-zero is stronger than ever. Most now see that a transitional approach isn't just the right thing to do, it's also the smart thing to do. It will bring benefits to the natural world, and value to our clients, our shareholders, and the communities we serve. We mustn't let this momentum slip. With COP26 coming up in November, we have a fantastic opportunity, as an industry, to drive real change on sustainability and to turn our net zero commitments into practical action. It's something I know many of my colleagues at HSBC care deeply about, and it's one of my biggest priorities. HSBC will continue to be at the forefront of this effort, both in the lead-in to COP26 and beyond.#RacetoZero #GFANZ"*

Table 1 below shows the corresponding result from the above input:

<b>Traditional LIWC Dimension</b>	Your Text	Average for Social Media Language
I-words (I, me, my)	1.78	5.44
Positive Tone	4.00	5.03
Negative Tone	0.00	2.34
Social Words	7.56	6.74
Cognitive Processes	5.33	8.86
Allure	6.22	8.62
Moralization	0.00	0.27
<b>Summary Variables</b>		
Analytic	93.64	47.06
Authentic	72.01	62.38

**Table 1. LIWC example output**

Of prominence in this example is the absence of a negative tone, coupled with a notably lower-than-average use of first-person singular pronouns ("I-words") typical in social media language. Analytic denotes analytical or formal thinking, and Authentic reflects unfiltered, off-the-cuff language. This is only a glimpse of what LIWC can do with one dictionary, we plan to use different dictionaries to explore other potentials of our data in the future. In addition, as shown in the table, LIWC produces quantitative analysis, therefore might open the door for our research to become a mixed-method study.

### **3.0 Preliminary Findings**

In presenting our preliminary findings, we show insights from a sub-set of data analysed to date. In particular, we outline five emergent discursive practices and summarise these in *Table 1* with example data in relation to indicative strategic issues, prominent linguistic features, and their potential implications for sensemaking.



Discursive practices	Example data	LIWC suggestions/ indication/features	Indicative strategic issues	Linguistic features	Sensemaking potential
Expressing sentiments	<i>"I'm proud that we're co-launching the Net Zero Banking Alliance today, which brings together 43 banks from 23 countries to accelerate the transition to net-zero across the sector...and it would be great to see you there. It's a big day for sustainability in the finance sector"</i>	Absence of a negative tone, lower-than-average use of "I-words" typical in social media language.  Highly analytic and authentic language use.	Sustainability / Net Zero	Positive language  Rhetoric	Resonance on important issues
Evidencing experience	<i>"I've had the honour of working with His Majesty through his Sustainable Markets Initiative, and specifically the Financial Services Task Force...I'm delighted, therefore, that HSBC is supporting the Terra Carta Accelerator, a new fund that will bring together private sector financing for nature and biodiversity projects."</i>	Absence of a negative tone, relatively high social words usage.  Highly analytic, extremely low authentic language use.	Sustainability	Positive language  Rhetoric	Rhetoric to boost alignment to new strategy and generate emotional reactions  Management of credibility
Personalising content	<i>"We're just as enthusiastic as you are about working together..." "...although it's clear a couple of them (children) want my job!"</i>	Absence of a negative tone, high positive tone.  Surprisingly low authentic language use.	Acquisitions Equality	Direct language  Comedic relief	Increased intimacy of communication and engagement
Coalescing others	<i>"Take a look and let me know what you think." "What's been on your mind?"</i>	High presence of social words, cognitive process, and allure.	Gender equality	Direct language  Opening the discussion	Enables strategic input
Directing virtuousness	<i>"Most now see that a transitional approach isn't just the right thing to do, it's also the smart thing to do."</i>	High presence of positive tone, allure.  Low authentic language use.	Sustainability/ Net Zero	Rhetoric  Comparison  Emphasis	Manages legitimacy  Impression management, strengthen faith

**Table 2. Discursive practices of strategists on social media and their potential for sensemaking**

For concision, we present the findings here and then discuss these further in relation to our intended future work to conclude the paper.

#### **4.0 Discussion And Emergent Contributions**

In this paper, we explore how social media is being used by strategists to leverage their linguistic skills in giving sense to strategic issues, and consequently influencing the sensemaking processes of stakeholders. From our data, we see potential to draw more explicitly on extant behavioural theories such as language-based views of strategy (Vaara and Fritsch, 2022) or emotion perspectives (Whittle et al., 2023). As seen in our emerging findings, linguistic features such as expressed emotion and rhetoric are prominent and can lead to notable outputs such as resonance or alignment. There are also rich possibilities in utilising theories of technology affordances where our emergent findings show potential in connecting strategy practice and information systems streams to harness diversity in illustrating how social media is a form of sociomaterial practice where technology features enable distinct outputs for collective sensemaking.

## References

- Bjerregaard, T., & Jeppesen, F. (2023). *Framing strategy under high complexity: Processes and practices of ongoing reframing in the becoming of strategy*, *European Management Journal*, 41(3) 258-270.
- Cornelissen, J. P., & Werner, M. D. (2014). *Putting framing in perspective: A review of framing and frame analysis across the management and organizational literature*, *Academy of Management Annals*, 8(1) 181-235.
- Costello, L., McDermott, M. L., & Wallace, R. (2017). *Netnography: Range of practices, misperceptions, and missed opportunities*, *International Journal of Qualitative Methods* 16(1) 1-12.
- Heavey, C., Simsek, Z., Kyprianou, C., & Risius, M. (2020). *How do strategic leaders engage with social media? A theoretical framework for research and practice*, *Strategic Management Journal*, 41(8) 1490-1527.
- Kohtamäki, M., Whittington, R., Vaara, E., & Rabetino, R. (2022). *Making connections: Harnessing the diversity of strategy-as-practice research*, *International Journal of Management Reviews*, 24(2) 210-232.
- Morton, J., Amrollahi, A. & Wilson, A.D. (2022). *Digital Strategizing: An Assessing Review, Definition, and Research Agenda*, *Journal of Strategic Information Systems* 31(1).
- Morton, J., Wilson, A.D. & Cooke., (2020). *The digital work of strategists: Using open Strategy for Organisational transformation*, *Journal of Strategic Information Systems* 29(2).
- Pan, L., McNamara, G., Lee, J. J., Haleblian, J., & Devers, C. E. (2018). Give it to us straight (most of the time): Top managers' use of concrete language and its effect on investor reactions. *Strategic Management Journal*, 39(8), 2204-2225.
- Quigley, S. (2012). The power of transparency, authenticity and empathy as drivers of corporate stakeholder engagement. *Managing Corporate Communication: A Cross-Cultural Approach*, 191.
- Tavakoli, A., Schlagwein, D., & Schoder, D. (2017). *Open strategy: Literature review, re-analysis of cases and conceptualisation as a practice*, *Journal of Strategic Information Systems* 26(3) 163-184.
- Vaara, E., & Fritsch, L. (2022). *Strategy as language and communication: Theoretical and methodological advances and avenues for the future in strategy process and practice research*, *Strategic Management Journal*, 43(6) 1170-1181.
- Whittle, A., Vaara, E., & Maitlis, S. (2023). *The role of language in organizational sensemaking: An integrative theoretical framework and an agenda for future research*, *Journal of Management*, 49(6) 1807-1840.

# Operationalizing Algorithmic Fairness: Empirical Study and Sociotechnical Framework Proposal

**Fredrik Wang**

*Department of Computer Science, Norwegian University of Science and Technology*

**Ilias O. Pappas**

*Department of Computer Science, Norwegian University of Science and Technology*

*Department of Information Systems, University of Agder*

**Polyxeni Vassilakopoulou**

*Department of Information Systems, University of Agder*

## **Abstract**

*This paper explores how organizations approach and operationalize algorithmic fairness in practice. Through semi-structured interviews with practitioners from organizations in Norway, insights were gained around their algorithmic fairness approaches and implementations. A thematic analysis revealed key considerations around starting early, law and regulations, the business value of fairness, challenges of identifying intersectional bias and technical solutions for ensuring and continuously monitoring fairness. An Extended Sociotechnical Framework for Algorithmic Fairness is proposed to help organizations address algorithmic fairness as a multifaceted issue. The framework categorizes general and case-specific factors across technical and social domains to provide structure while emphasizing context-specificity. It suggests harmonizing technical and social components to support practitioners navigating this complex area. The study provides empirical evidence of real-world fairness operationalization.*

**Keywords:** Algorithmic Fairness, Sociotechnical Systems, Responsible AI

## **1.0 Introduction**

The use of different types of Artificial Intelligence (AI) technologies including machine learning is today more widespread than ever. Ensuring that AI systems do not disproportionately favour or harm individuals or groups is critical. Discoveries of unfair algorithmic outcomes make headlines (Constantaras et al. 2023.; Asher-Schapiro 2020; Angwin et al. 2016). Algorithmic fairness means that algorithmic systems treat individuals and groups equitably, without discrimination or bias (Binns, 2018). Algorithmic fairness received increased attention from the research community, but mostly in terms of developing statistical definitions and mathematical

approaches for identifying and mitigating bias (Chouldechova 2017). Statistical notions of fairness are easy to measure, however, comprehensive operationalisations of the fairness concept require domain-specific expert input and opinion (Verma and Rubin, 2018). Hence, research beyond statistical formulations is needed to gain insights about algorithmic fairness in different application domains.

The objective of this paper is to develop a better understanding of how organizations approach algorithmic fairness, from initial discussions to deployed solutions. Specifically, the paper aims to answer the following research question: How do organizations approach and implement algorithmic fairness in practice? To answer this question, we collected and analysed empirical data collected by interviewing nine participants from eight different organizations. The insights are consolidated in The Extended Sociotechnical Framework for Algorithmic Fairness and recommendations for future work.

The remainder of this paper is structured as follows. Section 2 provides an overview of relevant background literature. Section 3 describes the method followed for data collection and analysis. Section 4 presents the main findings regarding organizations' fairness approaches and implementation experiences. Section 5 provides a discussion of these findings and introduces our proposed Extended Sociotechnical Framework for Algorithmic Fairness. Finally, Section 6 concludes the paper and outlines limitations and directions for future work.

## **2.0 Background**

AI systems affect many aspects of everyday life especially through algorithmic decision support (Adensamer, Gsenger, and Klausner 2021; Holten Møller, Shklovski, and Hildebrandt 2020). Such algorithmic support is used for instance in hiring (Langenkamp, Costa, and Cheung 2020), loan assessments (Sheikh, Goel, and Kumar 2020) and rankings used for recommender systems (Biega, Gummadi, and Weikum. 2018). However, studies conducted by researchers and regulators found algorithms to reflect and even amplify historical bias, and also potentially introduce biases of their own accord (Mehrabi et al. 2022). Algorithms containing bias can unfairly discriminate against minorities or discriminate on the basis of gender, age and language. Algorithmic unfairness has been identified across a wide range of fields including welfare (Constantaras et al. 2023), healthcare (Obermeyer et al. 2019),

judiciary services (Angwin et al. 2016), and education (Asher-Schapiro 2020). These serve as constant reminders that the use of AI comes with discrimination risks.

The concept of fairness is differently used across disciplines: philosophers consider fairness in terms of morality, social scientists often consider fairness in light of social relationships, power dynamics, institutions, and markets, quantitative fields have studied questions of fairness as pure mathematical problems (Mulligan, et al. 2019). For more than 20 years researchers have been studying bias in computer systems and pointing to the risks of biased systems (Friedman and Nissenbaum 1996).

The term algorithmic fairness refers to technological solutions designed to prevent systematic advantages or disadvantages to certain groups. In other words, algorithmic fairness means that algorithmic systems treat individuals and groups equitably, without discrimination or bias (Binns, 2018). From a technical standpoint, it is possible to introduce mathematical measurements of bias that can be used to develop computational approaches to minimize discriminatory outputs in machine learning against specific groups (Chouldechova 2017). However, as fairness is not merely a technical concept it has to be approached from a sociotechnical standpoint (Dolata, Feuerriegel, and Schwabe 2022).

### **3.0 Research Method**

In order to collect data, nine semi-structured interviews were performed with people from eight different organizations. Semi-structured interviews allow for the discovery of unforeseen information as they accommodate interviewees' decisions about what is important and relevant to talk about (Schultze and Avital 2011). The interviews were guided by an interview guide which was structured into general questions first, such as background, role, and fairness impressions, and then asking about the approach followed in the specific organization. It was also sometimes beneficial to ask follow-up questions that were not in the guide as issues emerged from the participant's answers. Hence, the interviewer allowed for development of the plot (Myers and Newman 2007) during each interview based on the input of the interviewees.

The interviews were performed between February and April 2023, were transcribed and recorded, and all participants signed consent forms. Interviews were conducted over Microsoft Teams with video and audio. The organizations were selected on the basis of their experience in developing and deploying AI solutions and we aimed to

cover different industries and also both public and private organizations. Table 1 provides an overview of the interviews performed.

Participants were identified in three ways. One way was by contacting those who had participated in public conferences where algorithmic fairness was a topic, or similarly had published articles or academic papers where algorithmic fairness was a topic or subtopic. The second way was using the authors' network. The third way was using LinkedIn to search for topics like 'algorithmic fairness' and similar, to find people who worked with machine learning and AI in companies where it would be logical for fairness to be a part of their projects. When participants were recruited, they were given some instructions about what to expect the interview to be about. In this way, they would have some time to think about their views regarding the topic. Sending information about topics and questions to allow the interviewee to prepare can have a positive effect (Oates 2005). Another benefit of this is that it can alleviate some of the pressure from interviewees.

IDs	Company	Role	Organization Size	Duration
R1	State-owned enterprise	Data Scientist	20000	55 min
R2	State-owned enterprise	Lawyer	20000	50 min
R3	State-owned enterprise	Data Scientist	500	50 min
R4	Private Research	Research Director	100	50 min
R5	Insurance	Director	4000	50 min
R6	Private Corporation	Data Scientist	500	45 min
R7	State-funded enterprise	Senior Advisor	100	55 min
R8	Private Corporation Company	Lawyer	100	55 min
R9	State-owned enterprise	Technologist	100	40 min

**Table 1. Overview of Interviews**

Thematic analysis was adopted to analyse the qualitative data collected (Oates 2005). Thematic analysis is a method for identifying, analysing, organizing, describing, and reporting themes found within a data set. All interview transcripts were thoroughly analysed to generate initial codes summarizing key concepts and patterns found. The codes were consolidated further to form overarching themes representing important patterns within the data in relation to the research topic and objectives. This thematic analysis process allowed for a rich, detailed, and nuanced interpretation of the perspectives and experiences described by participants.

## **4.0 Findings**

### **4.1 Starting Early with Fairness Considerations**

One aspect of working with algorithmic fairness is that you can't simply begin considering it when your model is already deployed and affecting people. Starting early with fairness is essential for success. One of the participants stated:

*“When working with fairness you need to start early, not just because of legal considerations, but also because it affects the design and product development of the solution.”*

Starting early also has the benefit of reducing the need for costly and time-consuming revisions further down the line of product development. Retroactively integrating fairness considerations in an AI system that is operational can be complex and costly, and it's generally more cost-effective to prevent unfairness in the first place rather than to face the aftermath of algorithmic bias.

### **4.2 The Role of Law and Regulations for Achieving Fairness**

It is normally agreed upon that fairness is something one wants to achieve, the question is how it should be achieved. The interviews revealed that there are legal considerations related to achieving fairness that need to be taken into account. One of the interviewees explained how the laws have significant impact in achieving fairness:

*“You can make a model and test its performance and fairness, but legislators can decide that certain groups should be prioritized over other groups, and then the model would have to be “unfair” first so that it complies with the law before it can be fair to other groups.”*

Rapidly evolving legal and regulatory landscapes surrounding fairness can cause uncertainty for companies. Understanding and complying with laws and regulations related to fairness can be challenging. Furthermore, achieving this understanding can be expensive if the company doesn't already have these resources.

### **4.3 The Business Value of Fairness**

An aspect that can serve as a motivation in several contexts, is fairness as a selling point, with fairness adding business value. Fair algorithms enhance brand reputation and foster trust among customers and stakeholders. In an era where customers increasingly value ethical business practices, companies demonstrating a commitment to fairness can differentiate themselves in the market. Investing in algorithmic fairness



is not just a matter of ethics and compliance, but also a sound business strategy that drives long-term value and competitiveness as explained by one of the participants:

*“We believe that implementing fairness, along with transparency and responsibility, will drive business value, and those who are best at it will have a competitive advantage. ... fairness will become a selling proposition.”*

Fair algorithms can also lead to better and more inclusive decision-making. They can uncover and correct biases that may have traditionally limited business opportunities, such as in hiring, lending, or marketing. This leads to a more diverse and inclusive customer base and workforce, which are known to improve creativity, innovation, and profitability. Lastly, fairness can reduce the risk of costly litigation and penalties associated with unfair or discriminatory practices.

#### **4.4 Challenges of Identifying Intersectional Bias**

Bias may not always be so easy to spot, proxies can make it difficult to identify bias. Similarly, discrimination that only happens at intersectionality makes it difficult to understand when unfair treatment is happening. Intersectionality refers to the way different aspects of a person's identity (such as gender, race, sexual orientation, etc.) combine and overlap to expose them to various forms of discrimination or unfair treatment. Bias can occur when multiple aspects of a person's identity intersect, such as discrimination against women with immigrant background but not necessarily men with immigrant background or women on their own. This type of intersectional discrimination may be difficult to identify because the bias is not evident when only examining one aspect of identity, such as gender or background alone. One would need to analyze how different personal attributes combine before the unfair treatment resulting from their intersection is detectable. So intersectional discrimination makes it more challenging to pinpoint precisely when and how unfair treatment is taking place within a system compared to bias along single identity dimensions. Several participants pointed out that there is a lack of systematic methods for discovering bias and unfairness, and discrimination happening at the intersectionality of attributes is an example of bias that won't be discovered easily.

#### **4.5 Technical Solutions for Ensuring and Continuously Monitoring Fairness**

Participants stated that with today's toolkits, the technical aspect is a very small part of implementing algorithmic fairness. For instance:

*“The technical implementation is a small part, the tools, and frameworks support you to check that your algorithm is implemented correctly and saves you*

*from a lot of troubleshooting. It's a small part, but it's reassuring to have it in place."*

For classification and regression problems one can use techniques such as feature importance to see what attributes the model utilizes the most in its prediction. Through these techniques, practitioners can identify potential biases in their model. One of the participants explained a project where they revealed bias by looking at the feature importance of the model:

*"By using feature importance methods we were able to see the model being discriminatory towards gender, and pointed out that this unfairness should be looked into even though the project is in an early phase."*

Once fairness metrics are determined it's important to continuously monitor the system against this. One of the participants stated the following:

*"... the AI must be checked against this limit continuously. This is, for example, because the composition of the group of people the AI is used on can change, or the algorithm can become biased over time if it learns from and systematizes biases gradually."*

## **5.0 Discussion**

The research on algorithmic fairness has mostly been concerned with statistically defining fairness and then proposing methods and techniques to mitigate undesirable biases, in relation to these definitions (Agarwal et al. 2018). Whilst practitioners to some degree were also concerned with implementing statistical metrics, the overall takeaway from the interviews is that the most difficult part of algorithmic fairness is to decide what constitutes fairness in each specific context (Selbst et al. 2019). This requires domain knowledge and also, understanding of regulatory provisions.

Evaluating the fairness of an AI model requires a definition of fairness. Thus, understanding the context and assessing the impact of the system is pivotal algorithmic fairness. Relying on intuition for discovering unfairness is a risky strategy, but is often the chosen strategy, due to the lack of support to address the issue. A study by Holstein and colleagues also found that most industry practitioners rely on their intuitions, even though these were often found to be wrong (Holstein et al. 2019).

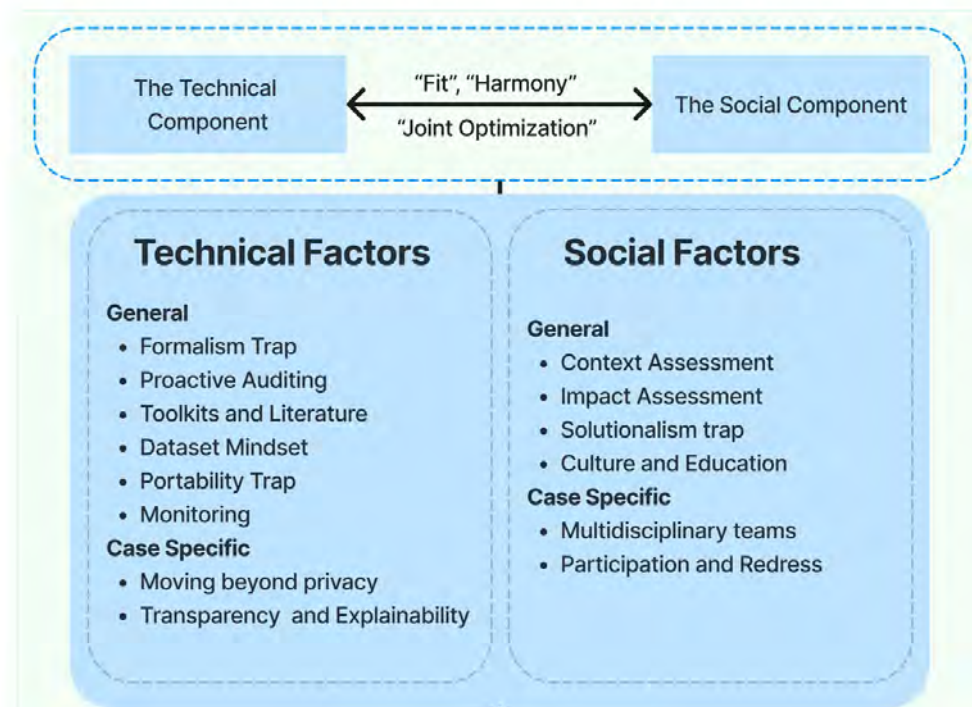
Data quality and sufficiency is a key challenge because data may contain historical bias which an AI model trained on these data will reflect (Roselli, Matthews, and Talagala 2019). Similarly, data may be affected by the conscious or unconscious bias in the people who collect the data. Having enough data is also a challenge as unprivileged groups are often underrepresented. There are also cases where data isn't available, such as when all outcomes aren't observable. An example is getting a rejection for a loan, where one still doesn't know if the loan would have been paid back if it had been approved (Verma et al. 2020). Expanding on this, if a loan was approved and later paid back, one does not know if the loan would have been paid back had circumstances been different (e.g., larger loan amount or longer loan term). Models that are biased against certain groups could continue to reject candidates from that group and we would never be able to have data on the outcomes if these decisions were not taken.

### **5.1 The Extended Sociotechnical Framework for Algorithmic Fairness**

Based on the results from the thematic analysis and the literature, a framework for understanding how practitioners can advance toward algorithmic fairness has been created (Figure 1). The framework is expanding on the work of Sarker and colleagues (Sarker et al. 2019) aiming to a harmonization between technical and social components. The technical components involve things like developing mathematical definitions of fairness, implementing algorithmic mechanisms to mitigate bias, and assessing models for unfair outcomes. Meanwhile, the social components pertain to high-level issues like organizational policies, legal/regulatory landscapes, sociocultural biases, and stakeholder values. Rather than seeing these as separate concerns, harmonization aims to bridge the gap between the technical and social domains. The goal is to develop an integrated approach where the technical solutions account for relevant social factors, and social/policy decisions are informed by technical considerations. This harmonization of the technical and the social is key to operationalizing fairness in real-world applications.

The suggested framework is split into four main categories, consisting of General Technical Factors, Case-specific Technical Factors, General Social Factors, and Case-specific Social Factors. This structure distinguishes factors that have broad relevance across all organizations addressing algorithmic fairness versus those more tailored to individual situations. The classification makes the framework less overwhelming and easier to apply while also showing the context-dependent nature of algorithmic

fairness initiatives, emphasizing that there is no one-size-fits-all solution to algorithmic fairness (Morse et al. 2021). By splitting factors in this way, users can identify baseline technical and social elements to address generally as well as those necessitating adaptation.



**Figure 1. Extended Sociotechnical Framework for Algorithmic Fairness.**

Factors such as performing *proactive auditing* in order to avoid bias from the start and having mechanisms in place to handle emerging bias as data and model parameters change are crucial. Using appropriate toolkits can help in properly implementing the technical part of the solution and ensuring that the outcomes are equitable. Recognizing that reusing algorithmic solutions that were designed for a specific context could lead to inaccuracies or cause harm can help prevent algorithmic systems from further marginalization and exclusion, and thus foster both inclusivity and diversity. Having a *dataset mindset* is an example of a crucial factor. Improving the quality of datasets is key for both better accuracy and fairness. Having a dataset that better represents the real world can increase diversity. Staying up to date with technical solutions, such as the described toolkits is one way that companies can take a more structured and active approach to fairness.

In the social factors, improving *culture and education* about fairness is key. Similarly, performing an *impact assessment* can help understand who is affected by the algorithmic outcomes, and help recognize that algorithmic systems can have significant effects on the life of individuals. An example of a case-specific social factor is: *Mechanisms for pooling knowledge across teams* so that one can develop the right solutions depending on the system and context. This factor is case-specific because it is only relevant for organizations that have multiple AI teams. It would not apply to a small company where sharing knowledge across teams wouldn't be an issue. Similarly, companies could, for several reasons, such as privacy or security, not have the possibility to allow affected individuals to participate and raise concerns, even though this would likely be beneficial. The full list of factors included in the framework along with their descriptions is provided in Table 2 below.

<b>Factor</b>	<b>Description</b>
Formalism Trap	Mathematical definitions eliminate the nuances of fairness.
Proactive Auditing	Aspire to implement fairness from the beginning, instead of mitigating unfairness later.
Toolkits and Literature	Using state-of-the-art toolkits for technical evaluation and implementation and staying updated on research.
Dataset Mindset	Ensuring that the data are complete and of good quality.
Portability trap	Recognizing that reusing algorithmic solutions, originally designed for a specific social context, could lead to misinterpretations, inaccuracies, or potentially cause harm when implemented in a different context.
Monitoring	Maintaining that the outcomes are fair and prevent bias and unintended consequences after initial development and deployment.
Moving beyond privacy	Understand that an AI system could respect privacy (by properly handling personal data) or be sheltered from sensitive attributes, but still be unfair (if it produces biased outcomes).
Transparency and Explainability	It is important to understand how we get the specific algorithmic outputs. Explainable AI techniques can help achieve this.
Context Assessment	Assessing the context of a system and how this affects how fairness is approached and defined, and who should be involved.
Impact Assessment	Assessing the impacts of an algorithm and potential negative outcomes necessitates understanding its social context and the varied notions of fairness within that system.
Solutionism trap	Overlooking the possibility that the optimal solution may not always involve technology can lead to missteps.
Culture and Education	Develop a culture for fairness. Necessary for developing domain-specific guides, algorithms, metrics, ethical frameworks, and case studies.
Multidisciplinary teams.	Contribute to a comprehensive understanding of biases, ethics, and social implications in algorithmic systems. Foster critical thinking, challenge assumptions, and promote creative problem-solving, leading to robust and equitable solutions.
Participation and Redress	Affected individuals and communities should have the opportunity to participate in decision-making about algorithmic systems, and there should be mechanisms for redress if the algorithm causes harm.

**Table 2. Factors in the Extended Sociotechnical Framework for Algorithmic Fairness**

## 6.0 Conclusion

Overall, this study aims to provide insights on real-world algorithmic fairness implementation. The sociotechnical perspective taken acknowledges fairness as a multifaceted issue and extends algorithmic fairness research by providing an understanding of actual practices and experiences. Based on the insights from interviews with practitioners we suggest the comprehensive *Extended Sociotechnical Framework for Algorithmic Fairness*. The framework can help practitioners and organizations understand how they can approach algorithmic fairness and gain a better understanding of their own situation and context. For the research community, it provides a first step towards operationalizing algorithmic fairness for organizations that develop and deploy AI systems. A key limitation of this study is the relatively low number of organizations that participated and the fact that they are all located in Norway. Factors in the proposed framework were also found in the literature from other countries, but further investigations in different contexts and geographical locations are certainly needed. Future research could also focus on one specific industry (e.g. healthcare) to create targeted frameworks accounting for contextual algorithmic solution factors.

## References

- Adensamer, Angelika, Rita Gsenger, and Lukas Daniel Klausner. (2021) “ ‘Computer Says No’: Algorithmic Decision Support and Organisational Responsibility.” *Journal of Responsible Technology* 7: 100014.
- Agarwal, Alekh, Alina Beygelzimer, Miroslav Dudík, John Langford, and Hanna Wallach. (2018) “A Reductions Approach to Fair Classification.” In *International Conference on Machine Learning*, 60–69.
- Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. (2016) “Machine Bias.” *ProPublica*. Accessed May 26, 2022.  
<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Asher-Schapiro, Avi. (2020) “Global Exam Grading Algorithm Under Fire for Suspected Bias.” *Reuters EverythingNews*. Accessed May 10, 2022.  
<https://www.reuters.com/article/us-global-tech-education-analysis-trfn-idUSKCN24M29L>.
- Biega, Asia J., Krishna P. Gummadi, and Gerhard Weikum. (2018) “Equity of Attention: Amortizing Individual Fairness in Rankings.” *41st International ACM SIGIR Conference on Research and Development in Information Retrieval*, 405-414.
- Binns, R. (2018) “Fairness in machine learning: Lessons from political philosophy”. *ACM Conference on Fairness, Accountability, and Transparency (FAccT 2018)*,

- Chouldechova, Alexandra. (2017) "Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments." *Big Data* 5 (2): 153–63.
- Constantaras, Eva, Gabriel Geiger, Justin-Casimir Braun, Dhruv Mehrotra, and Htet Aung. (2023) "Inside the Suspicion Machine." *Wired*. Accessed March 26, 2023. <https://www.wired.com/story/welfare-state-algorithms/>.
- Dolata, Mateusz, Stefan Feuerriegel, and Gerhard Schwabe. (2022) "A Sociotechnical View of Algorithmic Fairness." *Information Systems Journal* 32 (4): 754–818.
- Friedman, Batya, and Helen Nissenbaum. (1996) "Bias in computer systems." *ACM Transactions on Information Systems*, 14 (3): 330-347.
- Holstein, Kenneth, Jennifer Wortman Vaughan, Hal Daumé III, Miro Dudik, and Hanna Wallach. (2019) "Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need?" In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–16.
- Holten Møller, Naja, Irina Shklovski, and Thomas T Hildebrandt. (2020) "Shifting Concepts of Value: Designing Algorithmic Decision-Support Systems for Public Services." In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, 1–12.
- Langenkamp, Max, Allan Costa, and Chris Cheung. (2020) "Hiring Fairly in the Age of Algorithms." *arXiv Preprint arXiv:2004.07132*.
- Mehrabi, Ninareh, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. (2022) "A Survey on Bias and Fairness in Machine Learning." <https://arxiv.org/abs/1908.09635>.
- Morse, Lily, Mike Horia M Teodorescu, Yazeed Awwad, and Gerald C Kane. (2021) "Do the Ends Justify the Means? Variation in the Distributive and Procedural Fairness of Machine Learning Algorithms." *Journal of Business Ethics*, 1–13.
- Mulligan, Deirdre., Joshua Kroll, Nitin Kohli, and Richmond Wong. (2019) "This thing called fairness: Disciplinary confusion realizing a value in technology." *Proceedings of the ACM on Human-Computer Interaction*. CSCW, 1-36.
- Myers, Michael D., and Michael Newman. (2007) "The qualitative interview in IS research: Examining the craft." *Information and organization*. 17(1): 2-26.
- Oates, Briony J. (2005) *Researching Information Systems and Computing*. Sage.
- Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. (2019) "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science* 366 (6464): 447–53.
- Roselli, Drew, Jeanna Matthews, and Nisha Talagala. (2019) "Managing Bias in AI." In *Proceedings of the 2019 World Wide Web Conference*, 539–44.
- Sarker, Suprateek, Sutirtha Chatterjee, Xiao Xiao, and Amany Elbanna. (2019) "The Sociotechnical Axis of Cohesion for the IS Discipline: Its Historical Legacy and Its Continued Relevance." *MIS Quarterly* 43 (3): 695–720.
- Selbst, Andrew D, Danah Boyd, Sorelle A Friedler, Suresh Venkatasubramanian, and Janet Vertesi. (2019) "Fairness and Abstraction in Sociotechnical Systems." In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 59–68.
- Sheikh, Mohammad Ahmad, Amit Kumar Goel, and Tapas Kumar. (2020) "An Approach for Prediction of Loan Approval Using Machine Learning Algorithm." In *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 490–94. IEEE.
- Schultze, Ulrike, and Michel Avital. (2011) "Designing interviews to generate rich data for information systems research." *Information and Organization* 21(1): 1-16.

Verma, Sahil, and Julia Rubin. (2018) “Fairness Definitions Explained.” In 2018 IEEE/ACM International Workshop on Software Fairness (FairWare). IEEE Computer Society,

Verma, Sahil, Varich Boonsanong, Minh Hoang, Keegan E Hines, John P Dickerson, and Chirag Shah. (2020) “Counterfactual Explanations and Algorithmic Recourses for Machine Learning: A Review.” arXiv Preprint arXiv:2010.10596.



## **APPENDIX**

### **Contextual data about the study participants and their organizations**

This annex presents the participants in the study, including their title and experience, as well as information about the organization they work for and a description of the relevant systems or projects they have partaken in. Descriptions are made as accurate as possible without exposing sensitive information about the participant, nor the company they work for. The descriptions are provided to better understand the results.

#### ***R1 Data Scientist***

R1 works in a company with around 20000 employees. R1 works as a Data Scientist and has done so for the company in the last 5 years. The company R1 works for is a public agency that partakes in several fields, the most relevant to R1 is welfare. The company is working on an AI project that focuses on predicting the progression of individuals on sick leave. The system's prediction would not be the final output but instead, be given to a case manager who would use this information along with other information in order to make a final decision. R1 is thus concerned with algorithmic fairness in regard to an automatic decision-support system that would affect people's life. R1 also works on developing other AI systems, but this is the one that is the most relevant. R1 also follows the literature that is done on algorithmic fairness, such as by researching different toolkits that are available.

#### ***R2 Senior Advisor.***

R2 works in the same company as R1 and is a lawyer. They work with the same projects as R1 does but have a different role, as R2's main role is to give legal advice to different teams using machine learning. This includes making sure that the machine learning systems follow the law, and requirements such as fairness, explainability, transparency, and privacy. Assuring that the translation between law and code is correct is one task that is particularly important. R2 expertise does not lie in the technical aspects of algorithmic decision-making, instead, they use their legal expertise in order to oversee the translation between law and code that is done by developers and data scientists.

### ***R3 Data Scientist.***

R3 is educated as a sociologist but now works as a data scientist in a company with around 500 employees. R3 works in a company that specializes in auditing and controlling various systems and solutions. They work in the company's artificial intelligence department, where tasks include auditing machine learning systems and algorithms, and this is where the relevance of algorithmic fairness comes from in the work that R3 does. R3 follows the literature regarding algorithmic fairness and other publications about artificial intelligence and has also authored papers about artificial intelligence and fairness. They work both on implementing machine learning in their own systems and processes and also auditing other companies' use of machine learning and algorithms. Certain projects R3 has worked on were in relation to analyzing and auditing machine learning algorithms and checking for certain biases.

### ***R4 Research Director.***

R4 works as a researcher specializing in machine learning in a company with around 100 employees, R4 has 20+ years of experience. R4 works tightly with both companies and research institutions. R4 stays updated with algorithmic fairness research, and the increase in literature is part of why R4 has taken a special interest in the field. R4 often works on projects where R4 or R4's team only has partial responsibility such as only being in charge of the technical implementations, whereas another team has the superior responsibility, which may include deciding the fairness definition. Their task in these projects is usually to design the algorithm used in a solution and implement fairness accordingly.

### ***R5 Department Director.***

R5 works for an insurance company with around 4000 employees and has studied economics. They work as a department director and has 10 years of experience. In order to process insurance claims and decide insurance premiums, the company employs thousands of machine learning models. R5 has a long experience with insurance and the use of machine learning within the insurance context. Algorithmic fairness is vital for R5 along with other aspects of RAI. Fairness is a relatively new concept in regards to the use of algorithms, but at the same time seen as very important, and a key factor for the future in terms of reputation and business value.

### ***R6 Data Scientist.***

R6 has worked with algorithmic fairness both as a researcher as well as working as a Data Scientist. They work for a company that makes safety software and has around 500 employees. The current company of R6 is in the process of implementing more and more machine learning in order to streamline their solutions, although it's still at an early stage. R6 has previous experience working for an IT consulting company, where among other things they would provide solutions for implementing algorithmic fairness in AI systems. R6 also follows the literature and has attended several conferences on fairness in AI. Through this work as well as staying up to date with the literature, R6 has a good overview of existing solutions and toolkits.

### ***R7 Senior Advisor.***

R7 has a background in the social sciences and is now working as a senior advisor in a company with around 100 employees. They have more than 5 years of experience working with the use and effects of AI. R7 works for a company specializing in consumer rights, such as ensuring fair treatment when a system uses algorithmic decision-making. R7 thus provides a different view on algorithmic fairness, as they “represent” those affected by algorithms, as opposed to those who design and deploy them. As a consequence of this, R7 doesn't always have all of the tools for detecting algorithmic unfairness at their disposal, as they may not have all the data or outcomes available. Instead, they employ different methods for bias detection, such as algorithmic auditing and unsystematic approaches.

### ***R8 Lawyer.***

R8 is a lawyer who specializes in AI. R8 has worked at their current company for 3 years and the company has around 100 employees. R8 follows the research that is done and has a particular interest in algorithmic fairness. They work with client companies that wish to ensure that their AI systems are in line with legal regulations, which include ensuring algorithmic fairness. R8 is concerned with how the use of artificial intelligence challenges legal principles, and how bias in algorithms is a challenge to the principle of justice.

### ***R9 Data Scientist.***

R9 works as a Data Scientist for a company with around 100 employees specializing in digitalization and privacy. R9 has 10+ years of experience working with AI for different companies. Among other focus areas, the company that R9 works for leads artificial intelligence projects where different companies can try out and evaluate their systems. These projects often revolve around privacy and RAI, and around half of the projects are also concerned with fairness. R9 has partaken in these projects where algorithmic fairness is important, and the projects operate in several different contexts such as healthcare, welfare, and surveillance, where both technical and organizational solutions have been proposed to mitigate bias and implement algorithmic fairness.

# FACTORS INFLUENCING THE CAREER DECISIONS OF WOMEN SOFTWARE ENTREPRENEURS: PERSPECTIVES FROM INDIA AND IRELAND

*Complete Research Paper*

## **Abstract**

*Software entrepreneurship continues to be perceived as a sector dominated by men. Research points to key differences in the motivations to engage in entrepreneurship, stressing the significance of gender issues. Through a qualitative analysis approach, this research seeks to identify factors that influence women's decisions to pursue a career in software entrepreneurship in India and Ireland. Both countries rank highly as technology startup hotspots with rapidly growing women founder communities. This research undertakes a thematic analysis of six semi-structured interviews with women entrepreneurs from India and Ireland revealing two main themes of barriers and tactics. Six major barriers include gender stereotypes, confidence, male-dominated sector, family expectation, age, and power dynamics. Nine tactics emerge including building confidence, owning personal autonomy, family support and background, mentors, feminist mindset, educational impact, utilising specific skills, government policies, and flexibility in gendered roles. This study uncovers similarities and differences that merit further research.*

**Keywords:** Software entrepreneur, gender, women, barriers, tactics, India, Ireland.

## **1.0 Introduction**

The number of female entrepreneurs has been rising, significantly boosting economic growth and societal well-being (Kelley et al., 2017). However, many entrepreneurs are men (Brush et al. 2009; Sánchez-Escobedo et al. 2014; Dileo and Pereiro 2018). The recent Global Entrepreneurship Monitor (2022) report supports this highlighting that more women than men—2.7% versus 4.7%—establish businesses in the information, computer, and technology (ICT) sector.

One of the 17 Sustainable Development Goals (SDGs) of the United Nations is to achieve gender equality in all aspects of life (UN, 2021). However, several obstacles and limitations continue to prevent women from fully participating in entrepreneurial activities (Naidu and Chand 2017; Tur-Porcar et al., 2017). One of the main obstacles is gender stereotypes that characterise men as providers and women as homemakers,

these perceptions impede efforts to achieve gender equality (Brush et al., 2018), leading to the classification of men's and women's job trajectories. The lack of gender equality in entrepreneurship, particularly in the software field, and the fact that many women do not select entrepreneurship as a career path are evident when the subject of gender in entrepreneurship is examined using both qualitative and quantitative methodologies (Brush et al., 2018). Against this backdrop, our research aims to understand the factors that influence the career decisions of women software entrepreneurs in India and Ireland. This research focuses on women as a subject of study to investigate these factors by posing two questions:

1. *What barriers affect women software entrepreneurs?*
2. *How do women software entrepreneurs respond to these barriers?*

This study explores the challenges faced by women software entrepreneurs, particularly in the face of gender stereotypes and challenges such as accessing funding to support start-up ventures.

## **2.0 Literature Review**

Historically, gender and entrepreneurship studies have been divided into two approaches: the "Gender as a Variable approach" (GAV) and the "Gender as an influence" approach (Martinez and Marlow, 2018). The characteristics of a successful entrepreneur have been linked to masculine characteristics, leading to criticism when gender was used interchangeably with women (Schein, 1973). Research by Jennings and Brush (2013) focuses on the underrepresentation of women in entrepreneurship. Brush (1992) argues that women are "different," i.e., compassionate, and relational. Neergaard and Ulhoi (2007) bring a multiplicity of factors like gender, semiotics, media discourse, grounded theory, action research, ethnographic methodology, entrepreneurship as lived experience, critical realism, and discourse analysis to entrepreneurship, promoting the gradual shift towards the "Gender as influence" approach in qualitative research of gender in entrepreneurship.

According to Frederick et al. (2016), women encounter unique hurdles in entrepreneurship due to social conditioning differences. Socioeconomic obstacles can

prevent women from accessing necessary resources for business success (Watson, 2002). Female entrepreneurs are underrepresented in the business sector, leading to perceptions of inferiority or marginalisation in female-specific niches (Ahl and Marlow, 2021). According to Robb and Watson (2012), equal access to resources can help women achieve equal success in entrepreneurship, if not more.

## **2.1 Factors Influencing Women Software Entrepreneurs**

Women in the software sector face biases and underrepresentation (Kovaleva et al., 2023; Wilson and Patón-Romero, 2022). According to a study by Griffith (2010), the gender gap in STEM participation is not only a global issue but also one that is seemingly not diminishing over time. Between the 1980s and the turn of the century, the gender gap did not significantly narrow, with the gender gap in STEM widening as education levels rose (Griffith, 2010).

Many factors may contribute to the underrepresentation of women in entrepreneurship (Jennings and Bush, 2013), attitudes of parents, teachers, society, and the media at large are cited as key influencing factors (Adya and Kaiser, 2005). One early childhood bias may arise when parents select toys for their kids based on gender (Kollmayer et al., 2018). These toys may encourage different skill development and interests in boys and girls, which could in turn influence future behaviour (Kollmayer et al., 2018). Playing with building blocks, for example, fosters the development of spatial reasoning abilities, essential for engineering (Kovaleva et al., 2023). Although studies deny any gender difference in math abilities (Hyde, 2014), some teachers and parents still hold the belief that boys have higher mathematical potential in schools and girls are often encouraged to pursue the arts, whereas boys have different hobbies (Hyde, 2014). Powell et al.'s (2012) study on the selection process for engineering cited factors that influence students' decisions to pursue engineering as a career. Childhood interests in technology were one of the most important criteria identified (Powell et al., 2012). Students who chose to become engineers or work with technology as their future profession were more likely to be doing anything "hands-on" with technology (Powell et al., 2012).

The role of teachers in influencing students to pursue careers in technology and engineering is deemed an important one (Germeijs and Verschueren, 2006). Parental advice is one of the most significant influences on professional choice, particularly for

girls (Germeijs and Verschueren, 2006). Society sets the standards for acceptable behaviour for both boys and girls (Moss-Racusin et al., 2010). Additionally, children form their opinions regarding the world based on what is seen in the media and in their immediate surroundings, which still lack role models (Knoll et al., 2011). These elements influence females' decisions regarding their technical and scientific aptitude. Young women may pursue other careers because they have low expectations for their performance in technical fields like the software (Jirout and Newcombe, 2015).

In the software profession specifically, academic women are less inclined to pursue an entrepreneurial career, while non-academic women are more likely to "fit in" by adopting views associated with masculine culture (Pogesi et al., 2020). As one researcher stated, "*Founding a technology-based firm is commonly regarded as a male domain*" (Dautzenberg, 2012, p79). Wilson and Patón-Romero (2022) conducted a systematic mapping study of papers from 2010 to 2021 on female technology entrepreneurship and highlighted the factors responsible for underrepresentation in the field. They identified several factors including family life, adopting behavior, competencies, educational perceived challenges, male behaviour, stereotypes, and role models affecting the career path of women in software entrepreneurship (Wilson and Romero, 2022). The study revealed that entrepreneurial intention is noticeably higher in male students than female students there is a link between stereotypical masculine traits and successful entrepreneurs as bold, aggressive, logical (Wilson and Patón-Romero, 2022). However, they noted that by acknowledging the status of women entrepreneurs as "other," women learn how to turn barriers into advantages and to mobilise their "otherness" by creating their own norms and accepting different ways to perform tasks instead of trying to "fit in" (Wilson and Patón-Romero, 2022). Further studies point to the lack of role models and mentors as another major obstacle (Knoll et al., 2011, Kovaleva et al., 2022).

Employing a feminist lens allows us to examine the "otherness" of women by making them the center of study. The term "feminism" aims to end women's subjugation (Jaggar, 1983, p.5), one of the fundamental principles of feminist thought is the necessity to address women's subjugation. According to feminist scholars, our understanding of entrepreneurship and growth has to undergo a fundamental



ontological transformation (Jaggar, 1983; Rosser, 2005). A thorough examination of the characteristics (e.g., exceptional support, nurturing, and empathy) and environments (primary caregiver) of women entrepreneurs who are culturally identified as feminist will be possible (Ahl and Marlow, 2012). Furthermore, there is a misconception that women are less capable or motivated to operate growing enterprises compared with their male colleagues because they are still underrepresented in the companies that are expanding quickly. *"Somehow men get to be free riders on their few growth-oriented fellow businessmen while the women are marked out as the non-growers"* (Ahl, 2003, p.225).

While considerable research has been conducted to further understand feminist epistemology and the challenges faced by women entrepreneurs, often these studies assume male norms as a given without challenge or further interrogation (Ahl and Marlow, 2012). As a result, the conclusions from such studies inadvertently tend to attribute blame for deviating from the so-called "male norm" to women, and they recommend that women fix their flaws by adjusting their behaviour to adapt to the world. These adoptions may include getting better education, having higher aspirations, and networking differently (Brush et al., 2009; Mirchandani 1999; Jennings and Brush 2013; Henry et al., 2015). However, judging women's entrepreneurial aspirations based on male norms is overstating women's otherness and, as a result, strengthens women's subordination to males in fields related to entrepreneurship (Ahl 2006; Henry et al. 2015). Ahl (2006) developed a theoretical critique of the stereotypical masculinized discourse of contemporary entrepreneurship in her powerful critique of gender. This discourse asserts that males are given privileges just because of their gender because conventional masculinity is so prevalent in the entrepreneurial industry. This is a gendered discourse, according to Ahl (2006), where males who are identified with masculinity are given preference over women who are connected with femininity.

## **2.2. Gender in India and Ireland**

While Western developed nations have been the main focus of the majority of studies on female entrepreneurship (Kumar, 2013; Sestic and Ibrahimagic, 2015), little has been written about the difficulties faced by female entrepreneurs in developing countries. India, which has a rising number of startups but is also one of the lowest-ranked nations

in the Glass Ceiling Index, a gauge of the inclusion of women in the economy, offers a particularly fascinating context in which to study this phenomenon (Pandey, 2018). Due to sociological factors like religion and caste-making, women in India experience particular difficulties to pursue entrepreneurship (Bertaux and Crable, 2007). Despite these significant obstacles, some Indian women have overcome these and achieved success in business (Kumar, 2013). However, additional research is required as women from developing countries face different challenges than women from developed countries (Shah and Saurabh, 2015).

Ireland has experienced substantial economic growth since the 1980s. According to Grimes (2003), the influx of "Information" organisations into the "High-Tech" sector has significantly fueled the "employment boom" that has accompanied Ireland's economic success. While women's employment increased exponentially during this period (Cross and O'Brien, 2004), the same has not been mirrored in women-led businesses within the Irish economy (GEM, 2003). Even after ten years, Irish women entrepreneurs continue to emphasize that they are more responsible for their families than men are, and they use their businesses to help them strike a balance between their home and professional lives (Nevins and Hamouda, 2019; Anne and Eileen, 2010). According to GEM (2018), women in Ireland have seen fewer opportunities to start a business than males, despite men being 1.5 times more likely to be entrepreneurs. These statistics raise concerns about the status of women entrepreneurs in Ireland and make it space worth exploring.

This study begins to understand women software entrepreneurs in two different countries India and Ireland. Traditionally, software business is among the fastest-expanding industries in both countries (Arora et al., 2001; O'Gorman and Kautonen, 2004). The software industry has played a key role in both countries' recent economic development, despite that women in both countries find it difficult to start their businesses which allows one to question the pattern of low participation of women in software entrepreneurship despite high education and growth of software sector in countries. This opens up the option of researching on comparative study on women entrepreneurs in India and Ireland. Subsequently, this research aims to begin to shed

light on the important topic of gender and entrepreneurship, particularly focusing on the software in India and Ireland.

### 3. Research Approach

This research investigates the barriers that affect women software entrepreneurs in India and Ireland, focusing on gender diversity and innovation. India offers a fascinating context for this research, as sociological factors like religion and caste-making make women in India face difficulties in pursuing entrepreneurship (GEM, 2022). In Ireland, traditionally women's employment has not been mirrored in women-led businesses within the Irish economy which was largely attributed to the strict marriage bar and conservative values (McCooney, 2023). To address the research objective the study uses in-depth, semi-structured interviews inspired by Trauth and Connolly (2021). There are shortcomings in past research, and some argue for qualitative, in-depth interviews with a clear feminist agenda (Martinez and Marlow, 2018). The research team contacted potential participants via LinkedIn and email. From June to August 2023, six women software entrepreneurs were interviewed based on the following inclusion criteria: co-founded or was the founder of a software entrepreneurship venture based in India or Ireland, still actively involved in the organisation, and willing to share their experiences and insights through semi-structured interviews (Table 1).

Interviewees	Business start year	Nature of Business	Country	Education
Interview 1	2019	Mentors- mentees app	India	Bachelors
Interview 2	2008	Booking hotel online	Ireland	Masters
Interview 3	2011	Account based marketing for B2B enterprises	India	Bachelors
Interview 4	2000	Logistical software developing company	Ireland	Bachelors
Interview 5	2021	Neuroscience used to support mental health online	India	Masters
Interview 6	2019	Hen parties	Ireland	Masters

**Table 1. Overview of Interviewees**

Each interview lasted approximately 60 minutes. Open-ended questions were asked of participants about their businesses, the challenges they faced as women entrepreneurs, the influence of government policies, gender/regional stereotypes, the influence of friends, family, and mentors, and their opinions on feminism and women empowerment.

Data analysis was performed using Creswell's (2007) data analysis spiral, a recurrent and iterative process of meaning creation and moving from description to interpretation. The researcher conducted a series of coding cycles to analyse the data from six interviews. In the first cycle, codes were classified based on the research question, such as motivation and gender stereotypes. The second cycle involved creating an MS Excel spreadsheet with columns. Codes formed in the first stage were sub-divided into individual factors, such as "stereotype" and "motivation." In the third cycle, NVivo software was used to extract and filter data to reach the research aim. The lead researcher discussed her preliminary findings with the team, presenting 130 codes generated by thorough data analysis. Major codes formed included male-dominated, gender stereotypes, family support, and other factors. These codes emerged as subthemes, including family background and stereotypes in Indian and Irish culture.

## **4. Findings**

This qualitative study aims to identify the factors that affect the career path of software women entrepreneurs in India and Ireland. Cultural, societal, and institutional barriers influence women's entrepreneurial journey and business performance in the software sector. The following section highlights the findings relating to career barriers.

### **4.1 Barriers**

Institutional hurdles include messages about gender roles in the family that restrict women's mobility and decision-making ability regarding their business. Societal infrastructure, lack of confidence, and male-dominated structure constitute other barriers that affect women software entrepreneurs, displayed in a coding tree (See Figure 1). Six significant barriers emerged from coding: gender stereotypes, lack of confidence, male-dominated sector, family expectation, age, and power dynamics. These barriers were further subdivided into 15 sub-themes: Feminine Traits, Irish and Indian Culture forming the sub-theme of gender stereotype, lack of skills, lack of funding, less appreciation forming the sub-theme of lack of confidence, invisible bias and lack of personal autonomy forming the sub-theme of the male-dominated sector, role as mother, daughter, and spouse falling under family expectation, married and unmarried status becoming the sub-theme of age, and power dynamics having societal roles under it. Women's perceived skills and confidence are crucial for starting

entrepreneurial ventures, and their self-belief is significant in empowering themselves. Women must constantly remind themselves that they are good enough, which can lead to a lack of trust from others.

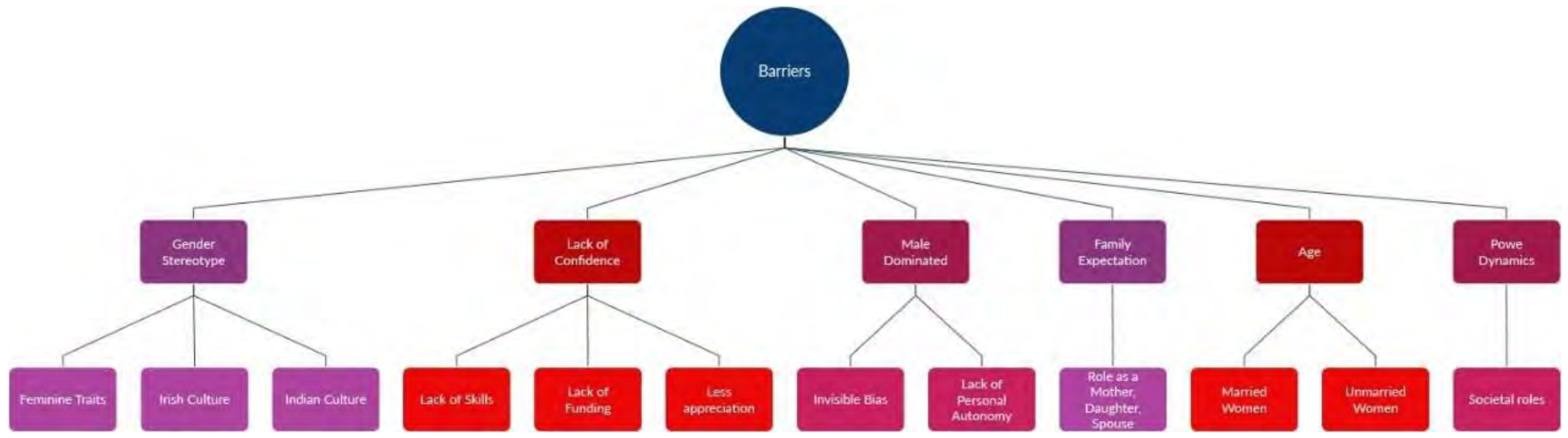


Figure 1. Barriers codes produced using NVivo

One of the most common obstacles that women entrepreneurs face is the lack of business funding, often a key driver for women starting their businesses. Women must meet requirements or provide guarantees, which may be related to a more complex phenomenon, as they must prove their worthiness and qualifications, which they believe does not happen with their male counterparts.

*“Women struggle to look for capital for their business. Women in India are often seen as dependent on men for their financial support if there is financial support if there is a property loan, but if they have their own money, it becomes easier for them.”* – Interview 3

The interviewee uses the term “dependent on men,” which leads to the question of men becoming the source and control capital for women and making them dependent on men.

The lack of female representation in the male-dominated sector of software entrepreneurship can negatively impact female entrepreneurs, creating an invisible bias and leading to feelings of self-doubt regarding their performance. Interviewees acknowledge the change in the field with changing times and the more acceptability of women leaders in IT, especially in software entrepreneurship. However, they also acknowledge the need for further improvement and more inclusivity.

*“I think software entrepreneurship is still a largely masculine domain though it is changing”*- Interview 4

Another significant effect of the male-dominated sector in IT is the lack of personal autonomy held by women. With most males around them wielding power, women may struggle to acknowledge their identity as the leader of their business. One interviewee felt that women must learn from their male counterparts, leading to whether women must become more like men to be effective in a male-dominated industry.

*“I did not call myself a CEO for 5-6 years”* – Interview 3

Research findings reveal that being a mother is a powerful identity that transcends geographical boundaries, and women face societal challenges in balancing professional advancement with the perception of women working outside the home when they

consider having children. Cultural discourses about parenthood often hinder women's potential to pursue careers, as they are often the primary caregivers of their children.

In addition to childcare, family-related responsibilities may also involve housekeeping and taking care of the extended family, which can lead to increased caregiving duties that hinder their potential to pursue entrepreneurship opportunities. Cultural norms and community beliefs play a significant role in these barriers, as Indian interviewees found that family permission to start a business becomes highly significant as their role as caretakers gets wrapped with the patriarchal norms of Indian society. For Irish interviewees, their role as a mother was important, but their role as entrepreneurs held as much relevance as their role as family members. Age also plays a significant role in their entrepreneurial journey. Indian interviewees found that childhood values had a significant impact on their careers, as they realised certain things were not expected of them due to their gender.

*“No matter how much we talk about it, some responsibility of household and household work is always more on women, even though my husband pick up a lot of the responsibility” – Interview 5*

Power dynamics among women entrepreneurs from underrepresented groups are also important, as being from a minority exacerbates the difficulties that women entrepreneurs face. In an Indian context, there are layers of power that create discrimination faced by both men and women, more so by women of their gender and social identity. Social class identity intersects with gender identity to suppress them. In contrast, another example of social class denied access to equal education for women belonging to a social class not accepted as privileged by Indian society. In an Irish context, power dynamics work differently, with diverse identities bringing diverse challenges. Irish women entrepreneurs face unique challenges such as pregnancy, being a single mom, and having twins, which they may not face as they are from India or any other ethnicity.

## **4.2 Tactics**

Nine themes emerged, including building confidence, owning personal autonomy, family support and background, mentors, feminist mindset, education impact, utilising feminine skills, government policies, and flexibility in gendered roles (See Figure 2).



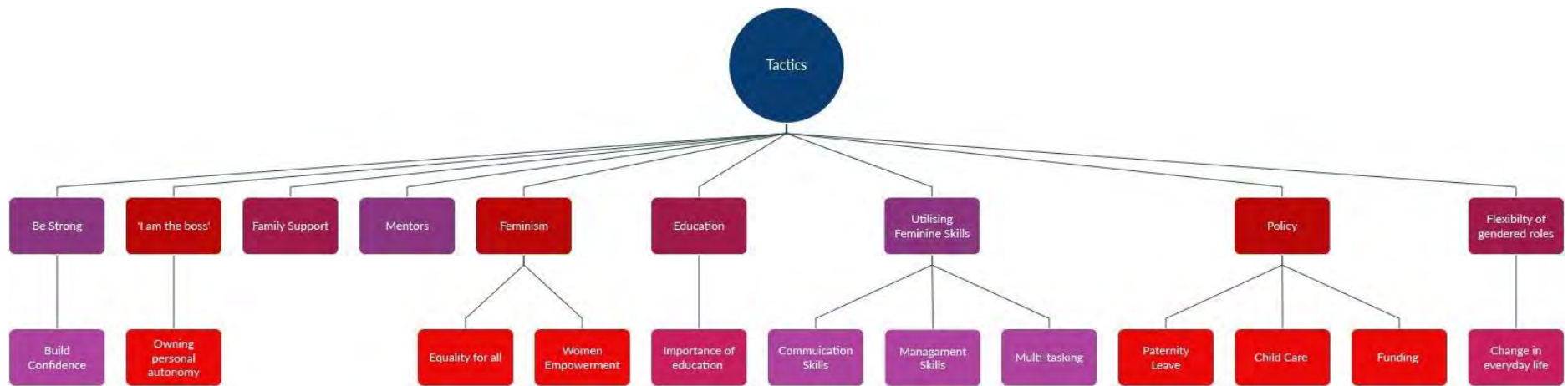
Building confidence is crucial for women entrepreneurs, as self-doubt in their skills reduces their capability in the field. Empathy and assertiveness are essential for women entrepreneurs, as they must be strong enough to face repercussions if decisions are made with emotions.

Women entrepreneurs are motivated to build a path for themselves by upholding their ideals and ambitions while refusing to submit to the norms of the sector carved out for them in a male-dominated society.

*“I think I grew up not embracing my femininity for the longest time because I recognise that I want to be ambitious”*—Interview 5

Family support is vital for women entrepreneurs, as parents want to support their children and actively encourage their daughters to break through gender boundaries while initiating cultural and infrastructure modifications. Family backgrounds in entrepreneurship also help pass entrepreneurial values to them as they grow up learning from them. A person's spouse's support and motivation have played a massive role in the successful growth of their career path.

*“My family had a major role in starting my business. When I was starting out, I needed to make connections in hotels, and as I had nothing to sell them because my grandparents had businesses in the same industry, my grandmother made personal calls to friends and set up meetings for me which helped me a lot”* – Interview 2



**Figure 2. Tactics codes produced using NVivo**

Mentors also play a significant role in the lives of women entrepreneurs, as they not only help them tackle these barriers, but also help them grow their businesses by giving them helpful advice. The influential people for women entrepreneurs include support from certain people in their firm, such as coworkers who ask other men to let them speak and hear their opinions. In an Irish context, there is no significant difference noticed, as all interviewees had male mentors. The importance of mentorship lies in the person, not their gender, as they teach them lessons through their experience, benefiting both themselves and their business.

Interviewees discussed their experiences with gender stereotypes and how they tackled these barriers using their feminine strengths. Some believed that women excelled in certain areas, such as management, communication, and discipline, while others believed in multitasking and that there were no distinct masculine or feminine skills. Interviewee 5 believed that everyone can do what they want if motivated.

*“Women are more disciplined and good in finances and management, so we can put those skills to use here better. Their support and caring nature can give different kinds of advantages to businesses if utilised wisely.” -Interview 3*

This quote from one of the interviewees represents her views on how they tackle those feminine stereotypes that divide the work into “men’s work” and “women’s work.”

The concept of feminism was also discussed, with some interviewees identifying as feminists while others did not publicly identify as feminists. Some Indian interviewees believed in equality but did not publicly identify as feminists, citing a misunderstanding. The diversity also emerged in Irish interviewees, as interviewee six did not identify as a "today's feminist" but believed in equality and equal rights for everyone. One of the interviewees shared her view on feminism.

*“People are like “Oh! She is a feminist; she does not care about equality; she cares about women.” So, I would not say that I am a feminist, but I fight for equal opportunity every time.” -Interview 6.*

Women's empowerment was another theme that emerged. Some interviewees believed that women must believe in themselves and be confident, while others believed that

financial independence and proper education are crucial steps towards empowering women. The diversity in their views on feminism and women's empowerment highlights the importance of understanding and embracing the unique strengths of women in various fields.

Government policies and statutes effectively promote change, but they often lack support for women entrepreneurs. In Ireland, there is a lack of encouragement for female entrepreneurship, and entrepreneurs believe that more significant incentives from the government should be provided. They propose policy changes around paternal leave and childcare facilities to mitigate family care responsibilities, particularly for women. In India, governments support women entrepreneurs and provide incentives to enter the field. However, policy amendments proposed differ from Irish responses, focusing on equal representation of women in policymaking. They suggest that the government consult with women business owners to understand their concerns and make more funding options available.

This study explored the experiences of female software entrepreneurs in Ireland and India. Shared issues emerged in this study, such as the need for self-confidence, identifying as a feminist and the availability of investment capital for Female entrepreneurs. Differences between Ireland and India emerged around the role of family and the nature of government policy.

## **5.0 Discussion**

Findings reveal that software entrepreneurship is a gendered phenomenon, with women entrepreneurs facing several barriers. Women entrepreneurs often struggle to balance work and family obligations, which can be challenging. They often work with their spouses to handle household chores and childcare and practice gender flexibility. Intersectionality is another significant issue for women entrepreneurs from underrepresented groups in India and Ireland. Being from a minority can exacerbate the difficulties women entrepreneurs face, as they may feel a perceived "lack of fit" in the group and may be disregarded during the entrepreneurial process. However, some women have overcome these obstacles and achieved success in entrepreneurship.

In Ireland, the church has traditionally played a role in determining women's placement in society, with, Catholic and rural cultural norms and values supporting this cultural perspective. Public regulations about marriage, reproductive rights, and employment also support this cultural perspective. It was revealed that in the Indian context, women from social class and religion face dual biases based on social identity as members of certain religions or social classes. Due to sociocultural issues like caste and religion making it difficult for them to pursue entrepreneurship, women in India confront difficulties (Bertaux and Crable, 2007). From a social feminist perspective, the intersectionality of gender, social class, and race has a dual influence on women entrepreneurs' access to resources and ability to succeed in the entrepreneurship industry.

This research highlights the importance of role models, connections, and mentors for women software entrepreneurs. However, none of the participants had female role models or mentors, which they identified as a barrier. Female mentors help women balance feminine and masculine features in their entrepreneurial identities, helping them overcome identity difficulties. Lack of confidence and difficulty in obtaining capital is another major obstacle faced by women entrepreneurs. Nieva (2015) emphasised that women business owners frequently experience a lack of self-belief, leading to hesitation when making decisions, avoidance of commitment, and a severe fear of criticism and failure. This lack of confidence is linked to external factors such as male dominance in the field and fear of failure.

In terms of social feminism, female business owners who participate in international accelerator programs encounter challenges that affect their capacity to secure funding from foreign investors, including severe competition due to the female entrepreneur's age, place of origin, financial needs, and the masculinised nature of the technology business and accelerator program (Tan, 2008). Parenthood and family expectations pose significant hurdles for women entrepreneurs, hindering their ability to start businesses. Winn (2004) highlighted the challenges women with children face in starting their businesses due to the demands of their families and childcare. Humbert and Brindley (2015) found that these expectations put women in entrepreneurship at risk and influenced their willingness to participate. Interviewees shared this perception, indicating that this is still a concern for worried women throughout Ireland and India.

Familism, a cultural ideal emphasising interdependent, loving, and sustaining family ties at the cost of individual desires and needs, impacts women software entrepreneurs. The cultural narrative that parenthood hinders women from pursuing a career in any field was presented to all women, but these messages were interpreted differently by different women. Significant individuals persuaded interviewees who were childless or had parenthood responsibilities to reinterpret the motherhood message. To navigate these barriers, women software entrepreneurs use coping mechanisms such as family support, mentors, and values like self-confidence, strength, and belief in feminism and equality. However, the lack of female role models and mentors was observed among the interviewees. This research suggests that government initiatives can help women entrepreneurs overcome obstacles by providing incentives and utilising government resources effectively.

Women entrepreneurs face challenges in networking and navigating the male dominated industry. They often rely on sisterhood and unity among women as their primary tactic to act against this male-dominated domain. Feminine stereotypical skills, such as management, communication, and multi-tasking, are used to overcome the lack of confidence in the male-dominated industry. Gender stereotypes, reinforced through socialisation during childhood and adolescence, are also influenced by childhood values taught by parents. Family support plays a vital role in women entrepreneurs' career paths, with Irish interviewees believing their family was open-minded to support them in pursuing software entrepreneurship. However, in developing nations like India, there is no distinction between the responsibilities of extremely small-scale female business owners at home and the company.

## **6.0 Conclusion**

This study investigates the factors affecting the career decisions of women software entrepreneurs in India and Ireland. It identifies several societal and cultural barriers that prevent women from engaging in entrepreneurship and the tactical response of these entrepreneurs to these barriers. The research findings align with Wilson and Patón-Romero's (2022) work, which identified factors like family life, adopting behaviour, competencies, perceived educational challenges, male behaviour, stereotypes, and role models affecting the career path of women in software entrepreneurship. The absence

of role models, social problems, institutional obstacles, gender stereotypes, lack of confidence, marginalisation, and individual factors like age, family background, and others were identified as barriers faced by women entrepreneurs. Gender stereotypes reinforced these barriers in various forms, and discrimination against women affected how they were perceived, and others largely viewed their social positions. The study also found that negative attitudes toward support varied between the two countries. In India, women were expected to fit in roles as family members and rarely received family support, while in Ireland, it was mostly a cultural construct. The adversarial culture toward marginalised social class and immigrant women was also identified as a barrier.

By asking questions about how they viewed feminism and whether they identified themselves as feminists and women's emancipation, the study reveals that women in entrepreneurship can own personalised opportunities, realise agentic potential, and create wealth and value for society. The study suggests that women should view entrepreneurship as a journey, seeking mentors and role models to help them succeed. Future research could explore the undoing or redoing of gender in entrepreneurship, examining social-cultural barriers to female entrepreneurship in Ireland and India. The socio-cultural differences between India and Ireland require further investigation. More context specific research could provide a comparative understanding of variables supporting or obstructing women entrepreneurs across different regions. Feminist theory can also be applied to investigate female entrepreneurship, and further studies on genders other than women could be pursued. The socioeconomic backdrop could also be investigated, including the economy, environmental conditions, and technological advances.

The small sample size and cross-sectional study are critical limitations of this research. Future research should focus on larger and heterogeneous samples, examining the influence of demographic characteristics on career choices. The research has shown that women have limited engagement in policymaking for entrepreneurship, which may offer some explanation for the dearth of female entrepreneurs (Shrivastava, 2021; Wu et al., 2019). The early involvement of female entrepreneurs in policymaking would be beneficial for promoting female-friendly entrepreneurial policies. Women software entrepreneurs are best positioned to empower other females through mentoring and

networking schemes. In doing so, they can support the next generation of women software entrepreneurs to navigate the existing barriers.

## References

- Adya, M., & Kaiser, K. M. (2005). Early determinants of women in the IT workforce: a model of girls' career choices. *Information Technology & People*, 18(3), 230259.
- Ahl, H. (2006). Why research on women entrepreneurs needs new directions. *Entrepreneurship: Theory and Practice*, 30(5), 595–621.
- Ahl, H., & Marlow, S. (2021). Exploring the false promise of entrepreneurship through a postfeminist critique of the enterprise policy discourse in Sweden and the UK. *Human Relations*, 74(1), 41-68.
- Anne, L. H. and Eileen, D. (2010) 'Gender, entrepreneurship and motivational factors in an Irish context', *International Journal of Gender and Entrepreneurship*, 2(2), 173–196. Doi: 10.1108/17566261011051026.
- Arora, A., Arunachalam, V.S., Asundi, J., and Fernandez, R. (2001). The Indian software services industry. *Research Policy*, 30(8), 1267 – 1287.
- Bertaux, N. and Crable, E. (2007). Learning About Women. Economic Development in Entrepreneurship and The Environment in India: A Case Study. *Journal of Developmental Entrepreneurship* 12(4): 467–478 <https://doi.org/10.1142/S1084946707000757> [accessed 30 March 2023].
- Brush, C. G. (1992). Research on women business owners: Past trends, a new perspective, and future directions. *Entrepreneurship: Theory and Practice*, 16(4), 5–30.
- Brush, C. G., de Bruin, A., & Welter, F. (2009). A gender-aware framework for women's entrepreneurship. *International Journal of Gender and Entrepreneurship*, 1(1): 8–24.
- Brush, C. G., Edelman, L. F., Manolova, T., & Welter, F. (2018). A gendered look at entrepreneurship ecosystems. *Small Business Economics*. <https://doi.org/10.1007/s11187-018-9992-9>.
- Creswell, J. W. (2007). Qualitative inquiry and research design: Choosing among five approaches. (2nd ed.). *Thousand Oaks, CA: Sage*.
- Cross, C., O'Brien, M. (2004). Shattered Dreams – The Resilient Glass Ceiling, Paper Presented at the 1st International Conference on Women's Studies Eastern Mediterranean University, Famagusta, Turkish Republic of Northern Cyprus, April.
- Dautzenberg, K. (2012). Gender differences of business owners in technology-based firms In *International Journal of Gender and Entrepreneurship* 4(1):79–98. <https://dx.doi.org/10.1108/17566261211202990> [ accessed 29 March 2023].
- Frederick, H., O'Connor, A., and Kuratko, D. F. (2016) *Entrepreneurship: Theory/Process/Practice*. 4th edn. Victoria: Cengage Learning.
- Germeijs, V. and Verschueren, K. (2006). High school students' career decisionmaking process: Development and validation of the Study Choice Task Inventory. *Journal of Career Assessment* 14:449471.



- Global Entrepreneurship Monitor (GEM) (2018) GEM Ireland 2018 report. Available at: <https://www.gemconsortium.org/report/gem-ireland-2018-report> [Accessed 09 January 2024].
- Global Entrepreneurship Monitor (2022). Global entrepreneurship monitor releases new research highlighting women's entrepreneurship trends across the globe, London, available from: <https://www.gemconsortium.org/reports/womensentrepreneurship> [accessed 01 September 2023].
- Griffith A. L. (2010). Persistence of women and minorities in STEM field majors: is it the school that matters? *Economics of Education Review* 29 (6) 911–922, Persistence of women and minorities in STEM field majors: Is it the school that matters? - ScienceDirect [ accessed 28 March 2023].
- Grimes, S. (2003). 'Ireland's Emerging Information Economy: Recent Trends and Future Prospects: *Regional Studies.*, 37(1):3-14.
- Henry, C., Foss, L., Ahl, H. (2016). Gender and entrepreneurship research: a review of methodological approaches. *International Small Business Journal*, 34(3):217–241.
- Humbert, A. L. & Brindley (2015). Challenging the concept of risk in relation to women's entrepreneurship. *Gender in Management: An International Journal*, 30(1), 2-25.
- Hyde J.S. (2014). Gender Similarities and Differences in Annual Reviews Psychology, 65(1): 373-398 Gender Similarities and Differences | Annual Review of Psychology (annualreviews.org) [accessed 28 March 2023].
- Jaggar, A. M. (1983). *Feminist politics and human nature*. Rowman & Littlefield.
- Jennings, J. E., & Brush, C. G. (2013). Research on women entrepreneurs: Challenges to (and from) the broader entrepreneurship literature? *Academy of Management Annals*, 7(1), 661–713.
- Jirout J., Newcombe N.S. (2015). Building Blocks for Developing Spatial Skills: evidence From a Large, Representative U.S. Sample, *Psychology Science*, 26 (3),302–310, <https://doi.org/10.1177/0956797614563338>
- Kelley, D. J., Baumer, B. S., Brush, C., Green, P. G., Mahdavi, M., Majbourni, M., et al. (2017). *Global entrepreneurship monitor 2016/2017 report on Women's entrepreneurship*. Babson College, Smith College, and the Global Entrepreneurship Research Association.
- Knoll S., Eisend M., Steinhagen J. (2011), Gender roles in advertising: measuring and comparing gender stereotyping on public and private TV channels in Germany, *International Journal of Advertising*, 30 (5) :867–888, <https://doi.org/10.2501/IJA-30-5-867-888> [accessed 29 March 2023].
- Kollmayer M., Schultes M. T., Schober B., Hodosi T., Spiel C. (2018). Parents' Judgments about the Desirability of Toys for Their Children In associations with Gender Role Attitudes, Gender-typing of Toys, and Demographics, *Sex Roles* 79 (5–6) 329–341, <https://doi.org/10.1007/s11199-017-0882-4> [ accessed 28 March 2023].
- Kovaleva Y., Saltan A., Happonen A. (2023). Becoming an entrepreneur: A study of factors with women from the Tech Sector in Information and Software Technology, 155(107110):1-12 *Becoming-an- entrepreneur--A-study-offactors-wit\_2023\_Information-and-Softwa.pdf* [ 29 August 2023].

- Kumar, A. (2013). Women entrepreneurs in a masculine society: Inclusive strategy for sustainable outcomes In *International Journal of Organizational Analysis*, 21(3): 373–84.
- Martinez, D., and Marlow, S. (2018). Annual review article: Is it time to rethink the gender agenda in entrepreneurship research In *International Small Business Journal: Researching Entrepreneurship*, 36(1): 3-22.
- McCooney, R. (2023). The Employment Boom in Ireland. <https://recruitonomics.com/the-employment-boom-in-ireland/>. [accessed 29 March 2023].
- Moss-Racusin, C., Phelan, J., & Rudman, L. (2010). “I’m Not Prejudiced, but...”: Compensatory Egalitarianism in the 2008 Democratic Presidential Primary. *Political Psychology*, 31(4), 543-561.
- Naidu, S., & Chand, A. (2017). National Culture, gender inequality and Women’s success in micro, small and medium enterprises. *Social Indicators Research*, 130(2), 647–664.
- Neergaard, H. and Ulhøi, J.P. (eds) (2007). *Handbook of Qualitative Research Methods in Entrepreneurship*. Cheltenham: Edward Elgar Publishing.
- Nevins, R. and Hamouda, A. (2019) ‘Generation Y females in Ireland: An insight into a new entrepreneurial phenomenon’, *Proceedings of the European Conference on Innovation & Entrepreneurship*, 716-723. 10.34190/ECIE.19.060.
- Nieva, F. (2015). Social women entrepreneurship in the Kingdom of Saudi Arabia. *Journal of Global Entrepreneurship Research*, 5(1).
- OECD. (2020). Inclusive Entrepreneurship Policies, Country Assessment Notes, <https://www.oecd.org/cfe/smes/Ireland-IE-2020.pdf>. [accessed 01 September 2023].
- O’Gorman, C and Kautonen, M. (2004). Policies to promote new knowledge-intensive industrial agglomerations. *Entrepreneurship & Regional Development*, 16 (6): 459 – 479.
- Pandey, N. K. (2018), An analysis of startup ecosystem in metropolitan city in India, *International Journal of Engineering and Management Research (IJEMR)* 8: 237–44.
- Powell, A., Dainty, A., & Bagilhole, B. (2012). Gender stereotypes among women engineering and technology students in the UK: lessons from career choice narratives. *European Journal of Engineering Education*, 37(6), 541-556.
- Robb, A.M. and Watson, J. (2012) ‘Gender differences in firm performance: evidence from new ventures in the United States’, *Journal of Business Venturing*, 27(5), 544–558.
- Rosser, S. V. (2005). Through the lenses of feminist theory: Focus on women and information technology. *Frontiers: A Journal of Women Studies*, 26(1), 1-23.
- Tan, J. (2008). Breaking the “bamboo curtain” and the “glass ceiling”: The experience of women entrepreneurs in high-tech industries in an emerging market. *Journal of Business Ethics*, 80, 547-564.
- Trauth, E.M. and Connolly, R. (2021). Investigating the Nature of Change in Factors Affecting Gender Equity in the IT Sector: A Longitudinal Study of Women in Ireland, *MIS Quarterly*, (45: 4), 2055-2100.
- Tur-Porcar, A., Mas-Tur, A., and Belso, J. A. (2017). Barriers to women entrepreneurship. *Different methods, different results? Quality & Quantity*, 51(5), 2019–2034.

- Sánchez-Escobedo, M. D., Díaz-Casero, J. C., Díaz-Aunión, A. M., & Hernández Mogollón, R. (2014). Gender analysis of entrepreneurial intentions as a function of economic development across three groups of countries. *International Entrepreneurship & Management Journal*, 10(4), 747–765.
- Shrivastava, M. (2021). “Women Tech Entrepreneurship in India.” *Women’s Entrepreneurship in STEM Disciplines: Issues and Perspectives*, edited by Michaela Mari et al., Springer International Publishing, 61–73. *Springer Link*, [https://doi.org/10.1007/978-3-030-83792-1\\_4](https://doi.org/10.1007/978-3-030-83792-1_4) [ accessed 24 September 2023]
- United Nations (UN). (2021). United Nations Sustainable Development Goals, <https://www.un.org/sustainabledevelopment/education/>. [ accessed 30 March 2023].
- Schein, V.E. (1973). A global look at psychological barriers to women's progress in management. *Journal of Social Issues*, 57, 675–688.
- Sestic, M. and Ibrahimagic, S. (2015). Business Problems in A Women’s Small Entrepreneurship–The Bosnia and Herzegovina Case of Post Conflict and Transition Context in *Economic and Social Development: Book of Proceedings*, 1: 491.
- Shah, H. and Saurabh, P. (2015). Women entrepreneurs in developing nations: Growth and replication strategies and their impact on poverty alleviation In *Technology Innovation Management Review*, 5(8): 34–43 <http://doi.org/10.22215/timreview/921> [ accessed 29 March 2023].
- Watson, J. (2002). ‘Comparing the performance of male- i and female-controlled businesses: relating outputs to inputs’, *Entrepreneurship: Theory & Practice*, 26(3), 91–100. doi:10.1177/104225870202600306.
- Wilson, A. W. and Patón-Romero, J. D. (2022). Gender equality in tech entrepreneurship: A systematic mapping study. In *Proceedings of the Third Workshop on Gender Equality, Diversity, and Inclusion in Software Engineering*, 51-58.
- Winn, J. (2004). Entrepreneurship: not an easy path to top management for women. *Women in management review*, 19(3), 143-153.
- Wu, J., Li, Y., & Zhang, D. (2019). Identifying women’s entrepreneurial barriers and empowering female entrepreneurship worldwide: a fuzzy-set QCA approach. *International Entrepreneurship and Management Journal*, 15(3), 905-928.

# Smart maintenance at offshore wind farms: A digital System of Systems approach

**David Wodak**

*Tilburg University, Tilburg, The Netherlands  
d.wodak@tilburguniversity.edu*

**Carol Ou**

*Tilburg University, Tilburg, The Netherlands  
carol.ou@tilburguniversity.edu*

## **Abstract**

*Critical infrastructures (CIs) are becoming increasingly interconnected, challenging existing governance approaches. This paper advocates a System of Systems (SoS) perspective for CI management by fostering collaboration, system alignment, and a digital ecosystem for data-sharing among stakeholders. This approach presents opportunities like efficient smart maintenance, which could help mitigate unexpected disruptions. This work-in-progress research embarks on a longitudinal case study on the collaboration among various CI managers in the context of the Digital North Sea project in Europe. Initial discussions with project stakeholders reveal significant maintenance challenges that a SoS approach could address. Further data collection and analysis will provide insights into the project's development and its potential impact on enhancing operations and contributing to the implementation of smart maintenance by CI management.*

**Keywords:** Digitalization, System of Systems, Critical Infrastructures, Smart Maintenance, IT Governance, Data-Sharing

## **1.0 Introduction**

Critical Infrastructures (CIs) often suffer from fragmentation, with CI managers working in isolated silos (Adjerid et al., 2018). Even though physical infrastructures are becoming increasingly intertwined, think about a bridge that has a road, a train track, and electricity cables on it. Although these elements of the bridge are all physically interconnected, it becomes clear that each asset component can be treated as a separate entity but all part of the same infrastructure asset. The boundaries of asset components are blurring even further due to digital advancements (Pursiainen, 2018). No longer does CI management solely rely on their own data and technology, but more often they form collaborations in which they share data and allow access to (partial) information produced by each other's technology (García de Soto et al., 2022), like sensors on roads. Even though these collaborations and data-sharing make the existing infrastructures more robust, they also increase the data interdependencies between CI managers (Rinaldi et al., 2001).

To address the complexity issue, a more holistic approach to infrastructure management is needed, where CI management can align their systems and collaborate as a unified infrastructure organization. In other words apply a digital System of Systems (SoS) approach (Maranghi et al., 2020). *A digital SoS* refers to the integration and management of interconnected digital systems, networks, and technologies that work together to achieve a shared goal (Wodak & Ou, 2023). A SoS approach promotes close collaboration among partners, aligning their systems toward a shared goal (Boardman & Sauser, 2006). This alignment blurs both the digital and organizational boundaries between partners, creating an ecosystem that enables the alignment of daily operations and data-sharing (Dimario et al., 2009).

CI management could realize benefits by adopting a SoS approach and fostering a digital ecosystem (Dawson et al., 2016). This innovative approach, which involves the integration of data and technology from various partners, unlocks new insights that were previously unattainable. A clear example of this is the implementation of smart maintenance practices for CIs, which is an organizational structure designed to oversee the maintenance of physical assets based on measurements from digital technologies, like sensors, that analyze the asset performance (Bokrantz et al., 2020a; Wang et al., 2021). By leveraging this approach, existing maintenance procedures can become more streamlined, paving the way for proactive maintenance measures. When CI management employ a digital SoS approach, they can delve into the potential for collaborative maintenance efforts. This is because this approach could enhance their ability to efficiently coordinate and manage maintenance tasks across interconnected systems. It also facilitates the attainment of synergies through shared resources and information (Mahulkar et al., 2009). This collaboration has the potential to significantly reduce maintenance expenses and minimize disruptions to CIs.

However, as digital advancements continue, SoS are becoming larger, more complex, and diverse, leading to uncertainties about their design, connections, utilizations, and required effective IT Governance structures (Gorod et al., 2008). Scholars have mainly focused on engineering aspects, often overlooking managerial concerns (Gorod et al., 2008). This leaves questions about the optimal IT Governance of SoS and the associated design. Simple systems might have preferred centralized governance (Ofe & Sandberg, 2023). But SoS, being complex and involving multiple owners with varying degrees of

freedom and connections, require careful analysis of their role and context for effective design, leading to the calls for more dynamic IT Governance mechanisms (Ofe & Sandberg, 2023; Wodak & Ou, 2023).

In short, due to digital advancements and data-sharing collaborations, the need for a holistic approach to infrastructure management is rising. That blurs digital and organizational boundaries among partners and allows for the creation of a digital ecosystem, an SoS. Therefore, the research question becomes: *How can IT Governance enhance collaborative smart maintenance in physical Critical Infrastructures through a digital System of Systems?*

To answer this question, this study focuses on the Digital North Sea project in Western Europe, where various managers of public and public-owned CIs collaborate to gather, manage, and maintain data from sensors on offshore windmills. Currently, ownership regarding data, sensors, and windmill assets is decentralized among various organizations, resulting in complex data-sharing and maintenance planning.

## **2.0 Background**

### **2.1 SoS in infrastructure management**

As indicated earlier, the boundaries between CIs are blurring, which requires a more holistic approach to infrastructure management (Maranghi et al., 2020). A digital SoS approach could be a solution. This allows for CI management to form an interconnected network that has its own capabilities the parties involved (Gopalakrishnan et al., 2022). Through seamless data-sharing, partners in this form of collaboration work together as if they are one to achieve goals they independently could not achieve (Delaurentis et al., 2017).

In theory, the SoS approach allows for various benefits. By aligning their operations, it becomes easier for CI managers to handle disruptions in their infrastructure (Hemme, 2015). For example, when crucial cargo is in transit by rail and an unexpected disruption hinders its transfer. The involved parties can collaboratively identify an alternative route, such as a highway or waterway, for seamless cargo transportation. As becomes clear via this example, these parties together have the shared goal of making sure this

cargo is transferred smoothly. The collaboration between these parties makes it possible to solve the issues and minimize the effects of the disruption in one infrastructure (Hemme, 2015).

However, even though the literature points out these potential benefits, various issues need to be solved before CI management can form these SoS. Research has indicated that aligning all the systems between the various organizations is a pivotal issue (Reynolds & Yetton, 2015). Especially since each organization has its own IT Governance mechanisms in place, simply aligning those would cause friction between partners (Adjerid et al., 2018). In addition, it is important to manage the parties in this SoS both individually and as a whole (Wodak & Ou, 2023). It is currently not clear what the right degree of autonomy is that parties can have within the SoS, to which extent coupling is required, and whether operating this way truly solves the issues CI management currently faces (Ofe & Sandberg, 2023). From the longitudinal perspective, the required IT governance mechanisms are also expected to evolve along the way of collaboration, raising additional challenges for the CI management.

## **2.2 Smart maintenance**

Smart maintenance in CI refers to an organizational design for managing the maintenance activities of physical infrastructure assets in environments with pervasive digital technologies (Bokrantz et al., 2020a). But at this moment it is a difficult endeavor due to various reasons. One of the most prominent issues is the scarcity of available data. CI managers have been gathering more data than ever before, but not all such data has proven to be useful. Not only that, but often CI management is also unaware of which data is available, which prohibits them from conducting any smart maintenance activities. The other problems are scalability issues (Bokrantz et al., 2020c). Many pilot projects focus on one CI asset, like a road or a bridge. However, expanding on these assets to a broader scope is a complex task as it requires organizations to expand their smart maintenance activities drastically, which organizations are currently unable to do (Bokrantz et al., 2020c; Shcherbakov et al., 2020).

Many of these issues can be potentially resolved if CI management adopt a SoS approach. Firstly, CI managers are no longer reliant solely on their data; instead, They

can access data from their partners, leading to a more comprehensive understanding of the infrastructure and how maintenance can be optimally conducted. This approach also allows for the consolidation of multiple maintenance activities. Secondly, scaling smart maintenance activities becomes easier as CI managers can share data (Bokrantz, et al., 2020c). This not only reduces complexity but also encourages the sharing of different perspectives on infrastructure. Through collaboration, complexities are minimized, and activities are streamlined (Adjerid et al., 2018).

### **3.0 Methodology**

This research is a work in progress. A longitudinal case study will be conducted, with the gathering and analysis of data spanning over a period of two years to answer the research question (Yin, 2009). This case study aims to explore and improve the collaborative practices of multiple organizations managing and maintaining windmills and sensors at the wind farm. The goal is to apply a SoS approach, enabling the organizations to function as a unified entity and enhance the nationwide energy transition. In addition, this research will follow the methodology based on Eisenhardt's (1989) approach, employing in-depth interviews with representatives from each organization involved in the Digital North Sea project. The interviews will explore their perspectives on collaboration, data-sharing, their vision for a unified SoS approach, as well as the inter-organizational IT governance mechanisms evolving in this project journey. Insights from these interviews will be key components for understanding existing challenges and potential solutions. Also, archival documents will be analyzed, like policy documents, contracts, and other relevant archival data. This allows for a deeper understanding of the existing collaboration mechanisms and identify any barriers that have hindered the organizations' joint efforts. Also, the minutes of previous meetings will be analyzed to acquire a historical context, understand past decisions, and recognize recurring issues or patterns. This would allow us to acquire a better overview of the development of the IT governance in place, and further highlight the dynamic changes that take place over a longer time period. Lastly, sounding boards and discussion sessions will be organized. These sessions will allow for direct observation of their interactions, communication, and decision-making processes.



Rigorous data analysis techniques will be used on the gathered data, including coding and thematic analysis. By doing so patterns, recurring themes, and crucial insights can be found in the collected data. This analysis will lead to the formulation of theoretical propositions to facilitate the development of a digital SoS. The data collected in combination with the literature available will provide the basis for interpreting the findings and providing recommendations for further research and practical applications.

#### **4.0 Discussion**

So far, preliminary discussions have taken place with the CI management involved in the Digital North Sea project. The goal of these preliminary discussions was to understand the issues which managers faced. These preliminary discussions are not yet sufficient to provide a full picture of the current situation of the Digital North Sea project. They do however highlight several central issues, related to maintenance, and how operating as one central system could potentially solve these issues. Therefore, more interviews will be conducted, sounding boards will be organized, and archival data will be collected and analyzed. Due to the longitudinal nature of this research, the first phase of this research will aim to acquire a solid overview of the as-is situation. The follow-up phases involve closely monitoring developments to understand the impact of choices on operating as a cohesive SoS and how this collaboration contributes to smart maintenance. Considering the rich information involved in this research, we also anticipate using quantitative methods to analyze the longitudinal data in the future.

#### **5.0 Contribution**

This paper proposes a SoS approach to recognize the advantages of close collaboration and the need for deeper system integration, particularly in scenarios like smart maintenance for CIs. The primary goal of this paper is to explore the initiation of a digital SoS collaboration among managers of physical CI and examine its impact on the daily operations of the involved parties. The longitudinal nature of this case study allows us to craft a theory of dynamic IT Governance for a digital SoS. Unlike previous articles we don't consider governance as a static phenomenon that is limited to a single organization, but as a cross-organizational dynamic concept that changes over time. Furthermore, we will shed light on the inherent frictions that manifest and necessitate resolution through a new IT Governance approach that no longer views organizations

in isolation but as interconnected entities. The anticipated practical outcome is to improve operational efficiency, which could translate into enhanced smart maintenance, as demonstrated in the case study, and potentially lead to better energy output, by minimizing unexpected disruptions.

## References

- Adjerid, I., Adler-Milstein, J., & Angst, C. (2018). Reducing Medicare Spending Through Electronic Health Information Exchange: The Role of Incentives and Exchange Maturity. *Information Systems Research*, 29(2), 341-361. <https://doi.org/10.1287/isre.2017.0745>
- Boardman, J., & Sauser, B. (2006). System of Systems - the meaning of of. *System of Systems Engineering*, 118-123. <https://doi.org/10.1109/SYSOSE.2006.1652284>
- Bokrantz, J., Skoogh, A., Berlin, C., & Stahre, J. (2020a). Smart maintenance: instrument development, content validation and an empirical pilot. *International Journal of Operations & Production Management*, 40(4), 481-506. <https://doi.org/10.1108/ijopm-11-2019-0746>
- Bokrantz, J., Skoogh, A., Berlin, C., Wuest, T., & Stahre, J. (2020b). Smart Maintenance: a research agenda for industrial maintenance management. *International Journal of Production Economics*, 224, 107547. <https://doi.org/https://doi.org/10.1016/j.ijpe.2019.107547>
- Bokrantz, J., Skoogh, A., Berlin, C., Wuest, T., & Stahre, J. (2020c). Smart Maintenance: an empirically grounded conceptualization. *International Journal of Production Economics*, 223, 107534. <https://doi.org/https://doi.org/10.1016/j.ijpe.2019.107534>
- Dawson, G. S., Denford, J. S., & Desouza, K. C. (2016). Governing innovation in U.S. state government: An ecosystem perspective. *The Journal of Strategic Information Systems*, 25(4), 299-318. <https://doi.org/https://doi.org/10.1016/j.jsis.2016.08.003>
- Delaurentis, D., Davendralingam, N., Marais, K., Guariniello, C., Fang, Z., & Uday, P. (2017). An SoS analytical workbench approach to architectural analysis and evolution. *INSIGHT*, 20(3), 69-73. <https://doi.org/10.1002/inst.12166>
- Dimario, M. J., Boardman, J. T., & Sauser, B. J. (2009). System of Systems Collaborative Formation. *IEEE Systems Journal*, 3(3), 360-368. <https://doi.org/10.1109/jsyst.2009.2029661>
- Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *The Academy of Management Review*, 14(4), 532-550. <https://doi.org/10.2307/258557>
- García de Soto, B., Georgescu, A., Mantha, B., Turk, Ž., Maciel, A., & Sonkor, M. (2022). Construction cybersecurity and critical infrastructure protection: New horizons for Construction 4.0. *Electronic Journal of Information Technology in Construction*, 27, 571-594. <https://doi.org/10.36680/j.itcon.2022.028>
- Gopalakrishnan, S., Matta, M., & Cavusoglu, H. (2022). The Dark Side of Technological Modularity: Opportunistic Information Hiding During Interorganizational System Adoption. *Information Systems Research*, 33(3), 1072-1092. <https://doi.org/10.1287/isre.2022.1100>

- Gorod, A., Gandhi, J., Sauser, B., & Boardman, J. (2008). Flexibility of System of Systems. *Global Journal of Flexible Systems Management*, 9, 21-31. <https://doi.org/10.1007/bf03396548>
- Hemme, K. (2015). Critical Infrastructure Protection: Maintenance is National Security. *Journal of Strategic Security*, 8, 25-39. <https://doi.org/10.5038/1944-0472.8.3S.1471>
- Mahulkar, V., McKay, S., Adams, D. E., & Chaturvedi, A. R. (2009). System-of-Systems Modeling and Simulation of a Ship Environment With Wireless and Intelligent Maintenance Technologies. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 39(6), 1255-1270. <https://doi.org/10.1109/TSMCA.2009.2025140>
- Maranghi, S., Parisi, M. L., Facchini, A., Rubino, A., Kordas, O., & Basosi, R. (2020). Integrating urban metabolism and life cycle assessment to analyse urban sustainability. *Ecological Indicators*, 112, 106074. <https://doi.org/https://doi.org/10.1016/j.ecolind.2020.106074>
- Ofe, H. A., & Sandberg, J. (2023). The emergence of digital ecosystem governance: An investigation of responses to disrupted resource control in the Swedish public transport sector. *Information Systems Journal*, 33(2), 350-384. <https://doi.org/10.1111/isj.12404>
- Pursiainen, C. (2018). Critical infrastructure resilience: A Nordic model in the making? *International Journal of Disaster Risk Reduction*, 27, 632-641. <https://doi.org/https://doi.org/10.1016/j.ijdr.2017.08.006>
- Reynolds, P., & Yetton, P. (2015). Aligning Business and IT Strategies in Multi-business Organizations. *Journal of Information Technology*, 30(2), 101-118. <https://doi.org/10.1057/jit.2015.1>
- Rinaldi, S. M., Peerenboom, J. P., & Kelly, T. K. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Systems Magazine*, 21(6), 11-25. <https://doi.org/10.1109/37.969131>
- Shcherbakov, M. V., Glotov, A. V., & Cheremisinov, S. V. (2020). Proactive and Predictive Maintenance of Cyber-Physical Systems. In (pp. 263-278). Springer International Publishing. [https://doi.org/10.1007/978-3-030-32579-4\\_21](https://doi.org/10.1007/978-3-030-32579-4_21)
- Wang, X., Wang, Y., Tao, F., & Liu, A. (2021). New Paradigm of Data-Driven Smart Customisation through Digital Twin. *Journal of Manufacturing Systems*, 58, 270-280. <https://doi.org/https://doi.org/10.1016/j.jmsy.2020.07.023>
- Wodak, D., & Ou, C. (2023). Inter-Organizational IT-Governance for a System of Systems. *AMCIS 2023 Proceedings*, 15. [https://doi.org/aisel.aisnet.org/amcis2023/sig\\_scuidt/sig\\_scuidt/15](https://doi.org/aisel.aisnet.org/amcis2023/sig_scuidt/sig_scuidt/15)
- Yin, R. K. (2009). *Case study research : Design and methods* (4th ed.). SAGE Publications.

# Data-driven Sustainable Cities in the UK: The Case of Public Space and Social Cohesion

*Research In progress*

## **Abstract**

*High quality public spaces are promoted in the UK on the ground that they facilitate social mix and contribute to cohesive communities. Although widely endorsed in both academic literature and policymaking, there is little evidence to critically evaluate how the relationship between public space and social cohesion is captured within UK policy contexts. In this work-in-progress paper, we address this gap by conducting a short review of UK policies and propose a methodology to understand the social outcomes of planning policies using crowdsourced information combining citizen science activities (participatory photo elicitation and collaborative field survey) and Volunteered Geographic Information (OpenStreetMap). The outcome of this ongoing study will fill the information gap in linking public space and social cohesion policies in the UK context and contribute to the data-driven solutions in understanding the social outcomes of planning policy.*

**Keywords:** Data-driven Sustainable Cities, Crowdsourcing Information, Citizen Science, Volunteered Geographic Information

## **1.0 Introduction**

Public spaces are the social settings where citizens' everyday life unfolds, offering opportunities for social interactions and community engagement (Cattell, Dines, Gesler, & Curtis, 2008). This is visible in the provisions of public spaces that are consistently associated with encouraging social encounters and promoting positive user experience (Zhang & Lawson, 2009; Zordan, Talamini, & Villani, 2019). The physical co-presence and conviviality captured from everyday kind of public spaces, shapes public perceptions of space quality (Carmona, 2019), all of which plays an essential role in facilitating social cohesion. In response, both public space and social cohesion have matured significantly as the strategic concepts in the UK planning policy and practice, i.e., innovations in the management of public space (De Magalhães & Carmona, 2006), continuous debate on the cohesion policy (Ratcliffe, 2012; Di Cataldo & Monastiriotis, 2020). Such efforts have often focused on devising new initiatives and presenting policy agendas that reconcile social sustainability and economic development in a way that is mutually beneficial (Cowell & Owens, 2010; Lennon, 2015). Notwithstanding the recognition of their prominence on the policy agenda, research linking planning and public policy on public space and social cohesion is

scarce due to the challenges in evaluating the social performance of physical space (Carmona, 2019) and knowledge acquisition for public participation and community involvement (Papadopoulou & Giaoutzi, 2014).

Crowdsourcing has become popular amongst researchers and institutions to enable collective intelligence and synthesize distributed knowledge in recent years (Wazny, 2017). Its potential in the policy making and evaluation process has also been increasingly recognised in the public sector (Taeihagh, 2017), especially within the planning field (Seltzer & Mahmoudi, 2013) for knowledge diffusion and citizen participation. Crowdsourcing can address traditional research challenges that are place-specific and citizen-centric, which have not been fully explored due to data limitation and technology constraints, in the form of volunteered geographic information (VGI) and citizen science to facilitate interactive citizen participation and collaboration (Jang & Kim, 2019; See, et al., 2016). Therefore, this work-in-progress paper seeks to address the information gap of linking public space and social cohesion with a short policy review proposing a crowdsourcing approach to understand the social outcomes of planning policies on public space and social cohesion in the UK.

Specifically, our research aim strives to answer the following research questions.

RQ1: What is the policy gap in linking public space and social cohesion in the UK?

RQ2: What is the strength and direction of relationships between citizens' perceptions of public space characteristics, social interactions, and social cohesion indicators?

RQ3: What is the relationship between the physical characteristics of public space and social interactions and cohesion experienced by citizens?

In this research in progress paper, we answer the RQ1 and present our methodology for future work that will help answer RQ2 and RQ3, discussing our citizen science and crowdsourcing approach that will aim to bring citizens' perspectives and experiences of public spaces.

## **2.0 UK Planning and Public Policy**

### **2.1 Public Space Policy**

Planning policies related to public space in the UK are primarily guided by the National Planning Policy Framework (NPPF) and the Planning Practice Guidance (PPG). In addition, policy documents like Town and Country Planning (2015), Neighbourhood Planning (2014), and National Model Design Code (2021) are established to provide more focused design guidance and assessment framework with high-quality design solutions and criteria, i.e., planning regulations, strategic development of local areas, and environmental impact assessment. Overall, these planning policies recognize the role of public space in local development but are often ill-equipped to understand:

- the motives and attitudes of the citizens and diverse communities towards engaging with public spaces (Aelbrecht, Stevens, & Kumar, 2021),
- the patterns and changes in how social life unfolds in urban public spaces (Mehta, 2019), and
- the quality of built outcomes, for example, how spaces are experienced, in response to different social and cultural contexts (Carmona, 2019).

## **2.2 Social Cohesion Policy**

The development of social cohesion policy in the UK emerged when “community cohesion” and “equality” became the dominant concepts in managing the cultural relationship and in response to the immigration and ethnicity agendas (Lewis & Craig, 2014). A wide range of urban policy agendas are set out to promote ethnic diversity and community cohesion. Some key policy documents include Integrated Community Strategy green paper (2018), Prevent Strategy 2011, Equality Act 2011, The Community Cohesion Fund (2022). These policies are mainly interpreted in terms of engaging with religious differences, cultural identity (Wetherell, 2007), and security and immigration discourse. In general, looking at the way centrally defined policy initiatives on the notion of community cohesion, recent debates over the policy interpretation of social cohesion slowly steer away from the discourse around solidarity and multiculturalism, becoming more aligned with the role of local communities in planning systems such as social sustainability practice in urban regeneration (Woodcraft, 2015), social integration and community engagement (Local Government Association, 2019). For example, in addition to adapting to the local social and cultural contexts, recent policy frameworks such as Integrated Community Strategy Green Paper (2018) and Levelling Up White Paper (2022) have a heavy emphasis on the

provisions of delivering accessible public infrastructure and inclusive public spaces that encourage effective community engagement and reflect local aspirations. And yet this line of policy conception and interpretation of social cohesion become somewhat diluted and undermine the policy capacity in building cohesive communities as current public policy agendas fall short of recognizing it at local scale in the light of everyday experience of the built environment (Forrest & Kearns, 2001; Head, 2019; Harris & Young, 2009; Ratcliffe, 2012).

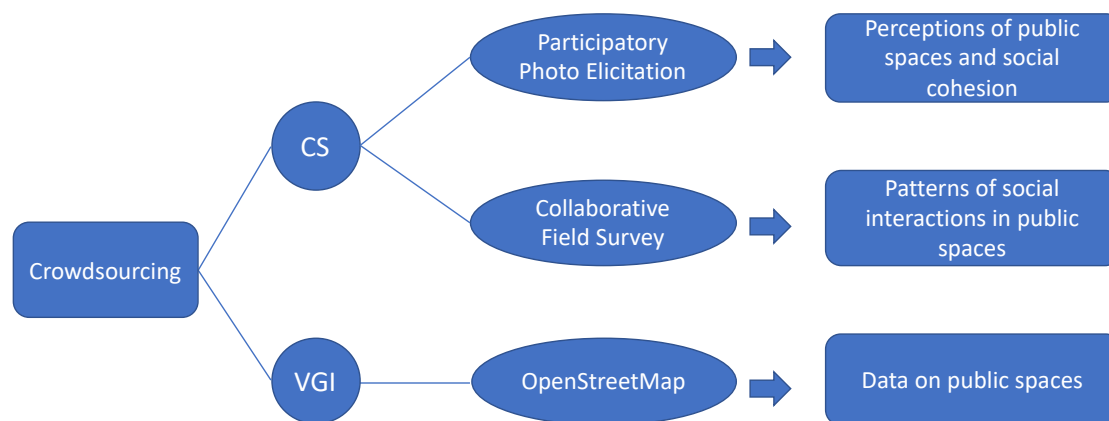
Public space often acts as the policy instrument of city development and urban regeneration to enhance liveability whilst the public policies to achieve social cohesion are largely concerned with racial, cultural, and societal challenges faced by the diverse communities and citizens in modern Britain (Ratcliffe, 2012). Together, they tend to fail to recognise the shared prosaic nature of public space and social cohesion (Amin, 2002) and to consider the spectrum of possible policy response underpinning the implications of “everyday urban” (Amin, 2002) that urban design and masterplanning add value to.

### **3.0 Data and Method**

To further understand the policy impacts on public space and social cohesion, we propose an integrative crowdsourcing approach to achieve a holistic understanding of the social outcomes of urban public spaces by applying citizen science (CS) and utilising Volunteered Geographic Information (VGI). Figure 1 demonstrates the three research methods employed and the corresponding research objectives being addressed. The participatory photo elicitation task will invite citizens to share their perceptions of public spaces, social interactions and cohesion with Google Street View images via completing structured survey questionnaires. The collaborative field survey will invite citizens to record the characteristics of social interactions in public spaces via field survey exercises using MerginMaps (a collaborative geospatial data capture platform). In addition, crowdsourced geospatial datasets will be extracted from OpenStreetMap (OSM). Participant assessment of the quality of public spaces using photo elicitation and field survey will be aligned with physical characteristics of the same locations derived from OSM data to evaluate the quality of public spaces. The findings from the

data analysis are expected to help us better understand the relationship between public spaces, social interaction and cohesion.

The novelty of this combination of methods lies in the streamlining of the data collection process by crowdsourcing the photo elicitation and field survey tasks to the citizens in addition to adopting a participatory (van Auken, Frisvoll, & Stewart, 2010) and collaborative approach. It allows for participant-driven data generation and collaborative knowledge production (Sui, Elwood, & Goodchild, 2014), but also allows for large, high quality (geospatial) datasets being created from scaling up the data collection process of traditional research methods such as photo interview and field observation in understanding indigenous experiences and local knowledge (Sobotkova, Roass, Nassif-Haynes, & Ballsun-Stanton, 2023).



**Figure 1: The research methods adopted to attain crowdsourcing information. CS – Citizen Science, VGI – Volunteered Geographic Information.**

### **3.1 Participatory Photo Elicitation**

#### 3.1.1 Study Design

A 2-minute video of how the author observes the public space sites and categorise social interactions (Table 2) will be provided for training purposes. 15 public space sites across Sheffield will be selected with 3 Google Street View links representing different images of each site will be used during the activity (Table 1). 5 images that represent 5 public space types will be assigned to each citizen to complete the survey questionnaires. The questionnaire is designed to understand their perceptions on quality of the public space, social interaction types, social relationship types, activity types,



group user group sizes, and user experiences of social cohesion by adopting 5 point likert scales.

<b>Public Open Space</b>	<b>Public Commercial Service</b>	<b>Street</b>	<b>Public Facility</b>	<b>Fourth Place</b>
Peace Gardens, Botanical Gardens, Western Bank Park	Starbucks (Western Bank Branch, Tudor Square Branch), Costa (Broomhill Branch),	Ecclesall Road, Sharrow Vale Road, West Street, Division Street	Millennium Gallery, City Library, Western Bank Museum	3 busiest bus stops, tram street corners, pedestrian precinct points, thresholds/transitional place between public space and private space around Sheffield train station within 800 metres radius.

**Table 1: 15 public spaces located in Sheffield are selected to cover 5 public space types respectively.**

### 3.1.2 Participants

A total number of 300 citizens will be recruited via Prolific to take part in the citizen science activities.

### 3.1.3 Data Analysis

Data will be analysed using structural equation modelling (SEM) technique to evaluate the moderation and mediation effects of demographic, perceptual factors, social interactions on the relationship between public spaces and social cohesion. We use the findings from this study to answer our RQ2.

## **3.2 Collaborative Field Survey**

### 3.2.1 Participants

A minimum of 30 participants will be recruited from the local nature and environment community groups in Sheffield to ensure the collaboration is motivated by citizens' shared interests and commitments to the local environments (Sheffield & Rotherham WildlifeTrust, 2023; Rotman, et al., 2012).

### 3.2.2 Study Design

The purpose of the collaborative field survey will be to map the characteristics of social interactions across multiple public spaces simultaneously drawing upon the collective intelligence of local communities (Zheng, et al., 2018). Five most popular public space sites will be identified in relation to the five typologies of public spaces (section 3.1). Participants will be provided with a training session to learn using MerginMaps and develop an understanding of how to map the social interaction by logging data points

to MerjinMaps interface. Table 2 provides a list of attributes and the corresponding information options are required to record during the field observation.

<b>Data Attribute</b>	<b>Information Options for Selection</b>
Timestamp	Automated generation
Location	Public space name
Public Space Type	Public Open Space, Public Facility, Public Commercial Service, Street, Fourth Place/Informal Public Space
Gender	Female, Male
Ethnicity	Asian, Black, White, Mixed
Life stage	Toddler, Teenager, Adult, The elderly
Social interaction type	Passive, Fleeting, Enduring
Social relationship type	Intimate, Personal, Social
Activity type	Walking, Lingerin, Stationary
Group size	1, 2, 3, 4, 5
Observation Index	The indexing number of the observation currently being made on the day.
Notes	Additional notes to describe the social interaction.

**Table 2: The attribute information required to collect during the collaborative field survey.**

### 3.2.3 Data Analysis

Spatial analysis and Exploratory data analysis (EDA) will be conducted to analyse the collected data points using QGIS to understand the characteristics of social interactions in different public space sites in addition to the temporal and spatial distribution. We use findings from this study to partially answer our RQ3.

## **3.3 OpenStreetMap**

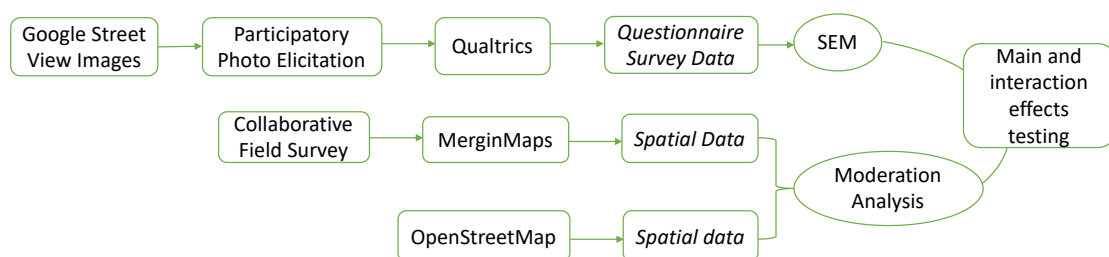
### 3.3.1 Data Collection

In seeking to understand the relationship between the quality of public spaces and the social performance of public spaces, a crowdsourced geospatial dataset downloaded from OSM will be integrated in the final stage of the research. Following the principle dimensions of built environment (“*design*”, “*density*” and “*diversity*”) developed by (Cervero & Kockelman, 1997), OpenStreetMap API will be used for data download in terms of composing the spatial layers representing the physical features of urban settings (i.e., street network, Point of Interests (POI), building footprints) that affect how the public experience urban environments.

### 3.3.2 Data Analysis

Spatial and statistical analysis will be conducted which consists of measuring the diversity of POIs (Shannon’s Evenness Index), density (street to building ratio) and design (connectivity) of urban amenity / infrastructure within the 800 m radius buffer approximately to the public space sites used in the Participatory Photo Elicitation (Kim & Hipp, 2021). The quality of the built environment in which the public spaces are situated in will be calculated by adopting a negative binomial regression approach (Kim & Hipp, 2021). Drawing upon the analysis results from the collaborative field survey (Section 3.2), moderation analysis will be conducted to evaluate the effects of demographic factors on the relationship between public space quality and social interaction characteristics.

Figure 2 describes the process by which different data sources are analysed and calibrated to assess the relationship between public space, social interaction and cohesion. This allows for a holistic understanding of the social performance of public spaces by integrating citizens’ perceptions and experiences of various characteristics typologies of public spaces, social interactions, and social cohesion. It further assesses how sociodemographic and perceptual factors interact with the physical characteristics of public space and in turn influence the level of social interactions and degree of social cohesion. We use findings from this study to answer our RQ3.



**Figure 2. Illustration of the integrative relationships of the research methods.**

## 4.0 Conclusion

This study provides an overview of the UK planning and public policy on public space and social cohesion. Our finding suggests a weak connection between policy provisions of public space and social cohesion. There is a lack of understanding of the social outcomes of urban space design and planning and limited information in exploring how social cohesion unfolds in public spaces adapting different socio-economic and spatial

contexts within UK policy contexts. To fill this information gap, in this research in progress paper, we propose our data-driven approach utilising crowdsourced information to evaluate the quality of urban built environments and assess citizens' experiences and perceptions of social interaction and cohesion in public spaces. This ongoing research aims to fill the information gap of understanding the social outcomes of urban planning and the impacts of planning and public policies. The findings are expected to inform policymaking in areas of public space and social cohesion whilst adding to our current knowledge of developing data-driven sustainable cities in terms of exploring the various means urban environments act as a medium to achieve social cohesion in the UK.

## Reference

- Aelbrecht, P. (2016). 'Fourth places': the contemporary public settings for informal social interaction among strangers. *Journal of Urban Design*, 124-152.
- Aelbrecht, P., & Stevens, Q. (2019). *Public space design and social cohesion: an international comparison*. New York: Routledge, Taylor and Francis Group.
- Aelbrecht, P., Stevens, Q., & Kumar, S. (2021). European public space projects with social cohesion in mind: symbolic, programmatic and minimalist approaches. *European Planning Studies*, 1-31.
- Amin, A. (2002). Ethnicity and the multicultural city: Living with diversity. *Environment and Planning A*, 959-980.
- Ballet, J., Bazin, D., & Mahieu, F. R. (2020). A policy framework for social sustainability: Social cohesion, equity and safety. *Sustainable Development*, 1388-1394.
- Brandon, P. S., & Patrizia, L. (2010). A Proposed Framework for Evaluating Sustainable Development. In P. S. Brandon, & L. Patrizia, *Evaluating Sustainable Development in the Built Environment* (pp. 122-146). Incorporated: John Wiley & Sons.
- Carmona, M. (2010). Contemporary public space, part two: Classification. *Journal of Urban Design*, 157-173.

- Carmona, M. (2018). The formal and informal tools of design governance. *Journal of Urban Design*, 1-36.
- Carmona, M. (2019). Principles for public space design, planning to do better. *Urban Design International*, 47-59.
- Carmona, M. (2019). Place value: place quality and its impact on health, social, economic and environmental outcomes. *Journal of Urban Design*, 1-48.
- Cattell, V., Dines, N., Gesler, W., & Curtis, S. (2008). Mingling, observing, and lingering: Everyday public spaces and their implications for well-being and social relations. *Health and Place*, 544-561.
- Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 199-219.
- Cowden, S., & Singh, G. (2017). Community cohesion, communitarianism and neoliberalism. *Critical Social Policy*, 268-286.
- Cowell, R., & Owens, S. (2010). Revisiting ... Governing space: Planning reform and the politics of sustainability. *Environment and Planning C: Government and Policy*, 952-957.
- De Magalhães, C., & Carmona, M. (2006). Innovations in the management of public space: Reshaping and refocusing governance. *Planning Theory and Practice*, 289-303.
- Department for Levelling Up, Housing and Communities. (2023, 10 23). *National Planning Policy Framework*. Retrieved from GOV.UK: <https://www.gov.uk/government/publications/national-planning-policy-framework--2>
- Di Cataldo, M., & Monastiriotis, V. (2020). Regional needs, regional targeting and regional growth: an assessment of EU Cohesion Policy in UK regions. *Regional Studies*, 35-47.
- Foreign, Commonwealth & Development Office. (2023, 10 23). *Policy Paper: Statement on freedom of religion or belief and civil society*. Retrieved from GOV.UK: <https://www.gov.uk/government/publications/freedom-of-religion-or-belief-and-civil-society-statement-at-the-international-ministerial-conference-2022/statement-on-freedom-of-religion-or-belief-and-civil-society>
- Forrest, R., & Kearns, A. (2001). Social cohesion, social capital and the neighbourhood. *Urban Studies*, 2125-2143.

- Forrest, R., & Kearns, A. (2001). Social cohesion, social capital and the neighbourhood. *Urban Studies*, 2125-2143.
- Francis, J., Giles-Corti, B., Wood, L., & Knuiaman, M. (2012). Creating sense of community: The role of public space. *Journal of Environmental Psychology*, 401-409.
- Harris, M., & Young, P. (2009). Developing community and social cohesion through grassroots bridge-building: an exploration. *Policy & Politics*, 517-534.
- Harris, N., & Thomas, H. (2004). Planning for a diverse society? A review of the UK government's Planning Policy Guidance. *Town Planning Review*, 473-500.
- Head, B. W. (2019). Forty years of wicked problems literature: forging closer links to policy studies. *Policy and Society*, 180–197.
- Hickman, M. J., & Mai, N. (2015). Migration and Social Cohesion: Appraising the Resilience of Place in London. *Population, Space and Place*, 421-432.
- Hislopa, M., Scottb, A. J., & Corbetta, A. (2019). What Does Good Green Infrastructure Planning Policy Look Like? Developing and Testing a Policy Assessment Tool Within Central Scotland UK. *Planning Theory & Practice*, 633-655.
- Jang, K. M., & Kim, Y. (2019). Crowd-sourced cognitive mapping: A new way of displaying people's cognitive perception of urban space. *PLoS ONE*, 1-18.
- Jiang, Y., Gu, P., Cao, Z., & Chen, Y. (2020). Impact of Transit-Oriented Development on Residential Property Values around Urban Rail Stations. *Transportation Research Record*, 362-372.
- Kearns, A., & Forrest, R. (2000). Social Cohesion and Multilevel Urban Governance. *Urban Studies*, 995-1017.
- Kim, Y. A., & Hipp, J. R. (2021). Density, diversity, and design: Three measures of the built environment and the spatial patterns of crime in street segments. *Journal of Criminal Justice*, 1-13.
- Kotzebue, J. R. (2016). The EU integrated urban development policy: managing complex processes in dynamic places. *European Planning Studies*, 1098-1117.
- Lennon, M. (2015). Green infrastructure and planning policy: a critical assessment. *Local Environment*, 957–980.
- Lewis, H., & Craig, G. (2014). 'Multiculturalism is never talked about': Community cohesion and local policy contradictions in England. *Policy and politics*, 21-38.
- Local Government Association. (2023, 10 23). *Public Spaces Protection Orders: Guidance for Councils*. Retrieved from Local Government Association:

[https://www.local.gov.uk/sites/default/files/documents/10.21%20PSPO%20guidance\\_06\\_1.pdf](https://www.local.gov.uk/sites/default/files/documents/10.21%20PSPO%20guidance_06_1.pdf)

- Local Government Association. (2019). *Building Cohesive Communities: An LGA guide*. London: Local Government Association.
- Lofland, L. H. (1998). *The Public Realm: Exploring the City's Quintessential Social Territory*. Hawthorne NY: Aldine de Gruyter.
- Mateo-Babiano, I. B. (2012). Public life in Bangkok's urban spaces. *Habitat International*, 452-461.
- Ministry of Housing, Communities & Local Government. (2023, 10 23). *Collection: Planning Practice Guidance*. Retrieved from GOV.UK: <https://www.gov.uk/government/collections/planning-practice-guidance>
- Natural England. (2023, 10 23). *Introduction to the Green Infrastructure Framework*. Retrieved from Natural England: <https://designatedsites.naturalengland.org.uk/GreenInfrastructure/Home.aspx>
- Oberndorfer, M., Dorner, T. E., Leyland, A. H., Grabovac, I., Schober, T., Šramek, L., & Bilger, M. (2022). The challenges of measuring social cohesion in public health research: A systematic review and econometric meta-analysis. *SSM - Population Health*, 1-14.
- Papadopoulou, C., & Giaoutzi, M. (2014). Crowdsourcing as a Tool for Knowledge Acquisition in Spatial Planning. *Future Internet*, 109-125.
- Paramita, S. A., Yamazaki, C., Hilfi, L., Sunjaya, D. K., & Koyama, H. (2021). Social cohesion and quality of life in Bandung: A cross sectional study. *PLoS ONE*, 1-11.
- Ratcliffe, P. (2012). 'Community cohesion': Reflections on a flawed paradigm. *Critical Social Policy*, 262-281.
- Ratcliffe, P. (2012). 'Community cohesion': Reflections on a flawed paradigm. *Critical Social Policy*, 262-281.
- Robinson, D. (2005). The Search for Community Cohesion: Key Themes and Dominant Concepts of the Public Policy Agenda. *Urban Studies*, 1411-1427.
- Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., . . . Jacobs, D. (2012). Dynamic Changes in Motivation in Collaborative Citizen-Science Projects. *Session: Civic and Community Engagement*, 217-226.
- Schiefer, D., & van der Noll, J. (2017). The Essentials of Social Cohesion: A Literature Review. *Social Indicators Research*, 579-603.

- See, L., Mooney, P., Foody, G., Bastin, L., Comber, A., Estima, J., . . . Rutzinger, M. (2016). Crowdsourcing, citizen science or volunteered geographic information? The current state of crowdsourced geographic information. *ISPRS International Journal of Geo-Information*, 1-23.
- Seltzer, E., & Mahmoudi, D. (2013). Citizen Participation, Open Innovation, and Crowdsourcing: Challenges and Opportunities for Planning. *Journal of Planning Literature*, 3-18.
- Sheffield & Rotherham WildlifeTrust. (2023, 10 30). *Local Groups in Sheffield and Rotherham*. Retrieved from Sheffield & Rotherham WildlifeTrust: <https://www.wildsheffield.com/nr-toolkit/looking-after-your-local-green-space/local-groups/>
- Soares, I., Weitkamp, G., & Yamu, C. (2020). Public spaces as knowledgescapes: Understanding the relationship between the built environment and creative encounters at dutch university campuses and science parks. *International Journal of Environmental Research and Public Health*, 1-30.
- Sobotkova, A., Roass, S. A., Nassif-Haynes, C., & Ballsun-Stanton, B. (2023). Creating large, high-quality geospatial datasets from historical maps using novice volunteers. *Applied Geography*, 1-13.
- Sui, D., Elwood, S., & Goodchild, M. (2014). Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In M. Haklay, *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice* (pp. 105-122). Springer Dordrecht.
- Taeihagh, A. (2017). Crowdsourcing: a new tool for policy-making? *Policy Sciences*, 629-647.
- van Auken, P. M., Frisvoll, S. J., & Stewart, S. I. (2010). Visualising community: Using participant-driven photo-elicitation for research and application. *Local Environment*, 373-388.
- Vasta, E. (2012). The controllability of difference: Social cohesion and the new politics of solidarity. *Ethnicities*, 262-281.
- Wazny, K. (2017). "Crowdsourcing" ten years in: A review. *Journal of Global Health*, 1-13.
- Wetherell, M. (2007). Community cohesion and identity dynamics: dilemmas and challenges. In M. Wetherell, M. Lafleche, & R. Berkeley, *Identity, Ethnic Diversity and Community Cohesion* (pp. 1-14). London: Sage.



- Wickes, R., Zahnow, R., Corcoran, J., & Hipp, J. R. (2019). Neighbourhood social conduits and resident social cohesion. *Urban Studies*, 226-248.
- Wong, C., & Watkins, C. (2009). Conceptualising spatial planning outcomes: Towards an integrative measurement framework. *Town Plan. Rev.* , 481–516.
- Woodcraft, S. (2015). Understanding and measuring social sustainability. *Journal of Urban Regeneration & Renewal*, 133-144.
- Zhang, W., & Lawson, G. (2009). Meeting and greeting: Activities in public outdoor spaces outside high-density urban residential communities. *Urban Design International*, 207-214.
- Zheng, F., Tao, R., Maier, H. R., See, L., Savic, D., Zhang, T., . . . Popescu, I. (2018). Crowdsourcing Methods for Data Collection in Geophysics: State of the Art, Issues, and Future Directions. *Reviews of Geophysics*, 698-740.
- Zhu, Y. (2015). Toward community engagement: Can the built environment help? Grassroots participation and communal space in Chinese urban communities. *Habitat International*, 44 - 53.
- Zordan, M., Talamini, G., & Villani, C. (2019). The association between ground floor features and public open space face-to-face interactions: Evidence from nantou village, shenzhen. *International Journal of Environmental Research and Public Health*, 1-19.

# Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs

Maria Kutar, Marie Griffiths, Tony Syme, Subrahmaniam Krishnan-Harihara,  
Aadya Bahl, Toluwanimi Ojutiku  
*Salford Business School, University of Salford, UK*

## Abstract

*The UK has suffered declining growth in productivity over many decades and there have been several attempts to study the reason for this declining growth in UK productivity. Various potential reasons including the transition to lower productivity sectors and the lack of sufficient high-skilled workers have been suggested to explain the productivity puzzle. One factor affecting productivity is firm investment in technology, an area where the evidence remains decidedly mixed. This paper presents the background to an ongoing project investigating the link between technology adoption, decision making and in-firm productivity. The project proposes to pair organisations with high productivity with organisations with low productivity and seeks to find out whether mentoring between senior managements teams from the high productivity category can yield a positive outcome for organisations with low productivity. The researchers are currently recruiting firms to form the mentor-mentee pairs and the results will be presented at the conference.*

**Keywords:** Technology adoption, digital adoption, productivity, productivity puzzle, technology leadership,

## 1. Introduction

The UK's productivity puzzle – the declining growth in productivity - remains unsolved over many decades and continues to be a conundrum for the foreseeable future. This paper, unapologetically, contributes to the growing volume of research and analysis on productivity. Conversely, we make the observation, that this important aspect of business growth, has sadly dropped in favour from Information System publications. Information System scholars such as Lucas (1999) and Pinsonneault and Rivard (1999) and Dos Santos and Sussman (2000) who have historically contributed to this area, maybe finding a more natural home in Operational Research journals? Considering the widely held perception of the relationship between productivity and IT investment as a crucial factor for economic development (Schreyer, 2001), we argue that this gap warrants further attention albeit in a different study. Scholars and policymakers alike have sought to understand the

determinants of productivity (Venturini, 2015), particularly in the context of the United Kingdom. With technology advancing faster than ever before, there has been an increased interest in understanding its impact on productivity. However, the relationship between technology and productivity is not straightforward and it these dynamics, specifically, that is the focus of our study. Despite the longstanding perception that technological innovation should lead to enhanced productivity (Brynjolfsson and Petropoulos 2021), the evidence to support this is mixed. In fact, the term ‘productivity paradox’ (also referred to as ‘Solow’s paradox’) emerged following studies that suggested that there was no significant effect on firm performance following investment in information technology (Loveman, 1994; Morrison, 1997; Barua, Kriebel and Mukhopadhyay, 1995). However, we argue that recent research proposes that the relationship between technology investment and productivity is more nuanced – it is not simply that financial investment will or will not lead to productivity gain, rather the organisational context including management and innovation approaches can influence the impact of that investment.

This project has been funded by The Productivity Institute, via UK Research and Innovation (UKRI). The study aims to understand the differences in approach to technology investment between SME organisations with high productivity and organisations with low productivity, and whether mentoring between senior managements teams from the high productivity category can yield a positive outcome for organisations with low productivity. The long-term goal is to improve strategic decision-making of SME management in relation to technology investment, enabling related productivity gains to be realised.

We have framed the following research questions:

- What are the organisational factors that contribute to technology investment decisions that improve productivity in organisations?
- How can this knowledge and expertise be captured and be shared with other SMEs?

The questions are addressed through a 6-month project to pilot a mentoring scheme for SMEs in a UK location recognised as having low productivity. The project firstly defines a firm selection methodology which assesses the productivity of SMEs, then closely oversees a clearly defined mentoring scheme, which is then evaluated. The outputs and evaluation of the mentoring scheme are used to gain an understanding of the organisational factors impacting technology investment decision making, and how

this influences firm level productivity. In the long term, the longitudinal data can be examined to identify whether there is an impact on firm productivity.

In this paper we set out the background to the research and provide an overview of the firm selection methodology and mentoring scheme. We explain the expected outputs which will be realised during project delivery. The project is currently in progress. It will be completed in March 2024; results will be presented at the conference. A literature review with two strands follows. The first strand takes the productivity puzzle as its theme, looking at economic growth and necessary skillsets. The second strand explores the nuanced relationship between technology investments and productivity. A full explanation of the project and associated methodology is provided, as are the project deliverables, but actual results will be presented at the conference following project completion.

## **2. Literature Review**

### **2.1. The Productivity Puzzle**

The “productivity puzzle” in the UK, which refers to a declining growth in productivity, has long intrigued researchers, with many theorising the reasons behind its decline. Remes, Mischke and Krishnan (2018) studied the trends in productivity, noting three major waves, with “digitalisation” being the most recent one. They stated that the benefits of the third wave have not fully materialised due to adoption barriers, lag effects and transition costs, and identified a shift to relatively low-productivity sectors such as services, and the public sector in areas of healthcare and education. Identified top sectors that also contributed to declining productivity in the UK were manufacturing, ICT, and finance and insurance. Coyle and Mei (2023), too, reached a similar conclusion by decomposing growth into contributions from different subsectors, and sectors, rather than adopting a firm-level perspective. Using the Tornqvist framework, they used data from the Office of National Statistics (ONS) to measure changes in productivity growth between 1998 and 2019. The findings attributed the productivity growth slowdown to transport equipment and pharmaceuticals within the manufacturing sector and to the computer software and telecommunication sub-sectors within ICT.

The UK Government, Skills and Productivity Board in the UK (which was dissolved in 2022, becoming Unit for Future Skills UFS) also proposed that productivity varies across a wide range of economic sectors, and that some regions may perform badly due to their specialisation in low productivity industries. The Levelling Up White Paper (HM Government, 2022), supporting this finding about regional productivity disparity, suggested six “capitals” that drove this including human capital (alongside social, financial, physical, institutional and intangible). McCann (2018) further added low levels of innovation and knowledge diffusion as causes of declining productivity, citing inadequate management and workplace relations as important contributing factors.

## **2.2. Skills and Qualifications**

When we investigate empirical literature around skills exclusively, there is an emerging theme around job polarisation and a skills mismatch. For the former theme, Acemoglu and Autor (2011) examine the relationship between technological change, the nature of tasks performed in the workplace, and the demand for different types of skills. The analysis primarily focuses on the American economy but also references countries in Europe, including the UK. The study found that technological change has a polarizing effect on the job market, as skill demand for routine tasks, both manual and cognitive in nature, is declining due to automation. In contrast, the demand for non-routine jobs, including knowledge roles that involve problem-solving and creative activities, such as professional, managerial, technical, and creative occupations, has increased. This shift in demand has favoured workers with higher levels of education and skills, leading to increasing wage inequality. The rise in skill-based technological change has contributed to this wage disparity, as the demand for skilled workers has outpaced the supply. However, the study also notes that there is a growing skill demand in non-routine manual roles that require interpersonal and environmental adaptability, such as food preparation and service, home health aides, and ground cleaning and maintenance. These jobs generally do not require extensive formal education beyond high school qualifications or extensive training, but necessitate adaptability and responsiveness that are challenging to automate or offshore. The combination of these effects has resulted in a hollowing out of middle-skill jobs, pushing labour into two polarising extremes. This shift has fostered employment growth in both high-wage, high-skill occupations and low-wage, low-

skill occupations. Cavaglia and Etheridge (2020) use price wage changes to draw the same conclusion, adding that the difference between price and average salary changes is the largest in high-skill non-routine occupations.

Montresor (2019) corroborates the claim that hollowing out for middle-skills jobs is due to growing technology; however, the paper attributes the growth of high-skill non-routine jobs to an increase in the number of graduates as opposed to technology. It finds that the rising supply of graduates intensifies competition for jobs along the employment distribution. Thus, middle-skilled individuals who lose their jobs, are more prone to finding a lower-skilled job, however, this loss can also reduce the chance lower-skilled workers have of climbing up the ladder. Multiple Organisation for Economic Co-operation and Development (OECD) working papers have further analysed the impact of skills and qualifications mismatch on productivity (Desjardins and Rubenson, 2011; McGowan, and Andrews, 2015). It suggests that qualifications alone may not reflect a skills mismatch, as they do not account for differences in the quality and orientation of education, additional training, and skills gained or lost beyond formal qualifications. The paper highlights the importance of considering the relationship between skills and productivity. The theoretical underpinnings of the research suggest indirect and direct channels through which the skills mismatch affects productivity.

According to other researchers like Green and Zhu (2010) and Battu, Belfield and Sloane (1999), overqualified or overskilled workers would have an incentive to move on to a job that better reflects their education and experience, suggesting that they experience lower job satisfaction. Lower job satisfaction would then lead to decreased job effort, higher absenteeism, and lower productivity. This can lead to a higher job turnover where over-skilled workers are more likely to change jobs, and less likely to take part in training compared to their well-matched workers with the same qualifications. Their review of the literature also suggests that skills shortages reported in various industry-level studies lower productivity growth, technological adoption, and investment drastically.

While previous research has primarily focused on within-firm productivity improvements, one paper takes a broader perspective and examines productivity at the economy-wide level. Desjardins and Rubenson's (2011) findings indicate that over-skilling within a given firm can harm overall productivity, as more productive firms may struggle to expand due to a lack of suitable labour. The skill level was measured

using a self-assessment method limited to literacy and numeracy. However, it does give us some insight into the relationship between skills and productivity. Desjardins and Rubenson (2011) use regression analysis to investigate the impact of the skills mismatch on productivity, controlling for country and industry fixed effects. The results show that over-skilled workers have a negative and statistically significant impact on overall productivity. Likewise, underqualified workers have a negative and statistically significant impact on within-firm productivity. Over-qualified and under-skilled workers, however, do not have a statistically significant impact on either. Furthermore, qualification mismatch has an inverse significant relationship with overall productivity, while skills mismatch does not. When controls for the overlaps between the components of qualification and skills mismatch are considered, the paper shows the following as having statistically significant and negative impact on overall productivity: overqualified and underskilled, underqualified and well-matched in skill, and well-matched in qualification but overskilled. However, underskilled and underqualified workers reduce within-firm productivity, while over-qualified and overskilled workers increase within-firm productivity. The paper also focuses on allocative efficiency, where workers who were well matched in terms of qualifications but over-skilled, have a negative relationship with allocative efficiency. This implies that being over-skilled alone may have a greater effect on productivity than just being over-qualified. Further, the research suggests that the strong association between under-qualification and within-firm productivity is explained by differences in managerial quality. However, the paper recommends focusing on policy factors that promote efficient reallocation of labour to tackle overall productivity. This includes improving residential mobility, and most importantly, investing in targeted training programmes.

Overall, these insights emphasize the complexity of the productivity puzzle in the UK and the multifaceted factors contributing to its decline. Addressing productivity challenges requires a comprehensive approach that considers sector-specific dynamics, regional disparities, job polarisation, and the role of skills and qualifications. Further, they reemphasise the importance of studying productivity at a local level.

### **2.3. Investment in technology and firm productivity**

A large body of research delves into the relationship between technology diffusion and economic growth and productivity (Mithas and Lucas (2014)). The growth model first explored by Solow (1956), used data from the US to conclude that about four-fifths of the growth in output per worker was attributable to technical progress. Romer (1990), too, contended that there is a close relationship between economic development and technical progress. While higher levels of input can lead to the bolstering of the economy, productivity improvements will be scant or non-existent in the absence of innovation. Technical innovation can lead to increasing returns to scale as the average cost of the infrastructure around it declines with an increase in uptake. Thus, the possibility of higher returns on investment is created by widespread technology dissemination (Arthur, 1996).

Productivity can be measured both at the aggregate level and at the firm level. While studies at the aggregate level point towards a growth in productivity due to investment in technology, results from firm-level studies have been inconsistent, varying based on the model specifications used, time periods covered and industries reviewed (Stiroh, 2002). In fact, as several studies undertaken in the past did not find evidence of a significant relationship between investment in technology and firm productivity, the term 'productivity paradox' was coined. For example, most studies based on the manufacturing sector suggested that information technology does not have any significant effect on firm performance (Loveman, 1994; Morrison, 1997; Barua, Kriebel and Mukhopadhyay, 1995). However, Dasgupta, Sarkis and Talluri (1999) undertook a similar analysis using different underlying assumptions. While a large number of studies assumed constant returns to scale, they tested under both constant and non-constant returns to scale assumptions and found that investment in technology in the manufacturing sector has a negative impact on firm productivity. Further, their findings for the service sector showed that investment in technology either had no effect, or a negative effect on productivity within firms. This result matched empirical literature which suggested the same.

More recent studies, on the other hand, have found that spending on information technology has a significant positive impact on firm productivity. Brynjolfsson and Hitt (1998) stated that productivity growth has historically stemmed from technological advancements including the steam engine and electricity. Using a firm effects model, they found that half of the realised benefits from IT investments were



due to characteristics unique to the firm, suggesting that organisational factors impact productivity greatly.

Building on this idea, Brynjolfsson and Hitt (2000) suggest that organisational transformation contributes to higher in-firm productivity in two ways. Investments in organisational practices including business processes are said to complement technology investments, which lead to improved outcomes, lower costs, and positive changes among intangible aspects of the output, and ultimately, improved productivity. According to Milgrom and Roberts (1990), for businesses to be successful, computer adoption must be a part of a "system" or "cluster" of organisational changes that reinforce one another. Any benefits of computerisation are far outweighed by adverse interactions with current organisational practises, so incremental change, either by investing in computers without implementing organisational change or only partially implementing some organisational changes, can result in significant productivity losses (Brynjolfsson, Renshaw, and Van Alstyne, 1996).

Research based on intra-firm level data from the UK, too, states that both investment and productivity have a favourable association with management and leadership quality (Ollivaud, Guillemette and Turner, 2016). Between the UK and its overseas counterparts, there are also glaring evidence-based inequalities in management and leadership quality and investment levels (Bender et al., 2016; Bloom, Sadun and Reenen, 2012; Bloom and Van Reenen, 2006). Thus, it becomes increasingly important for firms to have the right organisational setups and business practices to ensure that they are investing in innovation in a way to that helps them best realise the productivity gains. For instance, there is evidence of more positive innovation outcomes when firms invest in acquiring knowledge from other organisations (Klueter, Monteiro, and Dunlap 2017). Studies further display the high value of collaborating with other organisations by improving knowledge-sharing, upgrading innovative quality, and improving the managing or structuring of external collaborations (Driffield et al, 2021). These workplace practices impacting productivity positively involve work teams, training in multiple jobs, and flexible job assignments (Ichniowski, Shaw and Prensushi, 1995).

Many researchers believe that investment in and the adoption of digital technologies have a positive and significant impact on productivity (Cardona, Kretschmer Strobel, 2013; Cusolito, Lederman and Peña, 2020; Tastan and Gonel, 2020; Lopez, 2023).

Across sectors, the adoption of digital technologies has contributed to increasing efficiency and firm productivity. Cardona, Kretschmer Strobel (2013) conducted a literature survey on ICT and productivity and concluded that there is strong evidence for productivity enhancement resulting from ICT adoption. Cusolito, Lederman and Peña (2020) did a study on various developing countries and found that the adoption of technologies like email, business websites and subsequent integration between IT systems can help firms improve productivity not least because they could reduce production costs while also providing opportunities to expand their market although they argue that the adoption of different forms of technology is associated with varying levels of productivity gains. Moreover, not all of the productivity uplift is directly attributable to the adoption of technology itself (Boothby, Dufor and Tang, 2010) and the efficiency gains are at best only partially attributable to the adoption of general technologies such as email and websites, or even specialised systems such as ERP and CRM. One possible reason for firms being able to improve their productivity after the adoption of digital technologies is that technology plays an enabling role (Cusolito, Lederman and Peña, 2020). The results of a largescale study conducted by Gal et al (2019) also showed that digital adoption is strongly associated to productivity gains. When digital technologies are adopted, firms often make a shift towards capital intensive production practices, create new products and services and attract workers with more skills, all of which collectively contribute to improving efficiency (Boothby, Dufor and Tang, 2010; Tastan and Gonel, 2020).

These findings find resonance amongst other researchers (Lopez, 2023; Mosiashvili and Preussen, 2020), who have stressed the importance of 'complementarity', which is an important theme in productivity research. Complimentarity is the notion that mere adoption of technology does not itself boost productivity significantly; rather, it is the addition of complementing factors yield the most productivity benefits. Such complementary factors could be the use of technologies which complement each other or the provision of skills or training for staff when new technology is used. The UK's Office for National Statistics (ONS, 2018) did a review of the link between the information and communication technology and found that frequent use of technology that adopting complementary technologies and the intense use of the technologies that the firm has invested in are likely to afford the highest productivity premium. The ONS (ibid) also found that enabling factors for technology use, for example, the availability of high-speed internet also had a positive impact of firm productivity.

Other researchers have also explored the benefits of complementarity. Lopez (2023) analysed the importance of improving organisational practices along with technology adoption while Boothby, Dufor and Tang (2010) and Tastan and Gonal (2020) highlight the importance of training and skills development for employees. This could be because ICT adoption changes the nature of skills needed to work within the organisation and there is a need for workers with higher level skills. In their research, Gal et al (2019) found that although digital technology adoption does improve productivity, the results are weaker in the presence of skill shortages within the firm. Where firms invested in complementarities between digital technologies and skills development, they gained the most. Robinson, Siegel and Liao (2021) conducted a survey of SMEs based in Kent and concluded that the availability of skilled workers and/or development of skills were essential requirements for improving productivity.

Here are the key takeaways from the literature review above:

- **Shift to Low-Productivity Sectors:** The UK has seen a shift towards relatively low-productivity sectors such as services and the public sector such as healthcare and education. This transition has contributed to the declining productivity growth.
- **Drivers of declining productivity:** While the above shift has contributed to falling productivity growth, it is the manufacturing, ICT and, finance and insurance sectors that were identified as the top sectors responsible for this decline.
- **Impact of Subsectors and Industries:** Coyle's study decomposed growth into contributions from different subsectors and sectors. It attributed the productivity slowdown to specific subsectors within manufacturing and ICT, such as transport equipment, pharmaceuticals, computer software, and telecommunications. This suggests that addressing productivity challenges requires a sector-specific approach.
- **Regional Disparities:** The Skill and Productivity Board and the Levelling Up White Paper highlight regional disparities in productivity, emphasizing the role of specializations in low-productivity industries. Factors such as low levels of innovation, knowledge diffusion, inadequate management, and workplace relations contribute to declining productivity in certain regions.
- **Job Polarisation and Skills Mismatch:** Acemoglu and Autor's research indicates that technological change has led to job polarization, favouring high-skill non-routine jobs and low-skill occupations while hollowing out middle-skill jobs. The rise in skill-biased technological change has contributed to wage inequality. The emerging theme

of job polarization and skills mismatch underscores the need to address the evolving demands of the labour market.

- **Impact of Skills and Qualifications Mismatch:** The OECD working paper emphasizes that qualifications alone may not reflect skills mismatch accurately. Skills, including additional training and gained or lost skills beyond formal qualifications, play a crucial role in productivity. Over-skilled and underqualified workers have negative impacts on productivity, while overqualified and overskilled workers do not. Allocative efficiency is negatively affected by being over-skilled alone, suggesting its greater impact on productivity compared to being overqualified.
- **Importance of Policy Factors:** The research highlights the need for policy interventions to tackle the productivity challenge, such as investing in targeted training programs and improving residential mobility to promote efficient reallocation of labour. Additionally, enhancing managerial quality is crucial for addressing within-firm productivity issues associated with under-qualification.
- **Importance of organisational change on productivity:** Any benefits of computerisation are far outweighed by adverse interactions with current organisational practises, so incremental change, either by investing in computers without implementing organisational change or only partially implementing some organisational changes, can result in significant productivity losses.
- **Evidence that investment in technologies enhance productivity:** Many researchers believe that investment in and the adoption of digital technologies have a positive and significant impact on productivity. Though this is arguable.
- **Importance of management and leadership quality:** Increasingly important for firms to have the right organisational setups and business practices to ensure that they are investing in innovation in a way to that helps them best realise the productivity gains.

### **3. Project Design**

#### **3.1. Project Location Rationale**

The project is situated in Rochdale, a borough of Greater Manchester. This is an area which is identified as having lower productivity and is a priority area for regional improvement. Greater Manchester has a productivity deficit relative to the national average. Within Greater Manchester, the boroughs in Greater Manchester's North East - Rochdale and Oldham - have lower productivity than the other parts of Greater

Manchester. Rochdale also has some of the most deprived wards in England and is ranked the 15th most deprived in England. For all these reasons, Rochdale is a priority area for investment under the UK Communities Renewal Fund and UK Shared Prosperity Fund. Focusing on Rochdale allows us to investigate the barriers that businesses face in a low productivity location, coupled with the management decision making processes that govern the adoption of technologies that can assist in improving productivity. The findings from this project will be used to create a repeatable framework that can be mirrored across Greater Manchester and beyond.

### **3.2. Methodology for Identifying Firm Level Productivity**

Estimation of firm level productivity is possible using data within the profit and loss account and balance sheet within company accounts. This project sourced that data from FAME, an online database organised by Bureau van Dijk. Filters were applied such that only companies were selected that had a primary trading address in Rochdale and had between 10 and 250 employees. This resulted in a sample population of 76 active SMEs in Rochdale for which there was sufficient data to estimate firm-level productivity.

Gross Value Added (GVA) at the firm level can be calculated via the income approach. Namely,  $GVA = \text{Earnings Before Interest, Tax, Depreciation and Amortization (EBITDA)} + \text{Employee Costs}$  GVA per employee is the adopted measure of labour productivity at the firm level within this project. Using a three-year average, this metric provides an initial categorisation of Rochdale SMEs into different productivity bands from which the mentors and mentees can be drawn.

The next stage was to cross-reference this sample of firms with information held in Greater Manchester Chamber of Commerce's own CRM system so that firm level characteristics such as expansion plans, investment, recent increase (or decrease) in the number of employees and engagement with Chamber initiatives can be assessed. These aspects offer a window into the firm's decision making on productivity enhancing initiatives. This process supported the identification of organisations which have higher productivity to act as mentors to mentees from organisations that fall outside this category.

Along with this, data from other SME support projects delivered by Greater Manchester Chamber of Commerce will be used to understand whether the identified Rochdale based businesses have undertaken training to develop management competencies and if so, what specific types of training have been undertaken. Since one of the barriers to improving productivity is not merely the adoption of technology but the ability to utilise the adopted technology, it is important to understand whether firm management have the knowledge, aptitude, and skills for managing technology and digital transformation projects. Learning from this pragmatic approach will also inform future delivery of the scheme, but it is not intended to be a core element of the mentoring framework.

### **3.3. Methodology for Mentoring Pilot**

The mentoring scheme connects mentors and mentees from different SMEs in the same area, controlling for comparable size. The mentoring programme is designed to focus on improving productivity outcomes of strategic decision-making processes, and the participants are senior managers or staff with responsibility for investment decision making. Mentoring pairs are cross-sector to support a fertilisation of ideas and avoid matching direct competitors. The mentoring scheme is designed to operate over 6 meetings – a startup meeting followed by 5 meetings in which participants explore the functional aspects of productivity and how these feed into planning and strategy at the organisational level. The meetings are a collaboration from one business leader to another, sharing insights, transferring knowledge, agreeing on ideas how they can embed good practice it into their operations to add value. The participants are guided to discuss the following in the meetings:

- How do they engage in strategic planning and decision making for productivity?
- Where are the barriers and blockages to make or implement decisions to improve productivity, and how are they being addressed?
- How do business leaders make decisions around investment and activity on five key productivity drivers?
- How is productivity included in strategic planning?
- How do different functional leaders address productivity in their strategic planning?

The meetings are recorded and transcribed enabling a qualitative analysis to be conducted. This stage will focus on drawing together the findings from the meetings, and from an end of scheme workshop for all participants. Each topic will be dealt with individually and a summary across each of the five areas developed. These results will be provided at the conference. The Productivity Mindset Mentoring Framework will be developed for use in future projects based on the learning from this project.

### **3.3. Results**

To be presented at conference following project completion.

## **4. Conclusions**

The literature review highlights the complexity of the productivity paradox, with declining growth attributed to lower productivity sectors, regional disparities, job polarisation and a skills mismatch. However, much of this research has been conducted at a national level, overlooking potential variations in productivity at the local level. At the firm level, while earlier studies found no relationship between technology investment and productivity, recent research emphasises the critical role of complementary organisational changes in realising the benefits of technology investments. The latter evidence indicates that managerial quality, business processes, knowledge sharing, and workplace practices impact productivity outcomes.

Our research aims to bridge the gap by addressing the local variations and add to firm level insights. The mentoring pilot programme seeks to uncover differences in strategic decision-making around technology investments between higher and lower productivity firms. It will elucidate the organisational factors that enable more productive technology investments. The qualitative analysis of the mentoring meetings will provide rich insights into these issues at the local level. The repeatable framework we aim to design from this process will help us better understand the productivity challenges plaguing this region and equip firms with the tools and

techniques needed to overcome this challenge. This provides opportunities for IS scholars to examine the productivity paradox in more detail and, as emphasised in the introduction, we invite fellow IS researchers to develop and enhance the work of others works such as Brynjolfsson's. Overall, the research promises actionable insights for various stakeholders aiming to tackle the UK's productivity paradox and provides the opportunity for significant impact if an effective mentoring scheme can be developed and rolled out at scale.

## 5.0 References

- Acemoglu, D. and Autor, D., 2011. Skills, tasks and technologies: Implications for employment and earnings. In *Handbook of labor economics* (Vol. 4, pp. 1043-1171). Elsevier. DOI: [https://doi.org/10.1016/S0169-7218\(11\)02410-5](https://doi.org/10.1016/S0169-7218(11)02410-5)
- Arthur, W.B. (1996) *Increasing Returns and the New World of Business*, Harvard Business Review. Available at: <https://hbr.org/1996/07/increasing-returns-and-the-new-world-of-business>.
- Barua, A., Kriebel, C.H. and Mukhopadhyay, T. (1995) 'Information Technologies and Business Value: An Analytic and Empirical Investigation', *Information Systems Research*, 6(1), pp. 3–23. Available at: <https://doi.org/10.1287/isre.6.1.3>.
- Battu, H., Belfield, C.R. and Sloane, P.J., 1999. Overeducation among graduates: a cohort view. *Education economics*, 7(1), pp.21-38. DOI: <https://doi.org/10.1080/096452999000000002>
- Bender, S. et al. (2016) 'Management Practices, Workforce Selection and Productivity', *Journal of Labor Economics*, 36(1), pp. 371–409. Available at: <https://www.nber.org/papers/w22101>.
- Bloom, N., Sadun, R. and Reenen, J.V. (2012) 'Americans Do IT Better: US Multinationals and the Productivity Miracle', *American Economic Review*, 102(1), pp. 167–201. Available at: <https://scholar.harvard.edu/rsadun/publications/americans-do-it-better-us-multinationals-and-productivity-miracle> (Accessed: 20 September 2023).
- Bloom, N. and Van Reenen, J. (2006) 'Measuring and Explaining Management Practices Across Firms and Countries', *The Quarterly Journal of Economics*, 122(4), pp. 1351–1408. Available at: <https://doi.org/10.3386/w12216>.
- Boothby, D., Dufour, A. and Tang, J., 2010. Technology adoption, training and productivity performance. *Research Policy*, 39(5), pp.650-661. DOI: <https://doi.org/10.1016/j.respol.2010.02.011>
- Brynjolfsson, E. and Hitt, L.M. (1998) 'Beyond the Productivity Paradox', *Communications of the ACM*, 41(8), pp. 49–55. Available at: <https://doi.org/10.1145/280324.280332>.
- Brynjolfsson, E. and Hitt, L.M. (2000) 'Beyond Computation: Information Technology, Organizational Transformation and Business Performance', *Journal of Economic Perspectives*, 14(4), pp. 23–48. Available at: <https://doi.org/10.1257/jep.14.4.23>.



- Brynjolfsson, E. and Petropoulos, G. (2021) 'The coming productivity boom', MIT Technology Review, Available at: <https://www.technologyreview.com/2021/06/10/1026008/the-coming-productivity-boom/> (Accessed: 20 October 2023).
- Brynjolfsson, E., Renshaw, A. and Van Alstyne, M. (1996) *The Matrix of Change: A Tool for Business Process Reengineering*, MIT Sloan School of Management. Available at: <http://web.mit.edu/marshall/www/papers/MoC.pdf> (Accessed: 20 September 2023).
- Cardona, M., Kretschmer, T. and Strobel, T. (2013). ICT and productivity: conclusions from the empirical literature. *Information Economics and policy*, 25(3), pp.109-125. <https://doi.org/10.1016/j.infoecopol.2012.12.002>
- Cavaglia, C. and Etheridge, B., 2020. Job polarization and the declining quality of knowledge workers: Evidence from the UK and Germany. *Labour Economics*, 66, p.101884. DOI: <https://doi.org/10.1016/j.labeco.2020.101884>
- Coyle, D. and Mei, J.C., 2023. Diagnosing the UK productivity slowdown: which sectors matter and why?. *Economica*. DOI: <https://doi.org/10.1111/ecca.12459>
- Cusolito, A.P., Lederman, D. and Peña, J., 2020. The effects of digital-technology adoption on productivity and factor demand: Firm-level evidence from developing countries. World Bank Group, Middle East and North Africa Region, Office of the Chief Economist.
- Dasgupta, S., Sarkis, J. and Talluri, S. (1999) 'Influence of Information Technology Investment on Firm productivity: a Cross-sectional Study', *Logistics Information Management*, 12(1/2), pp. 120–129. Available at: <https://doi.org/10.1108/09576059910256493>.
- Desjardins, R. and Rubenson, K., 2011. An analysis of skill mismatch using direct measures of skills. DOI: <https://doi.org/10.1787/19939019> Available online: <https://www.oecd-ilibrary.org/content/paper/5kg3nh9h52g5-en> [Accessed on: 28 October 2023]
- Driffield, N. et al. (2021) *Understanding productivity: Organisational Capital Perspectives*. The Productivity Institute. Available at: <https://www.productivity.ac.uk/wp-content/uploads/2021/11/WP013-Organisational-Capital-scoping-paper-FINAL-151121.pdf>.
- Dos Santos, B., & Sussman, L. (2000). Improving the return on IT investment: the productivity paradox. *International journal of information management*, 20(6), 429-440.
- Durbin, S. (2004) *Review of Workplace Skills, Technology Adoption and Firm Productivity: A Review*, New Zealand Treasury Working Paper, No. 04/16, New Zealand Government, The Treasury, Wellington
- Gal, P., Nicoletti, G., Renault, T., Sorbe, S. and Timiliotis, C., 2019. Digitalisation and productivity: In search of the holy grail—Firm-level empirical evidence from EU countries. DOI: <https://doi.org/10.1787/18151973>
- Green, F. and Zhu, Y., 2010. Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. *Oxford economic papers*, 62(4), pp.740-763. DOI: <https://doi.org/10.1093/oeq/gpq002>
- HM Government (2022) *Levelling Up the United Kingdom: executive summary* available online <https://www.gov.uk/government/publications/levelling-up-the-united-kingdom> [Accessed: 25 October 2023]
- Ichniowski, C., Shaw, K. and Prennushi, G. (1995) 'The Effects of Human Resource Management Practices on Productivity', *American Economic Review*, 86, pp. 291–313. Available at: <https://doi.org/10.3386/w5333>.

- Klueter, T., Monteiro, F. and Dunlap, D. (2017) 'Standard vs. Partnership-Embedded Licensing: Attention and the Relationship between Licensing and Product Innovations', Mack Institute for Innovation Management, 6(9), pp. 1629–1643. Available at: <https://mackinstitute.wharton.upenn.edu/2017/standard-vs-partnership-embedded-licensing/> (Accessed: 20 September 2023).
- Loveman, G.W. (1993) 'An Assessment of the Productivity Impact of Information Technologies', in T.J. Allen and M.S.S. Morton (eds) *Information Technology and the Corporation of the 1990s: Research Studies*. Oxford : Oxford University Press, pp. 84–110.
- López, A. (2023): The role of information technology and workplace organization in firm productivity: evidence from Spanish firms, *Economics of Innovation and New Technology*, DOI: <https://doi.org/10.1080/10438599.2023.2172000>
- Lucas Jr, H. C. (1999). *Information technology and the productivity paradox: Assessing the value of investing in IT*. Oxford University Press.
- McCann, P. (2018) *Productivity Perspectives* Available online : <https://productivityinsightsnetwork.co.uk/app/uploads/2018/06/P-McCann-Final-synthesis.pdf> [Accessed on: 20 October 2023]
- McGowan, M.A. and Andrews, D., 2015. Skill mismatch and public policy in OECD countries. DOI: <https://doi.org/10.1787/18151973> Available online: <https://www.oecd-ilibrary.org/content/paper/5js1pzw9lnwk-en> [Accessed on: 28 October 2023]
- Milgrom, P. and Roberts, J. (1990) 'The Economics of Modern Manufacturing: Technology, Strategy, and Organization', *The American Economic Review*, 80(3), pp. 511–528. Available at: <https://www.jstor.org/stable/2006681>.
- Mithas, S., & Lucas Jr, H. C. (2014). *Information Technology and Firm Value: Productivity Paradox, Profitability Paradox, and New Frontiers*.
- Montresor, G., 2019. Job polarization and labour supply changes in the UK. *Labour Economics*, 58, pp.187-203. DOI: <https://doi.org/10.1016/j.labeco.2018.05.009>
- Morrison, C.J., 1997. Assessing the productivity of information technology equipment in US manufacturing industries. *Review of Economics and Statistics*, 79(3), pp.471-481. DOI: <https://doi.org/10.1162/003465300556887>
- Mosiashvili, N. and J. Pareliussen (2020), "Digital technology adoption, productivity gains in adopting firms and sectoral spill-overs: Firm-level evidence from Estonia", *OECD Economics Department Working Papers*, No. 1638, OECD Publishing, Paris, <https://doi.org/10.1787/ba9d00be-en>.
- Ollivaud, P., Guillemette, Y. and Turner, D. (2016) *Links between weak investment and the slowdown in productivity and potential output growth across the OECD*, *Library Catalog (Blacklight)*. Paris: OECD Publishing. Available at: <https://searchworks.stanford.edu/view/11731955> (Accessed: 20 September 2023).
- ONS (2018). *Information and communication technology intensity and productivity*. Office for National Statistics. available online at <https://www.ons.gov.uk/economy/economicoutputandproductivity/productivitymeasures/articles/informationandcommunicationtechnologyintensityandproductivity/2018-10-05> [accessed on 16 October 2023]
- Pinsonneault, A., & Rivard, S. (1998). Information technology and the nature of managerial work: From the productivity paradox to the Icarus paradox?. *MIS quarterly*, 287-311.

- Remes, J., Mischke, J. and Krishnan, M., 2018. Solving the productivity puzzle: The role of demand and the promise of digitization. *International Productivity Monitor*, (35), pp.28-51. Available online: <http://www.csls.ca/ipm/35/remes-mischke-Krishnan.pdf> [Accessed on: 28 October 2023]
- Robinson, C. & Siegel, C. & Liao, S. (2021). "Technology Adoption and Skills A Pilot Study of Kent SMEs," *Studies in Economics* 2114, School of Economics, University of Kent.
- Romer, P.M. (1990) 'Endogenous Technological Change', *Journal of Political Economy*, 98(5), pp. 71–102. Available at: <http://www.jstor.org/stable/2937632>
- Schreyer, P. (2001) Measuring productivity: measurement of aggregate and industry-level productivity growth. OECD Manual. Available at: <https://www.oecd.org/sdd/productivity-stats/2352458.pdf>.
- Solow, R.M. (1956) 'A Contribution to the Theory of Economic Growth', *The Quarterly Journal of Economics*, 70(1), pp. 65–94. Available at: <https://doi.org/10.2307/1884513>
- Stiroh, K.J. (2002) 'Information Technology and the U.S. Productivity Revival: What Do the Industry Data Say?', *American Economic Review*, 92(5), pp. 1559–1576. Available at: <https://doi.org/10.1257/000282802762024638>
- Taştan, H. Gönel, F. (2020), ICT labor, software usage, and productivity: firm-level evidence from Turkey. *J Prod Anal* 53, 265–285. <https://doi.org/10.1007/s11123-020-00573-x>
- Venturini, F. (2015) 'The modern drivers of productivity', *Research Policy*, 44(2), pp. 357–369. Available at: <https://doi.org/10.1016/j.respol.2014.10.011>

# Apparel Matters? A New First-Impression Static Appearance Dataset

## Abstract

*Automatic first impression estimation has garnered increased attention in recent years due to its potential applications in various fields. However, the momentum of this research is hampered by the significant lack of specialised datasets. Existing datasets largely focus on facial features, leaving out other critical elements that play a pivotal role in forming first impressions. Addressing this gap, we introduce the First Impression Static Appearance Dataset (FISAD). This comprehensive dataset comprises 6000 full-body images, each meticulously annotated with the Big Five first impression traits. Notably, FISAD opens up a new dimension for investigation, spotlighting the often-overlooked influence of clothing in first impression formation. Researchers can now delve deeper into understanding how attire interacts with and influences people's initial perceptions, paving the way for richer, more holistic studies in the realm of first impression estimation.*

**Keywords:** First impression, Dataset, Apparel

## 1.0 Introduction

People form first impressions of others in a brief moment, often relying on a quick glance. This phenomenon, where snap judgments can be both accurate and reliable, has been extensively

studied by psychologists (Celiktutan et al., 2017). Historically, numerous theories and models have been proposed to explain how first impressions are formed. Among these, traits have been identified as a crucial measurement in understanding individuals. According to Trait Theory (Costa & McCrae, 1998), these are stable characteristics that consistently influence a person's responses across various situations.

While strides have been made in understanding first impressions through traits, a significant challenge in accurately capturing these impressions, particularly in computer science, remains (Vinciarelli & Mohammadi, 2014). A notable limitation in current datasets is the

underrepresentation of apparel, a key element in shaping first impressions. Current datasets predominantly focus on features like facial expressions, background, and verbal cues, often

overlooking how clothing and style contribute to these judgments. This oversight is significant

given that apparel can convey a wealth of information about an individual's personality, social status, and even mood (Kodžoman, 2019). In many real-world scenarios, such as job interviews or professional networking events, apparel plays a crucial role in conveying professionalism, personality, and even the individual's attention to detail. Existing datasets

predominantly focus on facial expressions and verbal cues, neglecting how clothing

in the automatic recognition of first impressions, especially in situations where non-verbal cues like clothing are predominant in forming judgments.

Addressing these challenges, our study introduces the First Impression Static Appearance Dataset (FISAD), a unique full-body image dataset annotated with Big Five first impression traits. This dataset enables researchers to delve into the correlation between static appearance and first impression traits, independent of temporal visual information. By focusing on static images, FISAD offers a new dimension to first impression research, providing valuable insights that were previously unattainable due to the limitations of existing datasets and methodologies.

## 2.0 Literature Review

The first impression is one's initial perception of another person. Typically, appearance is the first piece of information available to a newly encountered person. Therefore, appearance can strongly influence the formation of first impressions from observers (Efran, 1974). The attempt at first impression estimation based on appearance can be traced back to the early days. Albright and his team tested the first impression based on the physical appearance of 259 subjects and concluded that extraversion and conscientiousness could be measured accurately (Albright et al., 1988). Another study by Peter Borkebau revealed extraversion strongly related to physical appearance. This study was done by 149 stimulus people and 24 perceivers. In the past two decades, first impression estimation has attracted more attention in the computer science field. The backbone reason is the first impression formation is stable enough to be modelled by computer science (Vinciarelli, 2016). Thus, with the development of computer vision techniques, image classification and pattern recognition have become mainstream, as most features can be recognised in videos and images.

Compared to the boom of the technical method, the lack of large public datasets with appropriate annotation and protocol becomes the bottleneck in the apparent first impression estimation research. As summarised in Table 1.

Name	Data Type	No. Data	Description	Annotation	Public
MAPTRAITS (Vinciarelli & Mohammadi, 2014)	Video	44	44 video clips	Big five traits and engagement, facial attractiveness, vocal attractiveness, and likability	Yes

ChaLearn First Impression (Ponce-López et al., 2016)	Video	10,000	10K 15sec video clips extracted from YouTube vlogs People facing and speaking to a camera, show their faces and shoulders.	Big five traits	Yes
Transcription of Youtube Video Blogs (Biel & Gatica-Perez, 2012)	Video with transcriptions	2260	2260 videos from 442 YouTube vlogs	Big five traits	Yes

**Table 1. List of Appearance First Impression Datasets**

The MAPTRAITS dataset (Vinciarelli & Mohammadi, 2014), created in 2014, was claimed to be the first dataset in the Audio/Visual mapping personality traits challenge. The dataset contains 44 video clip recordings of 11 different subjects interacting with four virtual characters. Raters were asked to rate along with the Big five traits and four other traits (engagement, facial attractiveness, vocal attractiveness, and likability). This challenge aimed to develop the automated method of personality and social traits estimation upon the visual, vocal, or combination of both information. The ground-breaking challenge attracted more attention to first impression computing. But there are limitations. First, the dataset size is small, with only 44 video clips with 11 subjects. Second, the subject's appearance is limited to face and shoulder only. In the same year, The workshop on Computational Personality Recognition published a customised dataset (Biel & Gatica-Perez, 2012) aimed to define the state-of-art approaches to personality recognition tasks. The dataset presented 2260 videos from 442 YouTube vlogs and collected 2200 annotations in Amazon Mechanical Turk. The sample videos were searched and collected using the keywords "vlogs" and "vlogging". The author filtered the video by restriction, which is a single person, showing their head and shoulders, and talking to the camera. This dataset has a similar shortcoming as the MAPTRAITS dataset.

In 2016, the ChaLearn dataset (Ponce-López et al., 2016) was designed for the First Impression challenge. It is used to automatically evaluate apparent personality traits (Big Five). The dataset consists of 10,000 short clips from YouTube and is labelled by Amazon Mechanical Turk. Each video is 15 seconds long and contains one person talking to the camera in a self-presentation context. Chalearn is one of the first data corpus in first impression computing and has become the most popular dataset. The team made the ChaLearn dataset publicly available, which significantly contributes to the community and offers the foundation of first impression research. But the research based on the ChaLearn dataset is limited by its content. (1) Only the face and upper shoulder of the body are exposed in the video, which restricts the studies base on the dataset can only focus on the facial features. (2) The videos are taken in natural private spaces, most in the host's own room; the furniture, ornaments, and other decoration may reflect the owner's taste. But such an advantage is doubtful to occur in most of real-life applications. In summary, the current appearance first impression dataset has several features in common. (1) Most of the raw data are video presentations from YouTube. The reason caused this is because most of raw data in the public domain are protected by copyright which takes a lot of time and effort for creators to contact and try to gain permission from authors. Among different online resources, YouTube is the most easily accessible resource platform since it allows video creators to mark their works with a Creative Commons CC BY licence<sup>1</sup>. Therefore, we noticed that most appearance first impression datasets are built upon YouTube videos. (2) Due to the fundamental requirement of first impression study, raw data needs to contain only one person appearing in the video, and the subject was talking to the camera in order to have voice feature. Therefore, the creators found Q&A videos particularly suitable. As a consequence, the person in the video has only face and shoulder appearance. However, this focus on face and shoulder appearance inherently limits the scope of first impression analysis. It omits critical aspects of non-verbal communication and personal style conveyed through full-body language and attire, which are significant factors in holistic first impression formation. (3) Most existing datasets we explored in first impression prediction utilize the Big Five traits – Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness. This is due to their established robustness in capturing key personality aspects, as noted in studies by (Atherton et al., 2022) and (Buecker et al., 2020). Our use of these traits ensures compatibility with prior research and aids in straightforward evaluation and interpretation of results. Aligning with this framework

---

<sup>1</sup> CC BY allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, so long as attribution is given to the creator. The license allows for commercial use.

also allows for effective comparison with existing work, particularly important as our dataset focuses on attire and static appearance, contrasting with those based on dynamic video data. The purpose of creating FISAD is to fill the following gaps (1) Clothing cues play an important role in first impression formation (Angerosa, 2014). It is a nonverbal tool that passes visual information to the observer. Unlike facial information, clothing is a manipulatable feature containing a complex array of information the wearer would like to deliver. In our dataset, full-body images allow the researcher to explore the clothing effect on the first impression formation. (2) The video presentation offers rich dynamic information. Such temporal features have been approved to positively impact the output (Junior et al., 2019). But since machine learning is a "Black Box" (Pamina & Raja, 2019), the importance of each feature is kept unknown. Our dataset focuses on static features by offering the still image, which eliminates the effect of visual temporal information.

### 3.0 The First-Impression Static Appearance Dataset

We contribute the FISAD (First Impression Static Appearance Dataset) to the community. FISAD is a large-scale dataset with 6000 full-body images and each annotated with Big Five first impression traits by the Amazon Mechanical Turk service. The dataset received 84,129 annotations from 2,211 participants (1,338 male and 872 female, with a mean age is 39.5 years old).

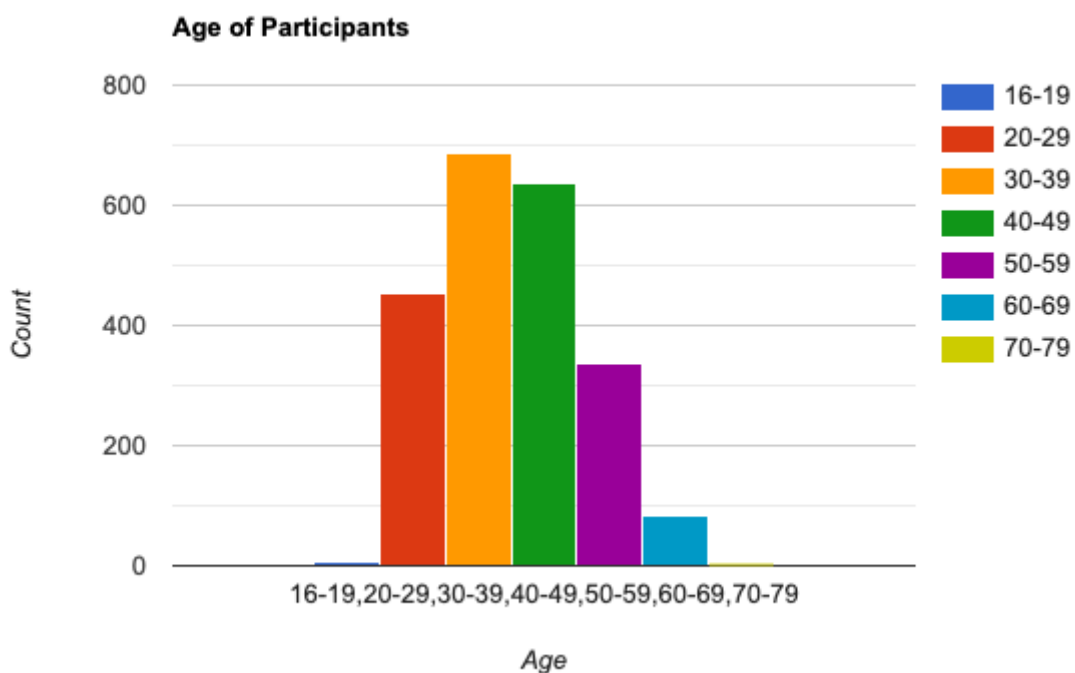
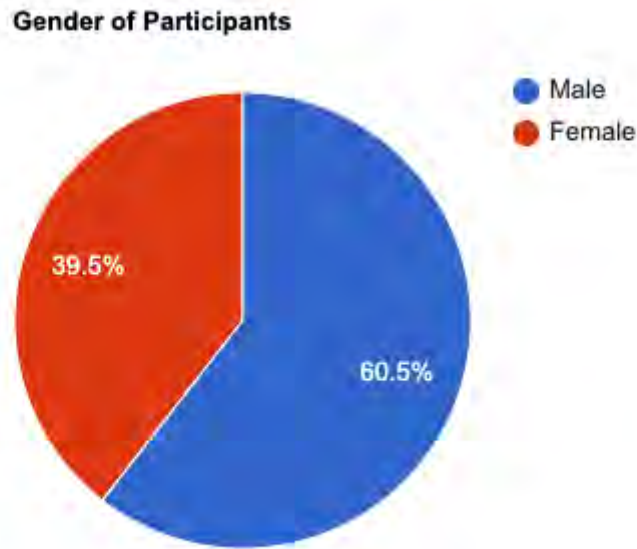


Figure 1. Age of Participants





**Figure 2. Gender of Participants**

### **3.1 Data image**

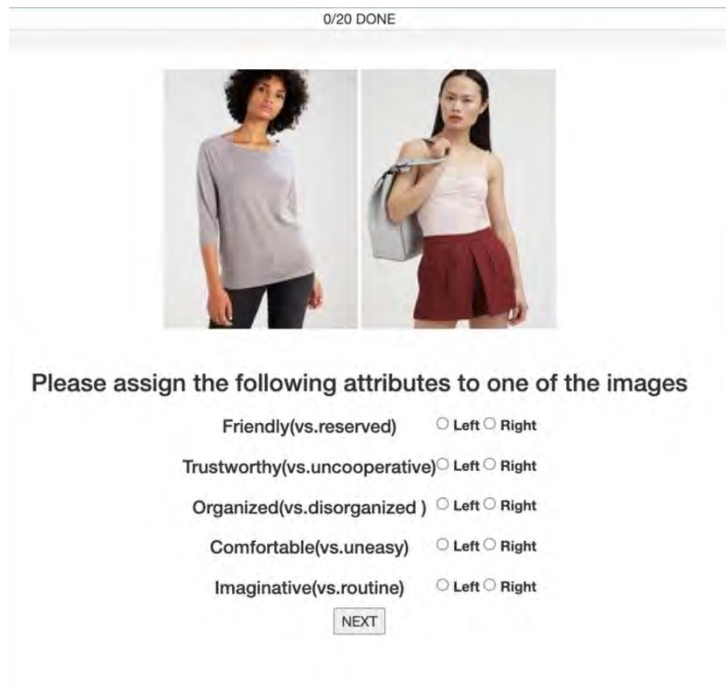
The FISAD images are sourced from a subset of the Multi-pose Virtual Try-on (MPV) dataset (Dong et al., 2019). The MPV dataset was initially designed for the 4th Look Into Person (LIP) Challenge, consisting of 37,723 person images and 14,360 cloth images, with a resolution of 256x192 pixels. First, we used Viola-Jones from OpenCV (Viola & Jones, 2004). to separate images containing either a person or cloth only, then we manually ran through the dataset and removed unusable images. The rules of image picks were as follow: (1) the person must show his/her front face. (2) both eyes must be visible. (3) the person in the image must have the entire upper half of the outfit visible, and the lower half of the outfit must be recognisable. In total, 6,000 person images are kept constructing the FISAD. Figure 3 illustrates sample of images in dataset.



**Figure 3.** Sample of dataset image

### **3.2 Acquisition of Ground Truth Annotation**

Obtaining the ground truth of first impressions poses a significant challenge, as it often involves stereotyping judgments based on the observer's personal experience (Cuddy et al., 2008). To reduce variance, we collected multiple votes per image (averaging 14 votes per image). For annotation, we utilised the Amazon Mechanical Turk (AMT) service, a common platform for quickly amassing a large number of labels. Voters receive a small payment in return for their contributions. However, each voter typically provides only a few annotations in a large dataset, which can introduce bias. As mentioned earlier, first impressions are personal judgments influenced by a voter's experiences and cultural norms, factors such as age, gender, race, and cultural biases, while unavoidable, are challenging to quantify. To address this issue, we developed a custom website that incorporates our algorithm to mitigate bias.



**Figure 4. Data collection comparison web page**

### **3.2.1 Data collection process**

All voters were recruited via the Amazon Mechanical Turk Service. Participants were required to meet two criteria (1) participants need to be 18+, and (2) participants need to be registered in one of the chosen country lists. Each AMT voter willing to participate was redirected to our data collection website and display the term and conditions. If voters agreed to start, they were assigned 20 pairs of images to compare. The layout of the comparison web page is shown in Figure 4. The voter is asked to compare two images and share their preference in Big Five traits: Extraversion=Friendly vs reserved; Agreeableness = Trustworthy vs uncooperative; Conscientiousness = Organised vs disorganised; Neuroticism = Comfortable vs uneasy; Openness = Imaginative vs routine. We adopted pairwise comparison to reduce statistical bias. Compared to direct rating, the pairwise comparison can significantly improve the accuracy (Koczkodaj, 1996). The comparison process repeats 20 times. Upon completion, the system recorded the voter's age and gender, and a token was generated for the voter to claim payment. The total process time was limited to 7 minutes to capture the observer's initial thoughts. The vote for the first two pairs will not be counted. This is the training process for the voters to be familiar with the work. For the remaining 18 comparisons, the image selected for each trait earned 1 point, while the other received 0 points. The overall score for each trait of an image was calculated by the following formula (1), where  $S_i$  is the score for each vote, and N is the total number of comparisons done on this image.

$$S_{(\text{one of the five traits})} = \frac{1}{N_S} \sum_{i=1}^{N_S} S_i \quad (1)$$

We used a custom algorithm to pick images for comparison. The left image was chosen as the one with the fewest comparisons to ensure equal exposure for all images, ensuring every image has a similar number of exposures. The right image was selected based on having the smallest difference in the five traits compared to the left image. The distance is calculated as:

$$D = \frac{1}{5} \sum_{i=1}^5 |l_i - r_i| \quad (2)$$

Where  $l_i$  is the score of the left image for a trait and  $r_i$  is the score of the candidate image for the same trait. The vote results will be stored in the database, and the mean of points is used as the trait score of each image.

### 3.2.2 Address Bias

First impression formation is subjective, influenced by observers' personal experiences and culture, which is where bias originates. Bias is common and challenging to avoid. We designed several strategies to address the following biases:

- The leniency-strictness bias: In our annotation collection process design, each image is rated by multiple observers, with each observer rating only a few images. Consequently, leniency-strictness bias becomes a significant issue. For instance, some observers may be lenient, rating images too highly, while others may be overly strict, resulting in low scores. We addressed this by implementing pairwise comparisons, allowing observers to compare two images instead of rating an individual image.
- Gender bias: Numerous studies indicate gender differences in personality. For example, females are often perceived as more agreeable than males [29], and generally score higher on neuroticism traits [30]. These findings suggest that comparing images of different genders leads to biased comparisons. To mitigate gender bias, we selected images of the same gender for comparison. Currently, our dataset contains only female images, but gender bias will be considered in future development.
- Operational bias: This occurs when an observer's behaviour changes over time during data collection, typically as familiarity with the process reduces carefulness. To address this, we limited the number of comparisons per observer, and the first two comparisons do not contribute to the final score.

## 4.0 Discussion and future work

The First Impression Static Appearance Dataset (FISAD) represents a pioneering step in the study of first impressions, focusing specifically on the role of apparel in static appearance features. Its potential goes beyond automated first impression prediction, providing valuable insights for applications in human resources, fashion retail, marketing, social psychology, and AI-driven personal styling. For instance, in human resources, FISAD can assist in developing tools to interpret non-verbal cues in job interviews. In the realms of fashion and marketing, it can inform design and advertising strategies by unveiling consumer preferences. Social psychologists can utilise it to comprehend the impact of clothing on perceptions of personality and status. Additionally, FISAD's distinctive focus on apparel has the potential to revolutionise AI personal styling, enhancing digital fashion experiences. Despite the inherently subjective nature of first impressions, our approach of securing multiple evaluations per image has aided in maximising consensus, averaging 14 votes per image. Looking forward, we aim to increase this number, aspiring for even greater accuracy and representation in our dataset.

To facilitate further research and application, we are committed to making FISAD publicly available online after the completion of our research. This will allow the research community to leverage our findings, promoting advancements in various fields related to first impressions and apparel. Ensuring easy access and comprehensive documentation of FISAD, we aim to support its widespread use in future academic and practical explorations.

FISAD is not merely a resource for current research; it establishes the foundation for future investigations into the intricate relationship between apparel and first impression formation, bridging academic research with practical real-world applications.

## References

- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of personality and social psychology*, 55(3), 387.
- Angerosa, O. N. (2014). Clothing as communication: how person perception and social identity impact first impressions made by clothing. *Rochester Institute of Technology, Rochester*.
- Atherton, O. E., Sutin, A. R., Terracciano, A., & Robins, R. W. (2022). Stability and change in the Big Five personality traits: Findings from a longitudinal study of Mexican-origin adults. *Journal of personality and social psychology*, 122(2), 337.
- Biel, J.-I., & Gatica-Perez, D. (2012). The youtube lens: Crowdsourced personality impressions and audiovisual analysis of vlogs. *IEEE Transactions on Multimedia*, 15(1), 41-55.
- Buecker, S., Maes, M., Denissen, J. J., & Luhmann, M. (2020). Loneliness and the Big Five personality traits: A meta-analysis. *European Journal of Personality*, 34(1), 8-28.
- Celiktutan, O., Skordos, E., & Gunes, H. (2017). Multimodal human-human-robot interactions (mhhri) dataset for studying personality and engagement. *IEEE Transactions on Affective Computing*, 10(4), 484-497.

- Costa, P. T., & McCrae, R. R. (1998). Trait theories of personality. In *Advanced personality* (pp. 103-121). Springer.
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in experimental social psychology, 40*, 61-149.
- Dong, H., Liang, X., Shen, X., Wang, B., Lai, H., Zhu, J., Hu, Z., & Yin, J. (2019). Towards multi-pose guided virtual try-on network. Proceedings of the IEEE/CVF international conference on computer vision,
- Efran, M. G. (1974). The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury task. *Journal of Research in Personality, 8*(1), 45-54.
- Junior, J. C. J., Güçlütürk, Y., Pérez, M., Güçlü, U., Andujar, C., Baró, X., Escalante, H. J., Guyon, I., Van Gerven, M. A., & Van Lier, R. (2019). First impressions: A survey on vision-based apparent personality trait analysis. *IEEE Transactions on Affective Computing, 13*(1), 75-95.
- Koczkodaj, W. W. (1996). Statistically accurate evidence of improved error rate by pairwise comparisons. *Perceptual and Motor Skills, 82*(1), 43-48.
- Kodžoman, D. (2019). The psychology of clothing: Meaning of colors, body image and gender expression in fashion. *Textile & leather review, 2*(2), 90-103.
- Pamina, J., & Raja, B. (2019). Survey on deep learning algorithms. *International Journal of Emerging Technology and Innovative Engineering, 5*(1).
- Ponce-López, V., Chen, B., Olliu, M., Corneanu, C., Clapés, A., Guyon, I., Baró, X., Escalante, H. J., & Escalera, S. (2016). Chalearn lap 2016: First round challenge on first impressions-dataset and results. Computer Vision—ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part III 14,
- Vinciarelli, A. (2016). Social perception in machines: The case of personality and the Big-Five traits. *Toward Robotic Socially Believable Behaving Systems-Volume II: Modeling Social Signals*, 151-164.
- Vinciarelli, A., & Mohammadi, G. (2014). A survey of personality computing. *IEEE Transactions on Affective Computing, 5*(3), 273-291.
- Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision, 57*, 137-154.

# A Conceptual Framework and Design Principles for Decision Support in Clinical Practice: Managing Knowledge and Uncertainty

**Jishnu Das**

*University of Agder*

**Geir Inge Hausvik**

*University of Agder*

**Carl Erik Moe**

*University of Agder*

*Research In progress*

## **Abstract:**

*This research proposes a methodology for developing Clinical Decision Support Systems (CDSSs) through design science and clinician co-design, focusing on mental health referrals. It includes a case study illustrating how this approach can enhance care coordination, highlighting possible improvements. Integrating decision theory and discrete choice theory, the methodology addresses mental health referral coordination challenges, thereby enhancing system responsiveness and clinician engagement. Our conceptual model uses multifaceted reasoning and design principles for effective uncertainty communication, aiding informed decisions amidst incomplete information. Findings suggest applying these design principles and communication strategies across healthcare can significantly improve system functionality, making CDSSs more adept at handling clinical complexities. This adaptable methodology and case study exemplify a systematic approach to refine CDSSs, suggesting advancements in clinical decision support. The approach contributes both to theoretical knowledge and has practical implications, offering a scalable solution to improve healthcare delivery and patient outcomes in diverse settings.*

**Keywords:** Clinical decision support systems, decision theory, design science research, conceptual modelling, uncertainty

## **1.0 Introduction**

In mental health, clinical decision-making is challenged by the integration of multifaceted data and evolving standards (Smith, Johnson, & Lee, 2021; Sutton et al.,

2020). Despite foundational contributions (Smith, Johnson, & Lee, 2021; Sutton et al., 2020; Papathanasiou, Ploskas, & Linden, 2016; Guo et al., 2020), gaps in managing clinical uncertainty with limited information persist. Our study introduces a theory-driven Clinical Decision Support System (CDSS) model targeting these gaps in mental health referrals, contributing to both the theoretical and practical aspects of the Information Systems (IS) field.

Struijk et al. (2022) emphasize the importance of applying robust IS research methods to real-world challenges. Aligning with this principle, our study introduces a novel CDSS model aimed at improving clinical decision-making.

Kane (2022) discusses the need to address theoretical gaps in IS research. By critically examining existing CDSS models and incorporating decision theory and discrete choice theory, our study contributes to refining theoretical frameworks within the IS field, with a particular focus on uncertainty management.

Davison (2022) highlights the critical role of practical outcomes in research. Reflecting this focus, our co-designed, theory-informed CDSS model aims to enhance mental health referrals, thereby positively impacting patient care and resource allocation.

Our research explores, *how a clinical decision support system can integrate normative models and clinical judgment to manage uncertainty in mental health referrals*. By applying the Design Science Research Methodology (DSRM), our objective is to develop a Clinical Decision Support System (CDSS) specifically for the referral management of mental health patients in a hospital setting. This initiative is designed to yield insights that are valuable in both theoretical frameworks and practical applications.

The paper reviews relevant literature, present our methodology, detail the proposed CDSS design, and discuss the implications and contributions of our findings.

## **2.0 Background**

In this section we discuss how clinical decision support systems facilitate decision-making, and theoretical approaches to support this in a clinical setting.

### **2.1 Clinical decision support systems**

Clinical decision-making is increasingly complex, necessitating systems that can handle vast data, probabilistic information, and diverse care objectives (Meyer, Kiernan, McManus, & Shih, 2014). CDSSs offer potential solutions by providing data-



driven recommendations for point-of-care decisions (Sutton et al., 2020). Yet, their development faces challenges, including unclear best practices, managing uncertainty, aligning with clinical practices, and integrating varying data sources (Prakash & Das, 2021; Vasquez et al., 2022). These challenges are compounded by interoperability, ethical, legal constraints, and the need to stay abreast of evolving medical knowledge. Addressing these challenges calls for innovative, clinically tailored approaches. Incorporating decision theory and discrete choice theory can provide solutions that align with evidence-based practices and real-world clinical scenarios (Arnott, 2006; Train, 2009). Although the existing literature on CDSS lays a solid foundation, it also highlights significant gaps, particularly in mental health referrals. These gaps signal the need for a more comprehensive approach that considers the unique challenges and uncertainties within clinical settings. Such an approach aligns with Kane's (2022) theoretical critiques and Davison's (2022) emphasis on practical applications in the IS field.

## **2.2 Theoretical Approach**

In clinical decision-making, particularly in mental health referrals, the integration of decision theory and discrete choice theory offers a comprehensive framework to improve clinical decision support systems (CDSS). These theories provide complementary insights to better align CDSS with the complexities of clinical settings. Decision theory assists in guiding referral decisions based on clinical evidence, probabilistic estimates, and utility maximization concepts (Wernz et al., 2021). It quantitatively assesses probabilities and outcomes based on available data, crucial for evidence-based decision-making. Meanwhile, discrete choice theory delves into subjective factors affecting clinicians' decisions, encompassing beliefs, perceptions, attitudes, emotions, and cognitive biases (De Brún et al., 2018; Kochenderfer et al., 2015; Campos, Neves, & de Souza, 2007; Savage, 1972; Train, 2009).

Our study seeks to address the frequently overlooked dynamic and uncertain aspects of clinical decision-making by integrating these theories. By blending decision theory with discrete choice theory, we aim to develop a CDSS that accurately reflects the complexities and variabilities in clinical practices, enhancing the decision-making process. This approach facilitates effective communication of uncertainties and reinforces clinician discretion in referral decisions, employing methods like probability

scales and ambiguity indicators (Arend, 2020) to convey varying levels of certainty to clinicians.

### 3.0 Methodology

Our study employs the Design Science Research Methodology (DSRM), a structured, six-step process designed for developing innovative solutions to complex real-world problems (Hevner et al., 2004; Peffers et al., 2007). This approach is particularly suited to the multifaceted nature of clinical decision-making in mental health, where patient conditions and treatment options vary significantly.

In our approach, the first step, **problem identification**, involves comprehensive literature reviews and User-Centered Design (UCD) workshops with patients, General Practitioners (GPs), and specialists, aiming to gather essential design requirements. These requirements then guide the **definition of clear objectives**, focusing on optimizing clinical decision-making processes.

In the **design and development** phase, a low-fidelity prototype is created, reflecting these stakeholder requirements. This prototype undergoes iterative refinement, with its functionality and effectiveness **demonstrated** through controlled simulated scenarios. The **evaluation** stage assesses the system's performance against metrics like accuracy, usability, and clinician satisfaction. Finally, **communication** of findings through academic channels and workshops with healthcare professionals facilitates feedback and continuous improvement.

DSRM's iterative and collaborative nature makes it ideal for addressing 'wicked problems' in clinical settings, characterized by ambiguous and shifting elements (Peffers et al., 2007). It allows for the development of artifacts that are not only theoretically robust but also practically applicable in real-world scenarios. Additionally, the establishment of design principles guides the systematic design and evaluation of solutions, encompassing various aspects of knowledge representation and uncertainty communication necessary for supporting clinical discretion (Chandra et al., 2016).

By aligning this human-centered methodology with clinical constraints, DSRM facilitates the creation of solutions that can significantly improve complex health systems.

### **3.1 Anticipated Challenges and Solutions**

In our approach, a primary challenge is ensuring the prototype aligns with stakeholder-derived design requirements. To tackle this, we will adopt an iterative design process for continuous improvement based on feedback.

Creating simulations that accurately represent clinical scenarios for performance evaluation is also crucial. We intend to develop these simulations to closely replicate real clinical conditions, ensuring comprehensive and relevant testing of the prototype. Clear communication about the prototype's scope and intended functionality is essential. Our goal is to set realistic expectations about its capabilities, especially its role in the initial phases of CDSS testing and feedback.

These strategies are key to developing a low-fidelity prototype that is both practical and pertinent for ongoing research in clinical decision support systems. They will help us address challenges related to design fidelity, the authenticity of simulations, and effective communication.

### **3.2 Application of decision theory and discrete choice theory**

Design science research seeks to create and evaluate innovative artifacts grounded in theoretical foundations (Arnott & Pervan, 2012). The application of theory informs design at every step, from problem formulation to evaluation (Hu, Rao, Tao, et al., 2019).

During problem identification, decision theory provides normative models while discrete choice theory illuminates real-world constraints (Arnott & Pervan, 2012). In conceptual modelling, decision theory shapes representations for clinical discretion. Further, Discrete choice theory informs designing for reasoning (Hu, Rao, Tao, et al., 2019), e.g., in a clinical setting. For uncertainty modelling, decision theory supplies probability estimates to convey risks transparently based on evidence (Arnott & Pervan, 2012). Discrete choice theory identifies key ambiguity factors shaping decisions. In evaluation, decision theory provides probabilities for choices with given trade-offs (Arnott & Pervan, 2012).

The integration of these complementary theories can lead to an informed design process tailored to clinical contexts. This promotes best practices while responding to clinical needs (Arnott & Pervan, 2012; Hu, Rao, Tao, et al., 2019). Previous research contributions show the importance of theories informing the design [e.g., Arnott & Pervan, 2012, Wang et al., 2021; Gladstone, 2012; Terris-Prestholt et al., 2019;

Schwartz & Hevner, 2008]. By applying both decision theory and discrete choice theory, innovative yet pragmatic solutions suited for complex clinical contexts may be achieved.

### **3.3 Case Description**

Sørlandet Hospital Trust (SHT) in southern Norway serves over 200,000 residents and offers specialized services, including emergency care and surgery, at its main hospital in Kristiansand. The referral system between General Practitioners (GPs) and specialists faces challenges, notably inadequate information exchange, leading to a roughly 40% referral rejection rate. This issue, stemming from a lack of integrated records and referral systems, results in GPs lacking comprehensive patient history and specialists having limited understanding of GPs' assessments, culminating in prolonged wait times for consultations, often exceeding 60 days, and increased readmissions.

To improve the referral process, integrated health records and structured referral templates are proposed to enhance information sharing. Additionally, implementing clinical decision support systems using patient data will help refine triaging and establish tailored care pathways. Tackling these barriers is key to enhancing SHT's referral efficiency and reducing workarounds. The following section will outline these proposed solutions in detail.

### **4.0 Proposed Design**

This outlines a web-based clinical decision support system (CDSS) to improve mental health referrals and care coordination at SHT by increasing referral quality and appropriateness from GPs to specialists. Following design science research, our conceptual model will integrate clinical guidelines within the specialized context. We will refine the model via participatory design with specialist feedback.

Developing robust design principles is also important to provide a guiding framework for the CDSS. Participatory design engaging GPs, patients and specialists via focus groups, interviews, surveys, and prototyping workshops will inform realizing an improved system.

The solution will utilize a phased evaluation framework tailored to clinical contexts to effectively apply and test the design principles. Controlled simulation studies assessing performance using referral case examples and varied conditions will be conducted. Finally, usability testing will examine workflow integration and ease of use if feasible.

## 5.0 Discussion and Conclusion

Our study introduces a theory-driven conceptual model for Clinical Decision Support Systems (CDSS), specifically targeting mental health referral decision-making. This model suggests potential enhancements for scenarios with imperfect information, contributing both to theoretical understanding and practical application in the Information Systems (IS) field, as underscored by Struijk et al. (2022). It also supports Kane's (2022) call for addressing theoretical gaps, particularly in managing clinical uncertainty.

Focused on mental health referrals at a single hospital, our research offers insights for effective CDSS development, though its generalizability may be limited. In conclusion, this research progresses towards bridging the gaps in existing CDSS models, tackling theoretical and practical deficiencies. Incorporating design science research principles, the proposed model aligns with Davison (2022)'s emphasis on research with practical outcomes. Our findings emphasize the importance of integrating multifaceted knowledge and uncertainty communication in CDSS. Further research is encouraged to broaden the evaluation across diverse clinical contexts, enhancing the potential impact of these approaches in healthcare.

Moreover, this research offers significant advancements in clinical decision support systems, particularly for mental health referrals. By integrating decision theory and discrete choice theory, it addresses both theoretical gaps and practical needs in healthcare. This model sets a precedent for future CDSS development, potentially transforming healthcare practices and improving patient outcomes in various clinical settings.

## 6.0 References

- Arend, R.J. Strategic decision-making under ambiguity: a new problem space and a proposed optimization approach. *Bus Res* 13, 1231–1251 (2020). <https://doi.org/10.1007/s40685-020-00129-7>
- Arnott, David and Pervan, Graham (2012) "Design Science in Decision Support Systems Research: An Assessment using the Hevner, March, Park, and Ram Guidelines," *Journal of the Association for Information Systems*, 13(11), DOI: 10.17705/1jais.00315
- Chandra Kruse, L., Seidel, S., & Purao, S. (2016). Making use of design principles. In *Tackling Society's Grand Challenges with Design Science: 11th International*

- Conference, DESRIST 2016, St. John's, NL, Canada, May 23-25, 2016, Proceedings 11 (pp. 37-51). Springer International Publishing.
- Davison, R. M. (2023). Impact and implications for practice. *Information Systems Journal*, 33(2), 187-191.
- De Brún, A., Flynn, D., Ternent, L. et al. A novel design process for selection of attributes for inclusion in discrete choice experiments: case study exploring variation in clinical decision-making about thrombolysis in the treatment of acute ischaemic stroke. *BMC Health Serv Res* 18, 483 (2018). <https://doi.org/10.1186/s12913-018-3305-5>
- F. Campos, A. Neves and F. M. C. de Souza, "Decision Making under Subjective Uncertainty," 2007 IEEE Symposium on Computational Intelligence in Multi-Criteria Decision-Making, Honolulu, HI, USA, 2007, pp. 85-90, doi: 10.1109/MCDM.2007.369421.
- Gladstone, N. (2012). Comparative Theories in Clinical Decision Making and their Application to Practice: A Reflective Case Study. *British Journal of Anaesthetic & Recovery Nursing*, 13(3-4), 65-71. doi:10.1017/S1742645612000435
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75–105. <https://doi.org/10.2307/25148625>
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2008). Design science in information systems research. *Management Information Systems Quarterly*, 28(1), 6.
- Hoffmann, M., Vander Stichele, R., Bates, D. W., Björklund, J., Alexander, S., Andersson, M. L., & Gustafsson, L. L. (2020). Guiding principles for the use of knowledge bases and real-world data in clinical decision support systems: report by an international expert workshop at Karolinska Institutet. *npj Digital Medicine*, 3(1), 1-10.
- Hu, Z., Rao, C., Tao, C. et al. A case-based decision theory-based process model to aid product conceptual design. *Cluster Comput* 22 (Suppl 4), 10145–10162 (2019). <https://doi.org/10.1007/s10586-017-1190-z>
- Kane, Gerald C. (2022) "How to Write an “A” Paper," *Journal of the Association for Information Systems*, 23(5), 1071-1079. DOI: 10.17705/1jais.00765 Available at: <https://aisel.aisnet.org/jais/vol23/iss5/9>

- Kochenderfer, M. J., Amato, C., Chowdhary, G., How, J. P., Reynolds, H. J., Thornton, J. R., ... & Wheeler, T. A. (2015). *Decision making under uncertainty: theory and application*. MIT press.
- Meyer, T. E., Kiernan, M. S., McManus, D. D., & Shih, J. (2014). Decision-making under uncertainty in advanced heart failure. *Current heart failure reports*, 11, 188-196.
- Papathanasiou, J., Ploskas, N., & Linden, I. (Eds.). (2016). *Real-world decision support systems: Case studies (Vol. 37)*. Springer.
- Peppers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, 24(3), 45-77.
- Prakash, A. V., & Das, S. (2021). Medical practitioner's adoption of intelligent clinical diagnostic decision support systems: A mixed-methods study. *Information & Management*, 58(7), 103524.
- Savage, L. J. (1972). Subjective probabilities. In *The foundations of statistics* (pp. 13-30). Dover Publications.
- Schwartz, A., & Elstein, A. S. (2008). Clinical reasoning in medicine. *Clinical reasoning in the health professions*, 3, 223-234.
- Smith, J., Johnson, L., & Lee, K. (2021). CDSS-RM: A reference model for clinical decision support systems. *Journal of Healthcare Informatics Research*, 5(2), 87-102.
- Struijk, M., Ou, C., Davison, R., & Angelopoulos, S. (2022). Putting the IS back into IS Research. *Information Systems Journal*, 32(3), 1-4. <https://doi.org/10.1111/isj.12368>
- Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., & Gibson, R. N. (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. *Journal of Medical Systems*, 44(4), 1-10. <https://doi.org/10.1007/s10916-020-01556-7>
- Takeda, H., Veerkamp, P., & Yoshikawa, H. (1990). Modeling design process. *AI magazine*, 11(4), 37-37.
- Terris-Prestholt, F., Neke, N., Grund, J. M., Plotkin, M., Kuringe, E., Osaki, H., ... & VMCC study team. (2019). Using discrete choice experiments to inform the design of complex interventions. *Trials*, 20, 1-11.

- Train, K. E. (2009). Risk, uncertainty and discrete choice models. *Marketing Letters*, 19(3-4), 269-285. <https://doi.org/10.1007/s11002-008-9050-8>
- Vasquez, H. M., Pianarosa, E., Sirbu, R., Diemert, L. M., Cunningham, H. V., Donmez, B., & Rosella, L. C. (2022). Human factors applications in the design of decision support systems for population health: a scoping review. *BMJ open*, 12(4), e054330.
- Wang Y, Wang Z, Wang Z, Li X, Pang X and Wang S (2021) Application of Discrete Choice Experiment in Health Care: A Bibliometric Analysis. *Front. Public Health* 9:673698. doi: 10.3389/fpubh.2021.673698
- Wernz, C., Song, Y. & Hughes, D.R. How hospitals can improve their public quality metrics: a decision-theoretic model. *Health Care Manag Sci* 24, 702–715 (2021). <https://doi.org/10.1007/s10729-021-09551-7>



# Temporal Dimensions in Cancer Clinical Decision-Making Through Machine Learning

**Yiyu Wang**

*J.E. Cairnes School of Business and Economics; University of Galway*

**Umair ul Hassan**

*Insight SFI Centre for Data Analytics; J.E. Cairnes School of Business and Economics; University of Galway*

**Anastasia Griva and Kieran Conboy**

*Lero - The Science Foundation Ireland Research Centre for Software, Ireland; J.E. Cairnes School of Business and Economics; University of Galway*

## **Abstract**

*This paper investigates the how time and temporality are considered in machine learning (ML) for medical decision-making through a literature review. It highlights the role of temporal dimensions in machine learning and clinical decision-making for cancer. Preliminary results indicate a lack of explicit attention to temporal dimensions in existing research, albeit their implicit inclusion, underscoring the necessity to incorporate time-related concepts to enhance the decision-making process. Future research can focus on the integration of time framework into ML design to improve clinical decision-making.*

**Keywords:** Machine Learning, Clinical Decision Making, Time, Temporal, Oncology

## **1. Introduction**

Medical decision-making in oncology is a complex process that requires careful consideration of various factors over time (Kourou et al., 2015). The integration of *machine learning* (ML) in this domain has shown promise, but there is a need to scrutinize the role of temporal and time-related characteristics of ML and decision processes more closely (Ancona et al., 2001). This review aims to fill the existing research gap by focusing on how time-related concepts are applied in ML across different stages of medical decision-making. These stages include but are not limited to disease detection, diagnosis, treatment planning, prediction of future medical risks and

treatment outcomes (Banegas-Luna et al., 2021). Focusing on oncology, it also summarizes how temporal dimensions affect medical decision-making.

Time plays a crucial role in the healthcare as it aids in understanding the chronological order of medical procedures. Physicians utilize their medical knowledge to make diagnostic and treatment decisions (Kamišalić et al., 2018). The complexity of temporal dimensions of clinical decision making is often overlooked in the majority of studies (Ancona et al., 2001; de Vaujany et al., 2021; Venkatesh et al., 2021), despite their crucial role in everything from the sequencing and duration of symptoms for diagnosis to the temporal changes in diseases for prognosis, and the ordering and event mapping in treatment plans (Augusto, 2005). Most research tends to focus only on the speed of decisions and simple clock-time as measured in hours and days. Researchers have developed a comprehensive set of dimensions to study time and temporality by categorizing them into three main classes (Ancona et al., 2001; de Vaujany et al., 2021; Venkatesh et al., 2021):

- *Time concepts* relate to how we think about and arrange time, such as traditional clocks and calendars, events, and the social construction of time (O'Connor et al., 2022).
- *Mapping activities* to time, i.e., how activities are arranged, including duration, frequency, synchronization, interruptions, or rhythm of activities.
- *Actors* related to time, i.e., how the individuals involved in these activities exhibit different temporal personalities and their perspectives on time.

While existing studies have highlighted ML's role in oncology, the integration of temporal dimensions remains inadequately explored. As such, the depth to which time mapping affects medical decision-making, especially in cancer, is not fully understood. Based on this, we propose the following two *research questions* (RQ) to study the application of ML in clinical decision-making for oncology, as well as to emphasize the role of introducing temporal dimensions in enhancing clinical decision-making.

- **RQ1:** How is the application of ML to clinical decision-making in oncology discussed in the current literature?

- **RQ2:** What and how temporal dimensions are addressed in the current literature of ML for clinical decision-making?

## **2. Background**

### **2.1 Clinical Decision and Machine Learning (ML)**

Machine learning has deeply permeated the healthcare sector. With its advanced algorithms and extensive applications, it has transformed many aspects of medical practice. It provides significant assistance to medical professionals in several areas such as health information systems, health data analysis, tracking of epidemics, symptom monitoring, predictive modelling, clinical decision support and medical imaging (Bohr & Memarzadeh, 2020).

Individual health is a domain of healthcare which focuses on the health status and needs of each patient. This field is not only concerned with the patient's current health status but also aims to prevent diseases, diagnose existing health problems (Bhardwaj et al., 2017; Secinaro et al., 2021) and provide the corresponding treatment and long-term care (CII, 2019).

Specifically, the application of ML in individual health is multifaceted. In prevention, ML complements routine checkups, vaccinations, and health education by analyzing extensive datasets to identify risk factors, enabling personalized strategies (CII, 2019). In the realm of diagnosis, ML enhances the accuracy of identifying and describing health issues, utilizing tools like laboratory tests, medical imaging, and patient symptoms. During acute treatment, ML aids in swiftly assessing patient data, assisting healthcare providers in making more accurate and timely decisions in critical situations. In terms of monitoring, ML contributes to patient observation, both inpatient and at facilities, by continuously analyzing health data to detect anomalies and anticipate potential issues. Finally, in follow-up and chronic treatment, ML supports ongoing management strategies, optimizing medication regimens and lifestyle modifications, and ensuring regular medical checkups are as effective as possible.

## 2.2 Time and Temporality

A significant question in the theory of time and temporality is determining which information will undergo change; what events can cause changes, and how these changes occur in a temporal sequence (Augusto, 2005). Upon delving into the literature on time reasoning in healthcare, it becomes evident that many articles might not directly discuss the notion of temporality. However, these concepts are implicitly present. For instance, when researchers discuss the sequence of an event's occurrence, they are essentially discussing the temporal relationships between events, even if they might not explicitly use the word “time”. Such concepts often lie latent within key dimensions, which might exhibit cyclical or sequential properties and map to the timeline of events. In a patient’s medical history, events can cause changes in states of information, such as biomarkers or genetic mutations, and these changes also occur in a temporal sequence (Augusto, 2005; Bhinder et al., 2021). Delving further into the literature on time reasoning and its application in ML for oncology, the complexity of temporal dimensions appears often underexplored (Ancona et al., 2001; de Vaujany et al., 2021; Venkatesh et al., 2021). While many studies might not directly discuss the notion of temporality, these concepts are implicitly present and crucial for improving the accuracy and efficacy of cancer diagnoses and prognoses.

## 3. Methodology

We conducted a literature review that delves into the consideration of temporal aspects of ML when applied to decision-making systems for oncology. The aim was to provide an overview of current applications of ML techniques by answering the key RQs and exploring the use of time-related and temporal concepts.

Category	Keywords
Machine Learning (ML)	Machine Learning, Deep Learning, Reinforcement Learning, Natural Language Processing, Supervised Learning, Unsupervised Learning, Neural Networks, Image Recognition, Computer Vision

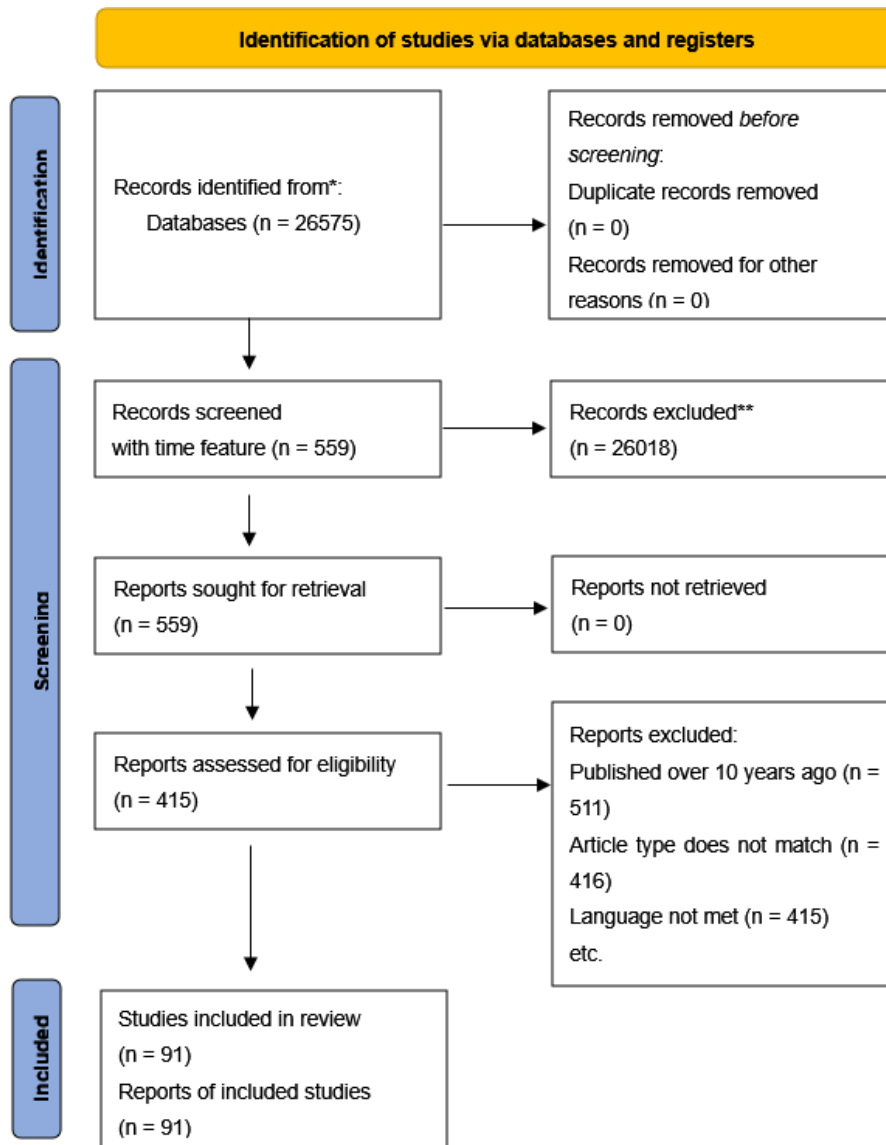
Clinical Decision-Making	Decision, Clinical, Prognosis, Diagnosis
Time-Related dimensions	Time, Temporal, Sequences
Tumor	Cancer, Oncology, Tumor

**Table 1. Categorized search keywords.**

Research articles for this review were identified through a combination of keywords related to ML, clinical decision-making, tumors, and time, as listed in Table 1. These articles were specifically selected from the core collection of journals in the Web of Science (WoS) database.

#### **4. Preliminary Results**

As illustrated in Figure 1, through the detailed search and the subsequent screening process, we identified 91 relevant publications. Based on Ancona's framework of time and temporality (Ancona et al., 2001), we systematically categorized these 91 publications to gain a deeper understanding of the temporal dimensions across different research studies.



**Figure 1. Literature selection process based on PRISMA flow diagram.**

The selected publications were preliminarily categorized into several temporal dimensions, including the time efficiency of clinical data annotation (Wang, Song, et al., 2022), survival prediction (Doppalapudi et al., 2021), time-to-event modelling (Danciu et al., 2022), time series analysis (Dorraki et al., 2019), the temporal application of treatment planning (Kalakoti et al., 2021), time prediction and decision-making (Saito et al., 2023), data acquisition and fusion at multiple time points (Li et al., 2020), as well as prediction and longitudinal data analysis under a temporal framework (Nitski et al., 2021). Table 2 presents the research-in-progress results.

Temporal Dimensions Name	Number of Articles
Clinical Data Annotation Time Efficiency	3

Survival Projection	14
Time-to-Event Modeling	17
Time Series Analysis	9
Application of Time in Treatment Planning	3
Temporal Prediction and Decision-Making in Patient Treatment	12
Data Acquisition & Fusion at Multiple Time Points	5
Time Framework of Prediction and Longitudinal Data	11
Time-Related Tumoral Biological Processes and Pathways	6

**Table 2. Matrix of temporal dimensions and corresponding number of articles**

#### **4.1 Temporal Prediction and Decision-Making in Patient Treatment**

Temporal prediction plays a pivotal role in patient treatment. The preliminary results of the review indicate that through ML and other advanced techniques, we can more accurately predict a patient's survival time, first treatment time, and other key events related to treatment (Lombardo et al., 2021). This offers doctors more precise and targeted treatment suggestions.

*Single Activity Transformation Mapping (aa')* (Ancona et al., 2001): In the context of diseases like *Chronic Lymphocytic Leukaemia (CLL)* (Mosquera Orgueira et al., 2019) and *Glioblastoma Multiforme (GBM)* (Zhu et al., 2023), life cycles and midpoint transitions are predicted that can influence the treatment plans.

*Actors Relating to Time (Temporal Perception)*: Clinicians and patients may experience time differently which may alter their perception of treatment duration based on ML predictions (Xu et al., 2022).

ML methods were utilized to predict the overall survival time for GBM patients, thus offering more tailored treatment suggestions (Tang et al., 2020). Similarly, ML methods were used to predict the time of first treatment for CLL patients and increased the interpretability of the predictions through unsupervised clustering techniques (Chen et al., 2019). Additionally, a prognostic prediction model was established for prostate cancer patients undergoing androgen deprivation therapy (Saito et al., 2023).

#### **4.2 Data Acquisition & Fusion at Multiple Time Points**

The key role of collecting and fusing data at multiple time points in medical imaging analysis by capturing colposcopy images at different points during an acetic acid test and using deep learning techniques for feature encoding and fusion, the importance of time-dimension data in the diagnosis of cervical cancer was emphasized (Li et al., 2020).

*Repeated Activity Mapping (aa)*: The rhythm and frequency of data acquisition, like capturing colposcopy images at regular intervals, play a crucial role in tracking disease progression.

*Actors Relating to Time (Temporal Perception)*: Clinicians' experience of time can influence the interpretation and diagnosis based on time-point data (Boers et al., 2020).

Similarly, a time-difference technique was employed to align and compare chest CT scans at two distinct time points, enabling the early identification of lung anomalies (Nagao et al., 2018). Both these studies demonstrate that leveraging data from multiple time points can offer a more comprehensive and accurate portrayal of disease progression, ensuring better medical care for patients.

Through such categorization, we can more clearly see the diversity of the temporal factors involved in ML and clinical decision-making. Although we have classified and analyzed the relevant literature, limitations in the number and quality of retrieved studies mean that we have not been able to capture all relevant research. Some studies that may implicitly involve temporal concepts but do not explicitly mention time may have been overlooked. These missed studies might contain insights that have the potential to further refine the temporal framework.

## **5. Conclusion and Future Work**

Based on our preliminary review and classification of the relevant literature, we recognize that researchers exploring clinical decision-making in oncology have already discussed multiple temporal dimensions, including annotation efficiency, survival prediction, time-to-event modelling, time series analysis, the application of time in



treatment planning, time prediction and decision-making, acquisition and fusion of multi-timepoint data, as well as prediction and longitudinal data analysis within a temporal framework.

This integrated approach enhances our understanding of ML by incorporating the relevant aspects of Ancona's framework, thereby providing a more holistic understanding of the role of time in medical treatment and data analysis.

The next step for extending this literature review will be an in-depth analysis of the integration of temporal dimensions in ML models for clinical decision-making in cancer research while focusing on what potential challenges and limitations might arise, and how could these be addressed to optimize model performance and improve outcomes of healthcare.

## 6. References

- Ancona, D. G., Okhuysen, G. A., & Perlow, L. A. (2001). Taking time to integrate temporal research. *Academy of Management Review*, *26*(4), 512–529. Scopus. <https://doi.org/10.5465/AMR.2001.5393887>
- Augusto, J. C. (2005). Temporal reasoning for decision support in medicine. *Artificial Intelligence in Medicine*, *33*(1), 1–24. <https://doi.org/10.1016/j.artmed.2004.07.006>
- Banegas-Luna, A. J., Peña-García, J., Iftene, A., Guadagni, F., Ferroni, P., Scarpato, N., Zanzotto, F. M., Bueno-Crespo, A., & Pérez-Sánchez, H. (2021). Towards the Interpretability of Machine Learning Predictions for Medical Applications Targeting Personalised Therapies: A Cancer Case Survey. *International Journal of Molecular Sciences*, *22*(9), Article 9. <https://doi.org/10.3390/ijms22094394>
- Bhardwaj, R., Nambiar, A. R., & Dutta, D. (2017). A Study of Machine Learning in Healthcare. *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)*, *2*, 236–241. <https://doi.org/10.1109/COMPSAC.2017.164>
- Bhinder, B., Gilvary, C., Madhukar, N. S., & Elemento, O. (2021). Artificial Intelligence in Cancer Research and Precision Medicine. *Cancer Discovery*, *11*(4), 900–915. <https://doi.org/10.1158/2159-8290.CD-21-0090>
- Boers, T., van der Putten, J., Struyvenberg, M., Fockens, K., Jukema, J., Schoon, E., van der Sommen, F., Bergman, J., & de With, P. (2020). Improving Temporal Stability and Accuracy for Endoscopic Video Tissue Classification Using Recurrent Neural Networks. *SENSORS*, *20*(15), 4133. <https://doi.org/10.3390/s20154133>
- Bohr, A., & Memarzadeh, K. (2020). The rise of artificial intelligence in healthcare applications. *Artificial Intelligence in Healthcare*, 25–60.

<https://doi.org/10.1016/B978-0-12-818438-7.00002-2>

- Chen, D., Goya, G., Go, R. S., Parikh, S. A., & Ngufor, C. G. (2019). Improved Interpretability of Machine Learning Model Using Unsupervised Clustering: Predicting Time to First Treatment in Chronic Lymphocytic Leukemia. *JCO CLINICAL CANCER INFORMATICS*, 3, 1–11. <https://doi.org/10.1200/CCI.18.00137>
- Danciu, I., Agasthya, G., Tate, J. P., Chandra-Shekar, M., Goethert, I., Ovchinnikova, O. S., McMahon, B. H., & Justice, A. C. (2022). In with the old, in with the new: Machine learning for time to event biomedical research. *JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION*, 29(10), 1737–1743. <https://doi.org/10.1093/jamia/ocac106>
- de Vaujany, F.-X., Leclercq-Vandelannoitte, A., Munro, I., Nama, Y., & Holt, R. (2021). Control and Surveillance in Work Practice: Cultivating Paradox in ‘New’ Modes of Organizing. *Organization Studies*, 42(5), 675–695. <https://doi.org/10.1177/01708406211010988>
- Doppalapudi, S., Qiu, R. G., & Badr, Y. (2021). Lung cancer survival period prediction and understanding: Deep learning approaches. *INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS*, 148, 104371. <https://doi.org/10.1016/j.ijmedinf.2020.104371>
- Dorraki, M., Fouladzadeh, A., Salamon, S. J., Allison, A., Coventry, B. J., & Abbott, D. (2019). Can C-reactive protein (CRP) Time Series Forecasting be Achieved via Deep Learning? *IEEE ACCESS*, 7, 59311–59320. <https://doi.org/10.1109/ACCESS.2019.2914473>
- Kalakoti, Y., Yadav, S., & Sundar, D. (2021). SurvCNN: A Discrete Time-to-Event Cancer Survival Estimation Framework Using Image Representations of Omics Data. *CANCERS*, 13(13), 3106. <https://doi.org/10.3390/cancers13133106>
- Kamišalić, A., Riaño, D., & Welzer, T. (2018). Formalization and acquisition of temporal knowledge for decision support in medical processes. *Computer Methods and Programs in Biomedicine*, 158, 207–228. <https://doi.org/10.1016/j.cmpb.2018.02.012>
- Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13, 8–17. <https://doi.org/10.1016/j.csbj.2014.11.005>
- Lee, C., Rashbass, J., & Van der Schaar, M. (2021). Outcome-Oriented Deep Temporal Phenotyping of Disease Progression. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, 68(8), 2423–2434. <https://doi.org/10.1109/TBME.2020.3041815>
- Li, H., Robinson, K., Lan, L., Baughan, N., Chan, C.-W., Embury, M., Whitman, G. J., El-Zein, R., Bedrosian, I., & Giger, M. L. (2023). Temporal Machine Learning Analysis of Prior Mammograms for Breast Cancer Risk Prediction. *CANCERS*, 15(7), 2141. <https://doi.org/10.3390/cancers15072141>
- Li, Y., Chen, J., Xue, P., Tang, C., Chang, J., Chu, C., Ma, K., Li, Q., Zheng, Y., & Qiao, Y. (2020). Computer-Aided Cervical Cancer Diagnosis Using Time-

- Lapsed Colposcopic Images. *IEEE TRANSACTIONS ON MEDICAL IMAGING*, 39(11), 3403–3415. <https://doi.org/10.1109/TMI.2020.2994778>
- Li, Z., Li, C., Long, Y., & Wang, X. (2020). A system for automatically extracting clinical events with temporal information. *BMC MEDICAL INFORMATICS AND DECISION MAKING*, 20(1), 198. <https://doi.org/10.1186/s12911-020-01208-9>
- Lombardo, E., Kurz, C., Marschner, S., Avanzo, M., Gagliardi, V., Fanetti, G., Franchin, G., Stancanello, J., Corradini, S., Niyazi, M., Belka, C., Parodi, K., Riboldi, M., & Landry, G. (2021). Distant metastasis time to event analysis with CNNs in independent head and neck cancer cohorts. *SCIENTIFIC REPORTS*, 11(1), 6418. <https://doi.org/10.1038/s41598-021-85671-y>
- Mosquera Orgueira, A., Antelo Rodriguez, B., Alonso Vence, N., Bendana Lopez, A., Diaz Arias, J. A., Diaz Varela, N., Gonzalez Perez, M. S., Perez Encinas, M. M., & Bello Lopez, J. L. (2019). Time to Treatment Prediction in Chronic Lymphocytic Leukemia Based on New Transcriptional Patterns. *FRONTIERS IN ONCOLOGY*, 9, 79. <https://doi.org/10.3389/fonc.2019.00079>
- Nagao, M., Miyake, N., Yoshino, Y., Lu, H., Kim, H., Murakami, S., Aoki, T., & Kido, S. (2018). Detection of Abnormal Shadows on Temporal Subtraction Images Based on Multi-phase CNN. *2018 18TH INTERNATIONAL CONFERENCE ON CONTROL, AUTOMATION AND SYSTEMS (ICCAS)*, 1333–1337. <https://www.webofscience.com/wos/woscc/full-record/WOS:000457612300204>
- Nitski, O., Azhie, A., Qazi-Arisar, F. A., Wang, X., Ma, S., Lilly, L., Watt, K. D., Levitsky, J., Asrani, S. K., Lee, D. S., Rubin, B. B., Bhat, M., & Wang, B. (2021). Long-term mortality risk stratification of liver transplant recipients: Real-time application of deep learning algorithms on longitudinal data. *LANCET DIGITAL HEALTH*, 3(5), E295–E305. [https://doi.org/10.1016/S2589-7500\(21\)00040-6](https://doi.org/10.1016/S2589-7500(21)00040-6)
- O'Connor, M., Conboy, K., & Dennehy, D. (2022). Time is of the essence: A systematic literature review of temporality in information systems development research. *Information Technology & People*, 36(3), 1200–1234. <https://doi.org/10.1108/ITP-11-2019-0597>
- Saito, S., Sakamoto, S., Higuchi, K., Sato, K., Zhao, X., Wakai, K., Kanesaka, M., Kamada, S., Takeuchi, N., Sazuka, T., Imamura, Y., Anzai, N., Ichikawa, T., & Kawakami, E. (2023). Machine learning predicts time-series prognosis factors in metastatic prostate cancer patients treated with androgen deprivation therapy. *SCIENTIFIC REPORTS*, 13(1). <https://doi.org/10.1038/s41598-023-32987-6>
- Secinaro, S., Calandra, D., Secinaro, A., Muthurangu, V., & Biancone, P. (2021). The role of artificial intelligence in healthcare: A structured literature review. *BMC Medical Informatics and Decision Making*, 21(1), 125. <https://doi.org/10.1186/s12911-021-01488-9>
- Tang, Z., Xu, Y., Jin, L., Aibaidula, A., Lu, J., Jiao, Z., Wu, J., Zhang, H., & Shen, D. (2020). Deep Learning of Imaging Phenotype and Genotype for Predicting Overall Survival Time of Glioblastoma Patients. *IEEE TRANSACTIONS ON MEDICAL IMAGING*, 39(6), 2100–2109.

<https://doi.org/10.1109/TMI.2020.2964310>

- The Rockefeller Foundation and the United States Agency for International Development's (USAID) Center for Innovation and Impact (CII). (2019). *Artificial Intelligence in Global Health: Defining a Collective Path Forward*. [https://www.usaid.gov/sites/default/files/2022-05/AI-in-Global-Health\\_webFinal\\_508.pdf](https://www.usaid.gov/sites/default/files/2022-05/AI-in-Global-Health_webFinal_508.pdf)
- Venkatesh, V., Sykes, T., Aljafari, R., & Scott Poole, M. (2021). *The Future is Now: Calling for a Focus on Temporal Issues in Information Systems Research* (SSRN Scholarly Paper 3958479). <https://papers.ssrn.com/abstract=3958479>
- Wang, Y., Lombardo, E., Avanzo, M., Zschaek, S., Weingaertner, J., Holzgreve, A., Albert, N. L., Marschner, S., Fanetti, G., Franchin, G., Stancanello, J., Walter, F., Corradini, S., Niyazi, M., Lang, J., Belka, C., Riboldi, M., Kurz, C., & Landry, G. (2022). Deep learning-based time-to-event analysis with PET, CT and joint PET/CT for head and neck cancer prognosis. *COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE*, 222, 106948. <https://doi.org/10.1016/j.cmpb.2022.106948>
- Wang, Y., Song, D., Wang, W., Rao, S., Wang, X., & Wang, M. (2022). Self-supervised learning and semi-supervised learning for multi-sequence medical image classification. *NEUROCOMPUTING*, 513, 383–394. <https://doi.org/10.1016/j.neucom.2022.09.097>
- Xu, Z., Lim, S., Shin, H.-K., Uhm, K.-H., Lu, Y., Jung, S.-W., & Ko, S.-J. (2022). Risk-aware survival time prediction from whole slide pathological images. *SCIENTIFIC REPORTS*, 12(1), 21948. <https://doi.org/10.1038/s41598-022-26096-z>
- Zhu, J., Ye, J., Dong, L., Ma, X., Tang, N., Xu, P., Jin, W., Li, R., Yang, G., & Lai, X. (2023). Non-invasive prediction of overall survival time for glioblastoma multiforme patients based on multimodal MRI radiomics. *INTERNATIONAL JOURNAL OF IMAGING SYSTEMS AND TECHNOLOGY*, 33(4), 1261–1274. <https://doi.org/10.1002/ima.22869>

# Teaching and Learning through the pandemic; the effects of remote work on women academics

**Maria Vardaki**

*Royal Holloway University of London, Matm015@live.rhul.ac.uk*

**Eleni Tzouramani**

*University of the West of Scotland, Eleni.Tzouramani@uws.ac.uk*

**Niki Panteli**

*University of Lancaster, n.panteli1@lancaster.ac.uk*

## **Abstract**

*During the pandemic, the Higher Education had to rapidly adapt to new ways of working with both staff and students having to swiftly work in a remote environment for a prolonged period. This research is focusing on how women in academia within the Higher Education environment in the UK have adapted to virtual working, online teaching and learning and in what way they were affected by the pandemic in their various professional and home life. The research approach is to investigate women in academia perspectives during the transitional period from face-to-face onto e-learning and the mechanisms that they have used, in order to endure the challenges of work and everyday life.*

**Keywords:** Remote work, pandemic, women academics, Higher Education, support, teaching

## **1.0 Introduction**

Higher Education institutions were compelled to rapidly move all their operations to a virtual environment, at the start of the Covid-19 pandemic lockdown. This has had an effect on technological developments, as institutions were forced to abruptly utilise platforms in a way, never used before, for specific functions such as teaching and learning, examinations, research activities, conferences, training activities, virtual and hybrid meetings. This has proven to be a new way of working which can be applied within a crisis situation in future. The transition to online teaching and learning due to the pandemic has led to a wider conversation on flexible education amongst the higher education institutions. This is creating opportunities for innovation and setting new boundaries for knowledge exchange (teaching and learning) and remote working.

It is inevitable that this virtual working era could generate the circumstances for new legislation regarding the processes and technology used for online learning (Basilaia

& Kvavadze, 2020) and for remote working in general. The concept of a widespread online learning during a crisis brings along challenges in the quality of teaching and learning, student engagement and the specific needs of the students.

## **2.0 Effects on academics and students**

The covid restrictions have affected how teaching and learning is conducted but it is a cause of reflection (Chertoff et al, 2020), in order to adapt and ultimately adopt innovative teaching methods. The rapid arrangements that were put in place for training on academic delivery of hybrid teaching and professional services staff supporting students remotely has become the norm in recent years. Academics and students had to quickly adapt with the majority of the practical elements of their studies to be restricted or cancelled, such as travelling, fieldwork, on some occasions, lab work and practical sessions (ibid, 2020). In the meantime, other practical subjects such as health and medical studies, where face-to-face interaction is necessary, have had to be adapted to virtual teaching while attempting to keep students engaged remotely.

Both staff and students would need to have some experience of using online platforms, in order to be able to use them positively for their learning (Aguilera-Hermida, 2020). The aspect of individual behaviour towards the technology and online teaching and learning is one that can be varied in people, and this could affect all aspects of engagement, resilience and motivation (Kemp et al, 2019). Aguilera-Hermida's research has identified that students prefer the face-to-face over online learning due to a variety of reasons in relation to lack of resources and support in an online environment. This is not an unexpected result due to the circumstances that the shift to online learning occurred and how prepared higher education institutions were with the infrastructure in place (or not). There is scope and room for improvement, as both parties, students and academics, had to make the transition simultaneously, without any prior preparedness and user testing practices. This experience has equipped them with knowledge on usage of technological tools and platforms to use in future (Murphy, 2020).

The effects on women academics, in particular, were evident, as they have had to find ways to cope during the lockdowns, in the midst of challenging situations they were potentially facing while working from home (Minello, 2020; Lutter & Schröder, 2020; Oleschuk, 2020). It is apparent that women academics have suffered increased pressure during this period and the situation deteriorated when caring responsibilities, increased housework and high levels of emotional stress were affecting them. Even before the pandemic, the increased use of technology in academic work, had contributed to blurring the boundaries between work and home, especially affecting academics with young children (Currie & Eveline, 2011) but the lockdown during the pandemic, intensified all this and increased academics' work – home conflict (Adisa et al. 2022).

The shift to online teaching and learning has impacted students' anxiety levels, depending on their demographics (Jehi et al, 2022), similarly to academics, especially women (Augustus, 2021). The elements of home infrastructure, social and academic life, wellbeing, and institutional support (Aristovnik et al, 2020) were crucial in the process of rapid change, notably during the first lockdown. For a number of academics, the lockdown has created efficiencies in the way they had to work from home and used their research time, in spite of the obstacles they may have faced due to the use of new technology and new ways of work (Acze, et al, 2021). Support from universities in the UK was varied and in some cases, there were some “inter-sectional impacts of the pandemic on minority staff, which consequently prevent the achievement of equity (Blell et al, 2023).

### **3.0 Technological effects**

Prior to the Covid pandemic, counter arguments of technological advances would focus on how the use of technology cannot replace face-to-face communications, the extreme reliance of data can be misused and that all employees must have access to such systems, in order for them to work effectively (Varma & Budhwar, 2011). The rapid shift to online and remote working has used, not only pre-existing technological advances but employers had to create new ones in order to identify ways of staff working together from a distance and in an effective way, in order to ensure that they

are beneficial and accurate to serve their purpose. Remote working for women academics has had a negative effect, according to Marchiori et al.'s (2019) research, due to the widespread and constant exposure to technology, thus creating higher levels of techno-complexity and techno-uncertainty than men.

Stoller (2021) has researched how the pandemic can provide opportunities and conducted a SWOT analysis for higher education institutions. His research has identified opportunities such as strategies to improve virtual teaching and learning, interface between different study levels, leadership opportunities within the crisis, as well as enforcing resilience and compassion. The focus on virtual participation in meetings and the usage of software to allow staff to attend, made it easier in some respects to promote inclusion (Nguyen et al, 2020). Stoller (2021) also referred to certain disadvantages in relation to home working, such as the individual conditions at home (Aguilera-Hermida, 2020) and the loss of face-to-face interaction which may affect training, recruitment, student/academic interaction, equipment issues (ibid, 2020) and virtual meeting “fatigue”.

One of the consequences of virtual education is that it minimises interaction and networking (Nguyen et al, 2020), with financial implications for universities in a variety of levels, from housing to tuition fees and recruitment of students. Stoller concludes his research by making suggestions on using best practises to develop virtual opportunities for expansion and development, such as suggestions for academic on teaching and assessment, for students on virtual interviews, for professional services staff to acquire extended knowledge on most virtual platforms, to review space and working patterns and communication and support for students and staff.

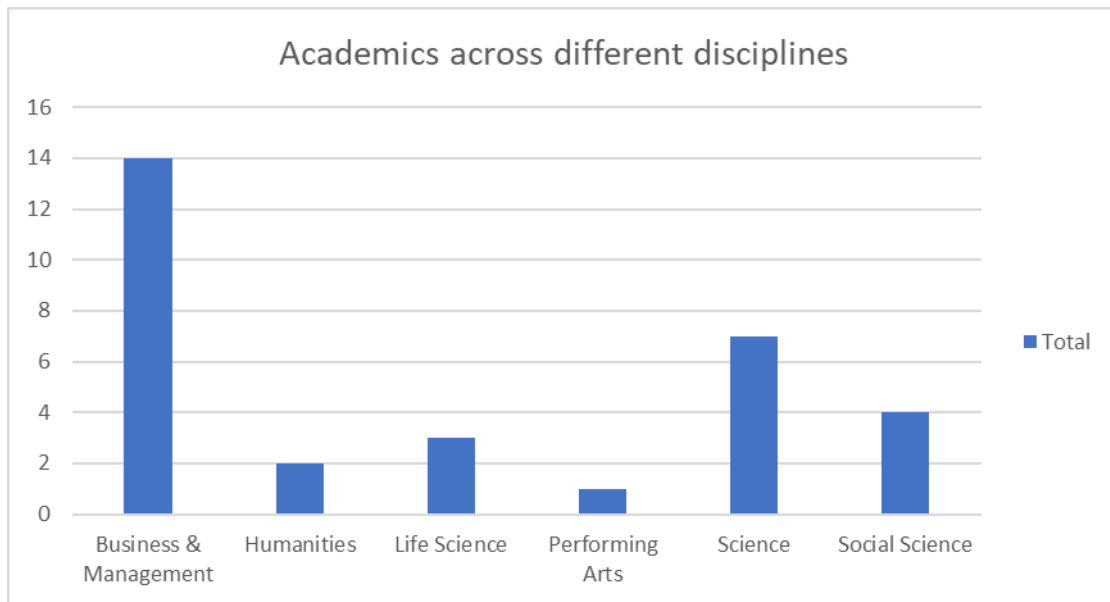
The transition to online learning has brought significant changes to academics, not only in relation to the online environment but in relation to the flexibility and willingness to adapt and change their teaching practises (Quezada et al, 2020). Academics and educators are required, even prior to the pandemic, to keep themselves informed of recent developments and technological advances in teaching and learning, as the demands from the students increase every year. The usage of technology in teaching and learning has become a necessity recently, so academics who were using



traditional methods of teaching could be potentially battling to adapt to virtual learning due to its nature and complexity of the environment. The speed in which the changes had to be made and the type of resources available gave little time for academics to familiarise themselves with new technologies.

#### **4.0 Methodology**

This research was conducted one year after the latest lockdown within HE institutions in the UK. Through an inductive approach, we identified a gap on how women academics experienced the period of the pandemic with a focus on teaching and learning. Academics have adapted their work processes, teaching methods and home working environment from the early stages of the first lockdown and up to recently, in order to work successfully and fulfil their duties towards the institution and realise their responsibilities towards students and colleagues in the midst of the academic year. We have used semi-structured interviews to collect data, so this research is focusing on one qualitative methods; 31 women academics from a variety of disciplines were interviewed; ethical approval was confirmed full information shared with the participants. The sampling method was initially convenience in order to identify academics from institutions familiar to the authors, but we also used the snowballing technique for access to academics from further institutions. The method used for this research is qualitative analytic, and specifically using thematic analysis (Braun & Clarke, 2006) to identify the themes and be able to draw conclusions based on the narrative.



**Figure 1.** Chart showing participants of the study from across different disciplines.

## 5.0 Preliminary Findings

### 5.1 Collegiality

The notion of academic colleagues working together and supporting each other is one that has been evident throughout the interview process. We have looked to explore the concept of support, whether this is support from the Universities, the Schools or departments and support from individual colleagues. The perception of collegiality was evident throughout the interview process, as participants reported productive cooperation between colleagues, not only the ones who may share similar responsibilities in a group, but also colleagues working together in different areas.

### 5.2 University Support

A number of academics have reported that what was extremely concerning for them was the support they felt they had to provide to students during this time, which meant on several occasions, emotional support to students out of “office” working hours. This has had an effect on the emotional wellbeing of the academics themselves, as they discovered they almost had to resume the role of counsellors to be able to assist and support students who were struggling psychologically during the lockdowns. They have reported that no specific training was provided for them from their own

institutions to be able to perform such a role and the ones who got drawn into this type of support, were negatively affected from the experience.

### **5.3 Work/life balance**

The home environment for single academics was quickly adapted to suit the working day, as opposed to academics with caring responsibilities (Collins et al, 2020), but the majority of the interviewees have admitted that the boundaries between work and home life were confused, especially in the first lockdown. In addition, the support academics were required to provide to students resulted in blurring the boundaries of work and home life, especially when students needed additional support to be able to cope with the pressures of the remote learning environment. In addition, the participants of the research had a notion of productivity that can be measured in terms of teaching, research, administrative duties and future effects of their productivity and career progression (Myers et al, 2020). They believe that their productivity has suffered (Gultom, 2021) due to the additional responsibilities of the teaching material they had to organise for students and the support they had to provide, which effectively resulted in long working hours from home (Amano-Patino et al, 2020).

## **6.0 Conclusion**

The effects of the pandemic can be visible amongst students and academics, especially during the lockdown period and home-working era. The issue of support is a crucial one in relation to isolation and home working, for all higher education employees, as the duration of the pandemic may have affected people in different ways (Salah & Al-Doghmi, 2023). In cases where academics were digitally advanced, this can be seen as the opportunity to be innovative and develop teaching and learning environments to improve remote student experience.

Through our research, we have identified that the issue of the response towards the pandemic as a crisis has been translated differently by the participants, according to their individual circumstances. A variety of participant responses have shown that universities have shown their support by offering information and practical solutions, such as extra days of annual leave. Responses at times of crisis, and preparedness can

take different forms and shapes. However, based on our preliminary findings, it seems a challenge to prepare for the unknown, especially when people or a group of individual's circumstances are so different that it cannot fit a specific model.

## References

- Aczel B, Kovacs M, van der Lippe T, Szaszi B (2021), Researchers working from home: Benefits and challenges. *PLOS ONE* 16(3).
- Adisa, T. A., Antonacopoulou, E., Beauregard, T. A., Dickmann, M., and Adekoya, O. D. (2022), Exploring the impact of COVID-19 on employees' boundary management and work-life balance, *British Journal of Management*, 33(4), 1694-1709.
- Aguilera-Hermida, A. P. (2020), College students' use and acceptance of emergency online learning due to COVID-19, *International Journal of Educational Research Open*, Vol 1, 100011.
- Amano-Patiño, N., Faraglia, E., Giannitsarou, C., and Hasna, Z., (2020), Who is doing new research in the time of COVID-19? Not the female economists - Publishing and measuring success in economics, Edited by S. Galiani and U. Panizza, CEPR Press, London.
- Aristovnik A, Keržič D, Ravšelj D, Tomaževič N, Umek L. (2020), Impacts of the COVID-19 Pandemic on Life of Higher Education Students: A Global Perspective. *Sustainability*. 2020; 12(20): 8438.
- Augustus, J. (2021), The impact of Covid -19 pandemic on women working in Higher education, *Frontiers in Education*, 14 May, Sec. Educational Psychology, Vol 6.
- Basilaia, G., & Kvavadze, D. (2020). Transition to Online Education in Schools during a SARS-CoV-2 Coronavirus (COVID-19) Pandemic in Georgia. *Pedagogical Research*, 5(4), em0060.
- Blell, M., Liu, S-J S., and Verma, A. (2023), Working in unprecedented times: Intersectionality and women of color in UK higher education in and beyond the pandemic, *Gender, Work & organisation*, 30:2, 353-372.
- Mwenz Blell, Shan-Jan Sarah Liu, Audrey Verma
- Braun, V and Clarke, V (2006), Using thematic analysis in psychology, *Qualitative Research in Psychology*, 3, 77-101.
- Chertoff, J. D., Zarzour, J. G., Morgan, D. E., Lewis, P. J., Canon, C. L., Harvey, J. A., (2020), The Early Influence and Effects of the Coronavirus Disease 2019 (COVID-19) Pandemic on Resident Education and Adaptations, *Journal of the American College of Radiology*, 17: 10, 1322-1328.
- Collins, C., Landivar L. C., Ruppanner L., and Scarborough W. J., (2021), COVID-19 and the gender gap in work hours, *Gender, Work & Organisation*, 28:S1 Supplement: Feminist Frontiers, January 2021, 101-112.
- Currie, J., & Eveline, J. (2011), E-technology and work/life balance for academics with young children, *Higher Education*, 62, 533-550.
- Gultom, E., (2021), The Effect of Working from Home on The Productivity of Female Lecturers during Covid-19 Pandemic at Private Universities in Indonesia, *ADPEBI International Journal of Business and Social Science*, 1(1), 53-63.
- Jehi T, Khan R, Dos Santos H, Majzoub N. (2022), Effect of COVID-19 outbreak on anxiety among students of higher education; A review of literature., *Curr Psychol.*, Jan 7:1-15.

- Kemp, A., Palmer, E. and Strelan, P., (2019), A taxonomy of factors affecting attitudes towards educational technologies for use with technology acceptance models, *British Journal of Educational Technology*, 50 (5), 2394-2413.
- Lutter, M., & Schröder, M. (2020). Is there a motherhood penalty in academia? The gendered effect of children on academic publications in German sociology, *European Sociological Review*, 36(3), 442–459.
- Marchiori, D. M., Mainardes, E. W., and Rodrigues, R. G. (2019), Do Individual Characteristics Influence the Types of Technostress Reported by Workers? *International Journal of Human–Computer Interaction*, 35(3), 218-230.
- Minello, A. (2020), “The Pandemic and the Female Academic.” *Nature*. Retrieved July 14, 2020, Available from: <https://www.nature.com/articles/d41586-020-01135-9>, [Accessed on 04 June 2021].
- Murphy, M. P. A., (2020), COVID-19 and emergency eLearning: Consequences of the securitization of higher education for post-pandemic pedagogy, *Contemporary Security Policy*, 41:3, 492-505.
- Myers, K. R., Tham, W. Y., Yin, Y., Cohodes, N., Thursby, J. G., Thursby, M. C., ... & Wang, D. (2020), Unequal effects of the COVID-19 pandemic on scientists, *Nature Human Behaviour*, 4(9), 880-883.
- Nguyen, K.D., Enos, T., Vandergriff, T., Vasquez, R., Cruz, P.D., Jacobe, H.T. and Mauskar, M.M., (2020), Opportunities for education during the COVID-19 pandemic, *JAAD international*, 1(1), 21-22.
- Oleschuk, M., (2020), Gender Equity Considerations for Tenure and Promotion during COVID-19, *Canadian Review of Sociology*, August 2020, 57:3, 502-515.
- Quezada, R. L., Talbot, C., & Quezada-Parker, K. B., (2020) From Bricks and Mortar to Remote Teaching: A Teacher Education Program’s Response to COVID-19, *Journal of Education for Teaching*, 46:4, 472-483.
- Salah, R., Al-Doghmi, N. (2023). 'The Impact of the Covid-19 Pandemic and Lockdown on Female University Employees: A Sociolinguistic Case Study', *International Journal of Society, Culture & Language*, 11(2 (Themed Issue on Language, Discourse, and Society)), 12-27.
- Stoller, J. K., (2021), A Perspective on the Educational “SWOT” of the Coronavirus Pandemic, *Chest*, 159: 2, 743-748.
- Varma, A and Budhwar, P S (2011), *Global Performance Management*, in Harzing, A-W and Pinnington, A H (eds), *International Human Resource Management*, 440-467, Sage Publications Ltd, London.

# Responsible AI Principles: Findings from an Empirical Study on Practitioners' Perceptions

**Pouria Akbarighatar**

*Department of Information Systems, University of Agder*

**Ilias O. Pappas**

*Department of Computer Science, Norwegian University of Science and Technology*

*Department of Information Systems, University of Agder*

**Polyxeni Vassilakopoulou**

*Department of Information Systems, University of Agder*

**Sandeep Purao**

*Department of Computer Information Systems, Bentley University*

*Department of Information Systems, University of Agder*

## **Abstract**

*As artificial intelligence (AI) continues its rapid evolution, ethical considerations become increasingly critical. This study presents an analytical approach to assessing the perceived importance, alignment, and implementation of Responsible AI (RAI) principles within organizations. An extensive survey collected insights from 82 AI experts across industries. The empirical data collected revealed interesting and diverse adherence patterns. The findings indicate a moderate overall level of operationalization, with variations across principles. Adherence is categorized into three groups, identifying strengths, weaknesses, and potential areas of overemphasis. Through iterations, organizations can refine their AI initiatives to align with evolving stakeholder expectations. This study contributes novel insights on the role of Responsible AI (RAI) principles within organizations in practice. However, it is essential to acknowledge limitations tied to subjectivity, sample representativeness, AI's dynamic nature, and the need for external validation. This research offers invaluable insights for translating RAI principles into operationalized practices.*

**Keywords:** *Responsible AI, Principles, Operationalized RAI, Perceptions*

## **1. Introduction**

Industries all over the world are adopting AI applications, which are extending into diverse fields like transportation, agriculture, healthcare, and security. For instance, AI is assisting with crop yield optimisation, diagnosing and treating illnesses, catching distracted drivers to improve road safety, detecting credit card fraud to protect finances, and identifying at-risk children to help provide support (Amugongo et al., 2023; Ho et al., 2019; Stilgoe, 2018; Van Esch et al., 2019; Wall, 2018). It is crucial to think about the ramifications and make sure that the development and application of AI proceed in a way that benefits both individuals and

society as these technologies continue to advance and become more integrated into our daily lives and key services and establish long-term sustainability (Clarke, 2019; Pappas et al., 2023; Vassilakopoulou et al., 2022).

While AI offers potential benefits, significant challenges also exist that must be addressed. There are risks of reinforcing unjust biases, violating privacy rights, and propagating false information online. Job displacement and reduced skills demand are also concerning (Mikalef et al., 2022). Additionally, mass surveillance, critical system failures involving autonomous technologies, and weapons applications pose risks. Even AI meant to help, like improving cybersecurity, could enable malicious uses such as cyberattacks if misapplied. These types of issues understandably cause public unease and raise valid questions about ensuring AI systems are responsible and appropriately managed (Akbarighatar et al., 2023b).

To prevent unintended negative consequences and foster positive outcomes in the deployment of AI systems and services to various stakeholders, both public and private sectors, as well as researchers, have proposed ethical and responsible AI principles (Clarke, 2019; Ess, 2009; Sojer et al., 2014). These principles, such as benevolence, non-malfunction, safety, and well-being, can guide organizations in their decision-making processes when implementing AI-driven technologies to achieve their strategic objectives (Mirbabaie et al., 2022). Incorporating these principles into strategic management and operationalizing them requires considering two perspectives. Firstly, it is imperative for managers and key personnel within organizations engaged in the development and deployment of AI systems to not only acknowledge the importance but also understand the synergistic nature of implementing and operationalizing these principles as a comprehensive system (which) (Akbari Ghatar et al., 2023a). Secondly, there is a need for clear and effective mechanisms (hows) that enable the practical application of these principles within the organization's processes and practices (Akbari Ghatar et al., 2023a; Whittlestone et al., 2019). Hence, in order to achieve the intended operationalization of the principles in the AI efforts, it is crucial to promote and operationalize the RAI principles to ensure alignment with these principles throughout the AI lifecycle and translate these principles into practices (Mittelstadt, 2019).

There appear to be three gaps in the research on operationalizing responsible or ethical AI, despite recent excellent research on operationalizing and translating the concepts into practices. First, the existing literature does not sufficiently explore or provide insights into how experts within organizations assess the relative importance of Responsible AI (RAI) principles (Vakkuri et al., 2019). The gap is related to limited empirical studies that directly investigate

the views of these experts of the relative importance of different principles across organizational contexts. Second, there is a gap in the understanding of how experts perceive the alignment of AI-infused initiatives with RAI principles (Munn, 2022). Previous research has not extensively examined the specific criteria or indicators that experts consider when evaluating this alignment. Additionally, there is limited research that delves into the potential challenges or barriers encountered in achieving alignment. Here, the focus is on assessing alignment from a higher, strategic standpoint, considering broader strategic factors. Finally, a gap in the literature exists regarding how experts perceive the operationalization of RAI principles within organizations (Morley et al., 2020). This gap is related to a scarcity of studies that provide in-depth insights into the practical implementation and integration of RAI principles into the daily operations and decision-making processes of organizations.

Our research is structured to provide evidence-based insights by conducting a survey gathered from AI experts who are actively involved in contributing to, managing, or consulting on AI initiatives. This approach allows us to gain a better understanding of how experts perceive the operationalization of RAI principles. To address this goal, we have framed three key research questions:

*RQ1: How do experts perceive the relative importance of RAI principles in their organizations?*

*RQ2: How do experts perceive the alignment of AI-infused initiatives with RAI principles?*

*RQ3: How do experts perceive the operationalization of RAI principles in their organizations?*

We expect to contribute to research in three areas. First, we extend knowledge of the existing literature by providing insights into how experts within organizations perceive the importance of Responsible AI (RAI) principles. Second, our research seeks to advance the understanding of how experts perceive the alignment of AI-infused initiatives with RAI principles. Third, we intend to contribute to the literature by shedding light on how experts perceive the operationalization of RAI principles within organizations.

The subsequent sections of the paper are structured as follows. In Section 2, we delve into the existing literature concerning responsible AI principles and discuss the journey from principles to practical application. Within this section, we also present a summary of the most important principles. In the upcoming sections, we delve into key aspects of our study. Section 3 details our data collection methods and analysis. In Section 4, we present empirical findings that directly address our research questions. Following these sections, our discussion section offers an extensive analysis of these results. It not only examines expert perspectives on RAI principles but also explores their theoretical and practical implications. We also discuss the limitations of our study and potential areas for future research.



## **2. Related Literature**

### **2.1. Responsible AI principles in practice**

It has been over five years since IS scholars initiated their investment in understanding how AI should be managed (Berente et al., 2021). Also, others highlighted the unintended consequences of the unethical use of AI and proposed some principles for being responsible AI. The AI4People recommendations, rooted in bioethical principles, serve as significant ethical guidelines in Western AI development. These principles, which encompass Autonomy, Beneficence, Non-Maleficence, Justice, and Explicability, have been adapted to address AI's unique challenges in healthcare. Specifically, transparency and explainability have been integrated into these recommendations. Transparency pertains to users' understanding of AI system development and functionality, while explainability focuses on the AI system's capacity to provide clear explanations for its decisions.

In practical literature, numerous inquiries also have taken place. A notable report, published by the Organization for Economic Co-operation and Development (OECD) in early 2019, stands out. This report synthesizes insights from over 70 documents that discuss ethical AI principles across various sectors. The documents originate from a range of sources, spanning industry players like Google, IBM, and Microsoft, governmental entities such as the Montreal Declaration and the Lords Select Committee, and academic institutions including the Future of Life Institute, IEEE, and AI4People. The standard comprises five complementary value-based principles: inclusive growth, fairness, transparency, security and safety, and accountability.

In a study that reviewed 84 ethical AI documents, the prevalent themes were transparency, justice and fairness, non-maleficence, responsibility, and privacy, each appearing in over 50% of cases (Jobin et al., 2019). Moreover, a systematic analysis of the ethical technology literature by (Royakkers et al., 2018) underscored recurring themes encompassing privacy, security, autonomy, justice, human dignity, technology control, and power equilibrium. As posited by these scholars, when considered collectively, these themes collectively 'define' ethically aligned machine learning as technology that is (a) beneficial and respectful towards individuals and the environment (beneficence); (b) resilient and secure (non-maleficence); (c) reflective of human values (autonomy); (d) fair (justice); and (e) transparent, accountable, and comprehensible (explicability).

When examining the European Commission's High-Level Expert Group report's ethical principles, a consistent pattern emerges. The report outlines four ethical principles, deeply rooted in fundamental rights, that must be upheld to ensure the trustworthy development,

deployment, and use of AI systems. The first principle prioritizes respecting human autonomy and freedom (respect for human autonomy). The second emphasizes that systems should neither cause harm nor worsen existing issues for humans (prevention of harm). The third underscores the necessity for fairness throughout AI's lifecycle (fairness). Lastly, explicability proves essential for establishing and maintaining user trust in AI systems. This mandates transparent processes, clear communication of AI system capabilities and intentions, and comprehensible decisions for those directly and indirectly impacted. The absence of such information impedes the ability to challenge decisions effectively (explicability).

ISO 22989:2022 and ISO 24038 provide definitions and detailed explanations of the concept of trustworthiness, encompassing elements such as robustness, reliability, transparency, explainability, interpretability, accountability, safety, privacy, and fairness. All these concepts align with the categories established by OCED and the European Commission (2019). For instance, transparency, interpretability, expandability, and accountability, share a common goal from varying perspectives, reinforcing each other. Collectively, these principles advance AI systems' understandability. Additionally, principles connected to avoiding harm and positive impacts, such as safety, privacy, benevolence, and non-maleficence, uphold AI's beneficence nature. Similarly, fairness and inclusiveness aim to eradicate disparities, ensure equal opportunities, and prevent marginalization. The harmonious combination of Responsible AI principles contributes to a better understanding of RAI and how they synergistically work together (Akbarighatar et al., 2023c). By sharing common objectives, these principles support and reinforce each other, forming a cohesive framework for Responsible AI. This means that the various principles of Responsible AI complement and enhance one another, resulting in an integrated approach to responsible AI development and deployment concisely presented in Table 1, offering a holistic grasp of these pivotal principles.

While recent research has made significant contributions to the field of AI ethics, particularly in the exploration of duty ethics and virtue ethics within sociotechnical systems, there remains a need to further elucidate the interconnectedness of these ethical viewpoints. (Heyder et al., 2023) have provided a theoretical framework in this regard. In our research, our emphasis is on duty ethics, which involves establishing ethical principles to guide human behavior, specifically in our context—experts. While virtue ethics cultivate character duty ethics better suit the governance needs of organizations through organizational principles and policies. Duty and virtue ethics complement each other. Organizational principles and rules aimed at duties/obligations (duty ethics) can help shape an ethical culture and virtuous behavior

over time (virtue ethics). Our research, focusing on duty ethics, aims to contribute to the ongoing discourse on AI ethics in practice.

<b>Principle</b>	<b>Literature descriptions</b>	<b>Refs</b>
Benevolence and Non-maleficence	Indicate that AI technology is designed to promote good and maximize benefits, all the while avoiding harm and minimizing risks.	(European Commission., 2019; Microsoft AI, 2020; Clarke., 2019; Floridi et al., 2018 ).
Reliability and Safety	AI systems should aim to prevent failures and accidents ensuring intended performance.	(ISO:24028, 2020; Microsoft AI, 2020; Clarke., 2019)
Privacy	Freedom from intrusion into an individual's private life or affairs when it happens due to improper or illegal collection and use of their data.	(ISO:24028, 2020; Microsoft AI., 2020).
Security	Security refers to protecting data and controlling access based on authorization levels.	(ISO:24028, 2020; Microsoft AI., 2020).
Accountability	Accountability refers to taking responsibility, providing justifications for actions, responding to inquiries, and being liable.	(ISO:24028., 2020; Microsoft AI., 2020; Clarke., 2019)
Explainability	Explainability refers to providing comprehensive information about AI's inner workings.	(ISO:22989., 2020; Microsoft AI., 2020; Clarke., 2019)
Intelligibility	Intelligibility refers to enabling humans who use or manage AI to understand the reasoning of an AI system.	(ISO:24028., 2020)
Transparency	Transparency entails disclosing AI system details, like performance, limitations, components, measures, design goals, data sources, for a decision, prediction, or recommendation.	(IS:22989., 2020; Microsoft AI., 2020; Clarke., 2019; Floridi et al., 2018)
Inclusiveness	Inclusiveness refers to involving diverse individuals and perspectives, regardless of their unique circumstances.	(OECD., 2018; Microsoft AI., 2020)
Fairness	AI systems must be designed to ensure impartial treatment, and prevention of discriminatory outcomes.	(OECD., 2018; 2020; Microsoft AI., 2020; Clarke., 2019; Floridi et al., 2018)

**Table 1. Responsible AI principles and their descriptions**

### **3. Data and Research Methodology**

#### **3.1. Instrument development**

To ensure the validity and robustness of the developed survey instrument, we followed the guidelines recommended by (MacKenzie et al., 2011). Our process began with the conceptualization of the constructs representing Responsible AI (RAI) principles in our study,

as outlined in Table 1. To evaluate the content validity of these principles, we engaged a panel of six experts with substantial academic and practical experience in responsible AI. Four of these experts had over 15 years of industry experience in data science and AI, while the remaining two were senior academics specializing in Information Systems in organizations. We provided the experts with definitions of each principle and asked them to answer the survey questions. Additionally, we sought their recommendations for improving or refining questions. Their feedback led to minor modifications and clarifications in the definitions, reinforcing the content validity of our instrument.

To assess convergent, discriminant, and nomological validity, we distributed the revised survey instrument to four C-level technology managers. These managers were selected from companies that specialize in the responsible development and deployment of AI and possess extensive experience in implementing RAI principles. Taking their valuable input into account, we carefully revised and refined the definitions of the principles to ensure they were more concise and understandable.

### **3.2. Data Collection**

A 'survey' is a research method where experts in a specific field are queried about their views on relevant organizational factors (Rungtusanatham et al., 2003). Surveys enable a stronger connection between academia and the real world by testing conceptual models with real-world data (Flynn et al., 1990), making it a suitable approach for our current research.

Our survey targeted AI and machine-learning experts involved in AI solution development and integration as a business enabler. The participants consisted of CEOs, managers, AI governance experts, and other relevant positions within these organizations. We identified and contacted potential respondents through professional groups on LinkedIn, such as the “Artificial Intelligence and Business Analytics” group, and the website “Ethical AI Database” to search for responsible AI or ethical AI companies and in general AI companies. This approach ensured a robust and representative sample for our study.

We reached out to selected respondents via email, specifically targeting those in high-level technology management roles who possessed knowledge of RAI operations and practices. Following an initial invitation and three subsequent reminders, each one week apart, we sent a total of 600 email invitations from September to October 2023 to potential participants experienced in AI-infused projects. From these invitations, we received a total of 82 complete and 13 incomplete ones, primarily due to respondents' unfamiliarity with certain initiatives. These responses came from various industries, including financial services, manufacturing, and

high-tech companies. The participants held a range of job titles, including head of AI or data science, chief data governance officers, directors of IT, co-founders, and chief data scientists.

## **4. Empirical results**

### **4.1 Demographic data**

In this study after giving short definitions of RAI principles and understanding them, we ask about the participant's perception of the importance, alignment, and operationalisation mechanisms and we use seven Likert points to measure them. To gain insights into participant demographics, we collect additional information such as their age, gender, professional background, years of experience, and organizational affiliations, allowing us to better understand the diverse perspectives within our participant pool. Across the total sample, the gender balance was 33% women 66% men, and 1% identified as non-binary or with other gender identities. 95 percent (80) of the respondents contributed, managed, or consulted to AI projects. The remained participants were excluded from the further analysis.

The sample (N=82) comprised professionals with significant experience in AI roles. The majority held graduate degrees, with 24% possessing a PhD and 56% a master's-level qualification. Over half (51%) had accrued more than 10 years of overall work experience. Regarding AI specialty, 33% reported 1-3 years spent in AI-related duties.

Geographically, Europe was most represented at 47% of respondents. North America accounted for 22% and the Australia/New Zealand region 21%. Participant organizations ranged in size, with 34% employed by large enterprises (>500 employees), and 48% by small-to-medium businesses ( $1 < x < 100$ ). A diversity of industries was sampled, including 36%, 10%, and 8% from technology, healthcare, and finance respectively. Additional sample characteristics are provided in Table 2. This delineation by demographics, roles, sectors, and geographies offered a breadth of expert insights across the global AI landscape. Overall, the sample comprised knowledgeable professionals well-positioned to offer informed perspectives on organizational responsible AI strategy and implementation efforts.

Category	Percentage of respondents N=80	Category	Percentage of respondents N=80
<b>Working experience</b>		<b>AI-related Working experience</b>	
Fewer than one year	1,2%	Fewer than one year	3,7%
1-3 years	11,0%	1-3 years	32,9%
4-6 years	15,9%	4-6 years	32,9%
7-10 years	19,5%	7-10 years	17,1%
More than 10 years	52,4%	More than 10 years	13,4%
<b>Familiarity with the concept of responsible AI</b>		<b>Education level</b>	
Expert	22,0%	PhD	24,4%
Very familiar and involved in RAI practices actively.	34,0%	Master	58,5%
Very familiar and some involvement in RAI practices	22,0%	Bachelor	17,1%
Familiar but never involved in RAI practices	14,6%		
Heard about it - slightly familiar	7,3%		
Never heard about it	0,0%		

**Table 2. Sample Characteristics**

## 4.2 Expert Perceptions on Responsible AI Implementation

To explore the three research questions, before responding to the questions, participants were given a description of these principles, which can be found in Table 1. To address all research questions a 7-point Likert scale was used to measure experts' perceptions. Regarding the questionnaire's reliability assessment, we utilized SPSS software (version 29.0). The results exhibited strong Cronbach's alpha values of 0.919, 0.928, and 0.896 for importance, alignment, and operationalization, respectively. The total Cronbach's alpha value of 0.962 confirms the reliability of our questionnaire data.

- RAI Principles' Perceived Importance

To address the first research question, we inquired about the extent to which participants perceived their organization's adherence to responsible AI principles. As previously mentioned, the surveys used a 7-point Likert scale ranging from 1 to 7 to rate each principle. A value of 1

represented "Never" in terms of importance/implementation, while 7 represented "Always". Table 3 presents a summary of the observed minimum, maximum, average, and standard deviation scores for various Responsible AI (RAI) principles. These metrics provide valuable insights into the participants' perceptions of each principle's importance within their organizations.

For example, the average scores indicate that participants viewed principles like Reliability and Safety (5.976) and Privacy (6.167) as more important than others like Intelligibility (4.643) and Inclusiveness (4.506). The standard deviation scores offer a detailed view of the variation in experts' responses, showing how closely or widely the principles were rated. For example, the average scores indicate that participants viewed principles like Reliability and Safety (5.976) and Privacy (6.167) as more important than others like Intelligibility (4.643) and Inclusiveness (4.506). The standard deviation scores offer a detailed view of the variation in experts' responses, showing how closely or widely the principles were rated. For instance, while Benevolence and Non-maleficence received a high average importance rating of 5.548, it also had a relatively wide standard deviation of 1.5, suggesting more dispersed views on their importance compared to principles like Reliability and Safety (std dev of 1.202). Overall, Table 2 provides a nuanced understanding of participants' prioritization of each RAI principle.

<b>Responsible AI principles</b>	<b>Observed minimum</b>	<b>Observed maximum</b>	<b>Average</b>	<b>Standard Deviation</b>
Benevolence and Non-maleficence	2	7	5.548	1.5
Reliability and Safety	2	7	5.976	1.202
Privacy	2	7	6.167	1.18
Security	2	7	6.012	1.247
Accountability	2	7	5.155	1.632
Explainability	2	7	4.952	1.693
Intelligibility	2	7	4.643	1.603
Transparency	2	7	4.833	1.642
Inclusiveness	1	7	4.506	1.87
Fairness	1	7	4.843	1.858

**Table 3. Perceived Importance of Observed RAI Principles**

- Experts' Views on AI Initiatives Alignment with RAI Principles

To address the second research question (RQ2), respondents were asked to rate on a 7-point Likert scale the degree of alignment between their organization's AI initiatives and responsible AI principles. As previously noted, the scale ranged from 1 to 7, with 1 representing "To a very little extent" and 7 being "To a great extent". Table 3 summarizes perceptions of alignment for

various principles. The averages provide insight into which principles on average are perceived to be best aligned. For instance, principles like Privacy (5.905) and Reliability and Safety (5.607) had higher average alignment scores than principles such as Intelligibility (4.512) and Inclusiveness (4.229).

The standard deviations in Table 4 also offer perspective into response variability. Fairness and Explainability exhibited wider standard deviations of 1.8 and 1.773 respectively, indicating more dispersed views on the alignment of these principles within organizations. In contrast, principles like Privacy (1.228) and Security (1.28) had tighter standard deviations, suggesting greater agreement among participants regarding their organizational alignment.

In summary, Table 4 analyzes experts' perceptions of how well-aligned their organizations are with responsible AI principles in practice. This sheds light on relative strengths and opportunities in operationalizing ethics.

<b>Responsible AI principles</b>	<b>Observed minimum</b>	<b>Observed maximum</b>	<b>Average</b>	<b>Standard Deviation</b>
Benevolence and Non-maleficence	2	7	5.357	1.588
Reliability and Safety	2	7	5.607	1.336
Privacy	2	7	5.905	1.228
Security	2	7	5.690	1.28
Accountability	2	7	4.786	1.636
Explainability	1	7	4.548	1.773
Intelligibility	2	7	4.512	1.639
Transparency	2	7	4.583	1.6
Inclusiveness	1	7	4.229	1.776
Fairness	1	7	4.614	1.8

**Table 4. Perceived Alignment of AI Initiatives with RAI Principles**

- **RAI Principles' Operationalization in Organizations**

As shown in Table 5, experts were asked to assess the operationalization of responsible AI principles using a 7-point Likert scale. Consistent with previous questions, the scale ranged from 1 to 7, with 1 representing "Never" and 7 being "Always". The data provides insights into both average ratings and standard deviations. Certain principles such as Privacy, Security, Reliability, and Safety received above-average scores of 5.8, suggesting stronger implementation compared to others. This suggests a stronger alignment of these principles with actual practices in comparison to others. In contrast, other principles like Accountability and Inclusiveness averaged below 5, implying greater room for improvement.

Furthermore, the standard deviations within the table provide additional insights. Principles like Fairness and Inclusiveness exhibited standard deviations exceeding 1.8, signifying varying



perspectives on how these concepts are put into practice. In contrast, principles like Privacy displayed tighter variability, with standard deviations near 1.3, indicating a higher level of consensus among participants regarding their operationalization. These findings suggest a nuanced understanding of the implementation landscape of responsible AI principles across organizations.

Responsible AI principles	Observed minimum	Observed maximum	Average	Standard Deviation
Benevolence and Non-maleficence	2	7	5.202	1.589
Reliability and Safety	2	7	5.798	1.315
Privacy	2	7	5.869	1.306
Security	2	7	5.869	1.17
Accountability	1	7	4.905	1.712
Explainability	1	7	4.607	1.672
Intelligibility	1	7	4.536	1.704
Transparency	2	7	4.643	1.573
Inclusiveness	1	7	4.446	1.796
Fairness	1	7	4.627	1.833

**Table 5. Operationalization of RAI Principles in Organizations**

## 5. Discussion

This section presents a discussion based on the findings of this study, offering valuable insights into practitioners' perceptions of responsible AI across importance, alignment, and operationalization dimensions. The diverse, experienced sample provides well-informed perspectives on progress and gaps in the operationalization of responsible AI principles.

- Patterns of Perceived Importance

Our survey responses revealed clear patterns in how responsible AI (RAI) principles are perceived, allowing us to classify them into three distinct categories, as illustrated in Figure 1. Firstly, principles such as privacy, security, reliability, and safety received significantly higher average importance ratings, with scores often approaching or exceeding 6. This alignment with the common emphasis in AI initiatives on addressing paramount concerns related to responsible use in algorithmic decision-making reflects our findings (Ashok et al., 2022). Notably, the recognition of privacy, security, reliability, and safety as foundational elements, as outlined in ISO:24028 (2020), underscores the necessity of prioritizing these aspects in AI development.

Secondly, we observed that principles like benevolence, non-maleficence, and accountability constitute the second group, with average importance ratings around 5.5 to 5.2. While these principles are still considered significant, they fall slightly below the top-tier

principles in terms of perceived importance. Conversely, the third group comprises principles such as intelligibility, transparency, explainability, fairness, and inclusiveness, which received lower average importance ratings, hovering around 4.5. These scores suggest that they may not be as highly prioritized within organizations compared to the principles in the first two groups.

- Alignment and Implementation

This pattern of prioritization based on perceived importance is also reflected in the alignment and implementation perceptions of these principles. Alignment between organizations' AI initiatives and RAI principles was assessed using the same scale. Privacy, reliability, and safety showed higher average alignment scores (approximately 5.9 and 5.6), while principles like fairness and inclusiveness had lower scores (around 4.5), implying room for improvement. The operationalization of RAI principles varied, with privacy, security, reliability, and safety demonstrating stronger implementation, while fairness and inclusiveness showed lower scores, indicating areas for enhancement.

- Highlighting Variations in Importance

The standard deviation scores, as depicted in Figure 2, offer valuable insights into the diversity of opinions among experts, highlighting the extent of variation in their views concerning the significance of specific principles. Notably, despite receiving high average importance ratings, certain principles displayed relatively wide standard deviations, signifying varying perspectives on their significance when developing and applying AI systems.

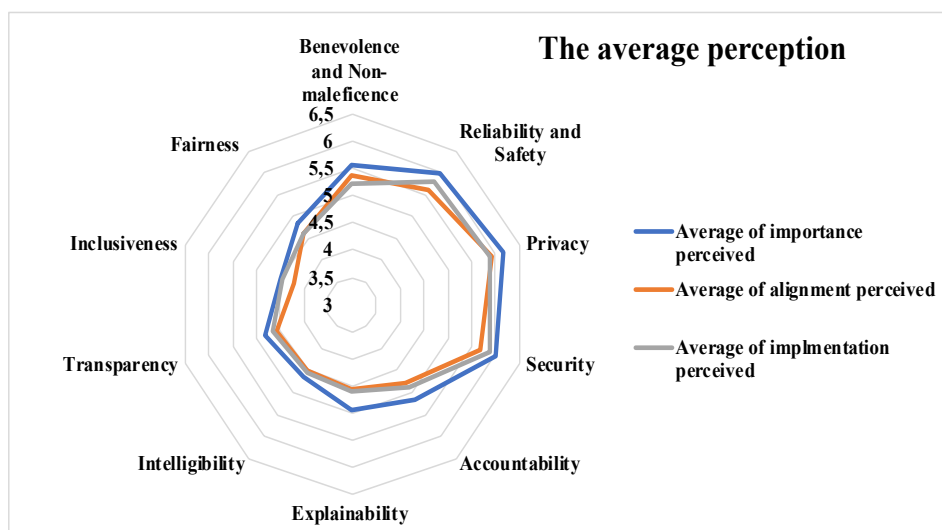


Figure 1. Average of perceived importance

For example, while benevolence and non-maleficence averaged a high importance score of 5.548, they also had a relatively wide standard deviation of 1.5. This variability can be partly explained by the subjective nature of these principles, where the definition of "benevolent" and "non-maleficent" system design may depend on contextual and cultural factors. Compared to more technical principles like reliability and safety, which received a narrower standard deviation of 1.202, experts likely had more diverse interpretations of how benevolence should be defined and prioritized.

A similar pattern emerged for the principles of fairness and inclusiveness, which exhibited even more substantial variability with standard deviations over 1.8. This wide dispersion in views can be attributed to the contextual nature of these principles, where their operationalization and understanding often depend on the specifics of the situation, the stakeholders involved, and the broader societal context. Experts may have evaluated these principles differently given their diverse backgrounds. Effectively incorporating fairness and inclusiveness may require a more nuanced consideration of social perspectives and adaptive, context-specific approaches that can accommodate varied stakeholder needs (Díaz-Rodríguez et al., 2023).

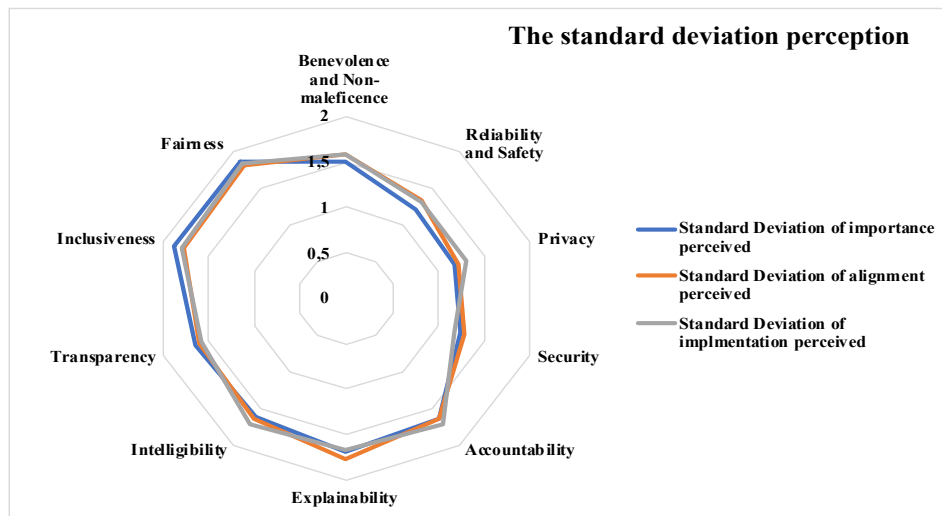


Figure 2. Standard deviation of perceived importance

- Consistency Across Facets

The results of the survey indicate a consistent prioritization of certain facets of responsible AI principles among organizations. Specifically, the principles of reliability and safety, privacy, and security were rated as high in importance. On the other hand, principles such as accountability, explainability, intelligibility, transparency, inclusiveness, and fairness were rated relatively lower.

The reasons behind these differences in prioritization can be multifaceted. One possible explanation is that experts may have varying levels of understanding or expertise in evaluating these principles, leading to different assessments. It is also possible that organizations face challenges in genuinely addressing certain principles, which may require more complex or resource-intensive measures. Another factor could be the clarity and articulation of these principles themselves, as certain principles may be less well-defined or have ambiguous guidelines compared to others.

In a more abstract sense, the higher-rated principles may represent more technical or tangible aspects of responsible AI, while the lower-rated principles may involve more nuanced or contextual considerations. This distinction could contribute to the differing prioritization observed. Further research and analysis are needed to delve deeper into the factors influencing the divergent ratings and to develop a comprehensive understanding of the dynamics between these two groups of principles.

For example, previous studies have consistently supported the value of explainability and transparency in achieving fairness. Vimalkumar et al. (2021) argued that transparency makes AI mechanics visible and known, while explainability describes decisions impacting individuals in human terms, significantly contributing to fairness by enhancing the understanding of model logic and its effects (Robert et al., 2020).

However, the relatively lower prioritization of inclusiveness and fairness in our survey results diverges from views that emphasize the role of principles like transparency, expandability, intelligibility, and accountability. These principles collectively aim to make AI systems understandable, ensuring fairness and inclusiveness in AI development within organizations (Haresamudram et al., 2023). This discrepancy highlights the need for further exploration of the factors influencing the prioritization of these principles in practice. Practitioners need to better understand the importance of fairness, its benefits, and the potential risks for organizations when this principle isn't prioritized.

## **6. Limitations and future research**

While the study offers valuable insights, it is important to acknowledge and address inherent limitations. Firstly, a significant limitation is its reliance on subjective assessments, which may introduce variability due to individual perceptions and biases. To mitigate this challenge, it is advisable to complement subjective assessments with objective metrics and external benchmarks whenever possible, promoting a more balanced evaluation.

Secondly, the study predominantly focuses on assessing perceptions around importance, alignment, and operationalization of RAI principles, potentially neglecting other critical dimensions such as legal compliance or industry-specific considerations. To address this limitation, organizations can consider a broader set of aspects that are relevant to their specific context, thus providing a more comprehensive evaluation.

Lastly, the study currently only includes an analysis of quantitative data. Further work can involve incorporating qualitative methods, such as in-depth interviews or case studies, to gain a deeper understanding of nuanced contexts, especially when evaluating principles like fairness. This balanced approach would allow for a more comprehensive assessment of RAI principles and their practical implementation.

## References

- Akbarighatar, P., Pappas, I., & Vassilakopoulou, P. (2023a). Practices for Responsible AI: Findings from Interviews with Experts. *In Proceedings of the American conference on information systems (AMCIS 2023)*.
- Akbarighatar, P., Pappas, I., & Vassilakopoulou, P. (2023b). A sociotechnical perspective for responsible AI maturity models: Findings from a mixed-method literature review. *International Journal of Information Management Data Insights*, 3(2), p.100193.
- Akbarighatar, P., Pappas, I., & Vassilakopoulou, P. (2023c). Justice as fairness: A hierarchical framework of responsible AI principles. *In Proceedings of the 31<sup>st</sup> European conference on information systems (ECIS 2023)*.
- Amugongo, L. M., Kriebitz, A., Boch, A., & Lütge, C. (2023). Operationalising AI ethics through the agile software development lifecycle: A case study of AI-enabled mobile health applications. *AI and Ethics*, pp.1-18.
- Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for Artificial Intelligence and Digital technologies. *International Journal of Information Management*, 62, 102433.
- Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS quarterly*, 45(3), 1433–1450.
- Clarke, R. (2019). *Principles for responsible AI*. <https://tech.humanrights.gov.au/sites/default/files/inline-files/4A%20-%20Roger%20Clarke.pdf>. Accessed 1 Nov 2020.
- Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., de Prado, M.L., Herrera-Viedma, E. and Herrera, F., 2023. Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. *Information Fusion*, p.101896.
- European Commission. (2019). Ethics guidelines for trustworthy AI. Retrieved from <https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/1.html>.
- Ess, C. (2009). Floridi's Philosophy of Information and Information Ethics: Current Perspectives, Future Directions. *The Information Society*, 25(3), pp.159–168.
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F. and Schafer, B., 2018. AI4People—An ethical

- framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), pp. 689-5.
- Flynn, B. B., Sakakibara, S., Schroeder, R. G., Bates, K. A., & Flynn, E. J. (1990). Empirical research methods in operations management. *Journal of Operations Management*, 9(2), 250–284. [https://doi.org/10.1016/0272-6963\(90\)90098-X](https://doi.org/10.1016/0272-6963(90)90098-X).
- Haresamudram, K., Larsson, S. and Heintz, F., 2023. Three levels of AI transparency. *Computer*, 56(2), pp.93-100.
- Heyder, T., Passlack, N., & Posegga, O. (2023). Ethical management of human-AI interaction: Theory development review. *The Journal of Strategic Information Systems*, 32(3), pp.101772.
- Ho, C. W. L., Soon, D., Caals, K., & Kapur, J. (2019). Governance of automated image analysis and artificial intelligence analytics in healthcare. *Clinical Radiology*, 74(5), 329–337.
- ISO:22989. (2022). ISO/IEC 22989:2022 Information technology — Artificial intelligence — Artificial intelligence concepts and terminology. Retrieved from \_.
- ISO:24028. (2020). ISO/IEC TR 24028:2020 Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence . Retrieved from <https://www.iso.org/standard/77608.html>)
- MacKenzie, Podsakoff, & Podsakoff. (2011). Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques. *MIS Quarterly*, pp.35(2), 293.
- Martilla, J. A., & James, J. C. (1977). Importance-Performance Analysis. *Journal of Marketing*, 41(1), pp.77–79.
- Matzler, K., Bailom, F., Hinterhuber, H. H., Renzl, B., & Pichler, J. (2004). The asymmetric relationship between attribute-level performance and overall customer satisfaction: A reconsideration of the importance–performance analysis. *Industrial Marketing Management*, 33(4), pp.271–277.
- Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and ‘the dark side’ of AI. *European Journal of Information Systems*, 31(3), pp.257–268.
- Mirbabaie, M., Brünker, F., Möllmann Frick, N. R. J., & Stieglitz, S. (2022). The rise of artificial intelligence – understanding the AI identity threat at the workplace. *Electronic Markets*, 32(1), pp.73–99.
- Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), pp.501–507.
- Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Science and Engineering Ethics*, 26(4), pp.2141–2168.
- Munn, L. (2022). The uselessness of AI ethics. *AI and Ethics*, 1-9.
- Oh, H. (2001). Revisiting importance–performance analysis. *Tourism Management*, 22(6), pp.617–627.
- OECD. (2019). The OECD AI Principles. Retrieved from <https://oecd.ai/en/ai-principles>.
- Pappas, I. O., Mikalef, P., Dwivedi, Y. K., Jaccheri, L., & Krogstie, J. (2023a). Responsible digital transformation for a sustainable society. *Information Systems Frontiers*, 25, 945–953.
- Robert, L. P., Pierce, C., Marquis, L., Kim, S., & Alahmad, R. (2020). Designing fair AI for managing employees in organizations: A review, critique, and design agenda. *Human-Computer Interaction*, 35(5-6), 545-575.
- Rungtusanatham, M. J., Choi, T. Y., Hollingworth, D. G., Wu, Z., & Forza, C. (2003). Survey research in operations management: Historical analyses. *Journal of Operations Management*, 21(4), pp.475–488.

- Sojer, M., Alexy, O., Kleinknecht, S., & Henkel, J. (2014). Understanding the Drivers of Unethical Programming Behavior: The Inappropriate Reuse of Internet-Accessible Code. *Journal of Management Information Systems*, 31(3), 287–325.
- Stilgoe, J. (2018). Machine learning, social learning and the governance of self-driving cars. *Social Studies of Science*, 48(1), pp.25–56.
- Vakkuri, V., Kemell, K.-K., Kultanen, J., Siponen, M., & Abrahamsson, P. (2019). *Ethically Aligned Design of Autonomous Systems: Industry viewpoint and an empirical study* (arXiv:1906.07946). arXiv.
- Van Esch, P., Black, J. S., & Ferolie, J. (2019). Marketing AI recruitment: The next phase in job application and selection. *Computers in Human Behavior*, 90, 215–222.
- Vassilakopoulou, P., Parmiggiani, E., Shollo, A., & Grisot, M. (2022). Responsible AI: Concepts, critical perspectives and an Information Systems research agenda. *Scandinavian Journal of Information Systems*, 34(2).
- Wall, L. D. (2018). Some financial regulatory implications of artificial intelligence. *Journal of Economics and Business*, 100, 55–63.
- Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 195–200.
- Zhang, H. Q., & Chow, I. (2004). Application of importance-performance model in tour guides' performance: Evidence from mainland Chinese outbound visitors in Hong Kong. *Tourism Management*, 25(1), 81–91.

# Design and Development of a Digital Transformation Canvas for SMEs in Developing Countries: A Case Study of Oman

## Abstract

*Small and Medium-sized Enterprises (SMEs) across the globe are grappling with formidable challenges brought about by the relentless pace of technological change. This is particularly pronounced in developing countries, where these businesses face a unique set of obstacles in embracing digital transformation. Challenges include limited financial resources, a lack of awareness regarding the implications of digital transformation, a deficit in technological expertise, and the constantly evolving nature of digital technologies. Using Oman as a case study, this research presents an empirically grounded digital transformation framework tailored to the specific needs and constraints of SMEs in developing nations. The study introduces innovative concepts such as the "digital transformation canvas" and "digital interaction," forming the fundamental building blocks of digital transformation. The digital transformation canvas, developed through a design science research approach, offers a practical and adaptable framework for SMEs in these regions. It encompasses nine essential components: ecosystem, strategy, people, technology, innovation, revenue, customer (external interactions), internal operations (internal interactions), and continuous development. This framework provides a valuable roadmap for SMEs in developing countries, seeking to thrive in the digital age.*

**Keywords:** Digital Transformation Canvas, Digital Innovation, Design Science, Organisational Transformation.



## 1.0 Introduction

Gartner defines, Small and Medium Enterprises (SMEs) classification is based on employee count and turnover, typically up to 100 employees and \$5 million in low and middle-income economies (Gartner 2023). European Commission (2023) defines SMEs differently, with small enterprises having fewer than 50 employees and a turnover below €10 million, and medium-sized enterprises having fewer than 250 employees and a turnover below €50 million (*need references here*). SMEs are crucial to economies, constituting 90% of companies, employing over half of the global workforce, and contributing to over 55% of GDP in OECD countries (Alqassabi, 2020). Governments globally support SMEs through various initiatives to foster sustainability. In Oman, SMEs are a key focus, Oman's strategy includes entities like Sharakah, SME Fund, and Riyada. However, entrepreneurship is in its early stages, facing challenges in human capital, government support, financial capital, technology, and regulations. Global challenges for SMEs include limited capital, and technological and human resource gaps (Veronica *et al.*, 2020). Oman faces similar challenges and focuses on capability building, understanding competitors, technology, and government support. Innovation is vital for SME survival. Two perspectives emerge for SME development: the entrepreneurial ecosystem view, focusing on external factors and government support, and the internal view, considering SME components, processes, behaviours, and dynamics. Technology plays a crucial role, especially in digitalizing accounting records, but barriers like technical capabilities and costs persist, hindering ICT adoption in SMEs. The success of the SME sector is vital for all economies, especially developing ones, including Oman. The government of Oman promotes SME growth, enhancing their contribution to the economy, and ensuring their sustainability but these efforts are faced with many challenges. SMEs themselves struggle to grasp the necessary structure, business model, and interrelationships among key components. This research aims to address these challenges, particularly in the context of digital transformation

While the government has supported SMEs at the incubation stage through financial aid, consulting, and training, technology management within SMEs remains fragmented. Technology advancements, particularly in digital services, have primarily occurred within the government sector, emphasizing infrastructure development and e-

government initiatives. Existing research on Oman's SMEs has primarily examined challenges (Muthuraman *et al.*, 2020) and provided strategic recommendations to the government (Al-Abri, Rahim and Hussain, 2018). However, the SME ecosystem and external environmental factors are still in the early stages of development, and support is primarily directed toward business aspects. Consequently, there is a pressing need for more comprehensive studies to help SMEs navigate challenges throughout their lifecycle, adapt to changing external and internal factors, and thrive within the evolving ecosystem. The challenges faced by SMEs in developing countries are exacerbated by the rapid and unstoppable wave of technological transformation affecting businesses across sectors, sizes, and industries. This transformation touches all aspects of the business model, including Product, Customer, and infrastructure (Osterwalder and Pigneur, 2013) However, SMEs face specific hurdles in adopting digital transformation due to factors like insufficient growth capital, limited awareness of its impact, a lack of inherent technology expertise, and the dynamic nature of digital technologies (CRN Team, 2020).

Digital transformation is no longer a trend but a necessity (Butt, 2020). It has disrupted traditional business models, as predicted by Nicholas Negroponte in his 1995 book "Being Digital," where he foresaw the transformation of packaging logistics into universal data communication (Bongiorno, Rizzo and Vaia, 2018). Companies are compelled to navigate this transformation to stay competitive, but this race demands skills and tactics yet to be fully defined, influenced by various factors.

This research seeks to adopt the principles and practices that shape digital transformation and propose an appropriate framework for SMEs, particularly in developing nations, to accomplish their objectives

The central problem identified is the need for digital transformation in SMEs in developing countries specifically in Oman, but there is no viable framework for its direct implementation in terms of alignment, comprehensive elements, processes, and behaviours. The research argues that developing such a framework is crucial to enable successful digital transformation within the SME sector in developing countries, which will be discussed in the following sections.

## 2.0 Digital Transformation (DT)

This session discussed digital transformation frameworks, starts with the introduced definitions, how the previous frameworks were developed, which data sources were used to develop these frameworks, and how they were designed.

### 2.1 Digital Transformation Definitions

Various studies have attempted to establish a unified definition of digital transformation, reflecting the growing interest in this topic, especially in the wake of the COVID-19 pandemic in 2020. (Gong and Ribiere, 2020) conducted an extensive review of 134 digital transformation definitions, leading to a refined and empirically verified definition: "A fundamental change process, enabled by the innovative use of digital technologies accompanied by the strategic leverage of key resources and capabilities, aiming to radically improve an entity and redefine its value proposition for its stakeholders." This definition underscores the transformative nature of digital technologies and their strategic impact. McKinsey's definition, as discussed by (Libarikian, 2019) emphasizes two levels: core transformation, which enhances existing processes using technology, and new business creation, where entirely novel offerings are developed. (Ulas, 2019) defines digital transformation as the reconfiguration of technology, business models, and processes to create new value in a dynamic digital economy, highlighting the multifaceted components involved.

Westerman and McAfee (2012) describe digital transformation as an ongoing process of advancing digital maturity through technology adoption and cultural change. Successful transformation enables organizations to provide better services, gain competitive advantages, and respond effectively in complex environments, leading to improved profitability.

Data and value creation are recurring themes in various definitions. Young and Rogers define digital transformation as "a technology-driven process of change derived from ubiquitous data, connectivity, and decision-making" (van Tonder *et al.*, 2020) characterizes it as a process that leverages information, computing, communication, and connectivity technologies to trigger significant changes in an entity's properties.

These definitions collectively emphasize the profound changes digital transformation brings to organizations, driven by technology innovation, strategic alignment, and a focus on value creation in an ever-evolving digital landscape.

## **2.2 Data Sources in Digital Transformation Frameworks**

The existing literature on digital transformation frameworks relies on various methods to gather primary data for analysis. The majority of studies use surveys and interviews to collect information from a diverse group of companies across different sectors and sizes. During these interviews, a set of characteristics is identified and conceptualized, as observed in the work of (Soule *et al.*, 2015). Some studies take a theoretical approach by analyzing existing designs and elements of business models. For instance, (El Sawy and Pereira, 2013) examined 25 different business models to develop their framework known as VISOR. Others, such as (Bharadwaj *et al.*, 2013), (Gupta Sunil, 2018), (Kavadias, Stelios; Ladas, Kostas; Loch, 2016), and (Adams, 2005), rely on the study of existing cases to formulate their frameworks.

(van Tonder *et al.*, 2020) used (Fenton, Fletcher and Griffiths, 2020) analytical data from transforming companies to create Cognizant's Digital Transformation Framework. Additionally, conducted a more in-depth interaction with organizations using an action case approach over the long term to develop their framework. These diverse methods demonstrate the various approaches taken in the literature to construct digital transformation frameworks.

## **2.3 DT Frameworks building blocks**

The development of digital transformation frameworks has yielded various models of the business landscape, operating at different organizational levels. These frameworks employ different terminology and nomenclature to describe similar elements. To simplify these concepts, a unified classification is proposed here. There is a consensus on the need to embrace change, with change being a constant factor in organizations (Perkin and Abraham, 2017) However, digital transformation has brought about a transformed scale, scope, and pace of change across sectors and industries. Existing frameworks refer to the driving forces behind this change, described by (Fenton, Fletcher and Griffiths, 2020) as the VUCA world (volatile, uncertain, complex, and ambiguous). These drivers were also characterized by (El Sawy and Pereira, 2013) as

the primacy of customer experience, distributed co-creation of value, and continuous sense-and-response experimentation.

Most existing frameworks break down digital transformation into various components. For instance, (El Sawy and Pereira, 2013) introduced five components in their VISRO framework (Value Proposition, Interface, Service Platform, Operational Model, Revenue Model). (Soule *et al.*, 2015) referred to these components as characteristics (Customer experience, Operational efficiency, and Workforce enablement). (Fenton, Fletcher and Griffiths, 2020) described them as subsystems (People, Data, Hardware, Software, Process, and Communication). Similarly, (Kavadias, Stelios; Ladas, Kostas; Loch, 2016) referred to them as elements (Personalized Product, Closed-Loop Process, Asset Sharing, Usage-Based Pricing, Collaborative Ecosystem, Agile and Adaptive Organization). (van Tonder *et al.*, 2020) identified components required for digital transformation as Customer, Product, Organization, and Processes and Systems. Another common area in these frameworks is the strategic approach. This approach is described differently across different frameworks, such as stages and subsystems (e.g., HING by (Fenton, Fletcher and Griffiths, 2020), or as themes (e.g., Vision, New Perspectives, Organization as an Energy Field, Leadership, Human Empowerment, Performance Excellence in (Adams, 2005). (Gupta Sunil, 2018) outlines it as Reimagining your business, Reevaluating your value chain, Reconnecting with customers, and Rebuilding your organization. (Nwaiwu, 2018) simplifies it into Hyperawareness, Informed Decision Making, and Fast Execution. These diverse frameworks collectively contribute to a comprehensive understanding of digital transformation.

## **2.4 Innovative Model**

In the context of our previous categorization of transformation into core changes and innovative changes, innovation stands out as a crucial aspect with its own components. (Kavadias, Stelios; Ladas, Kostas; Loch, 2016) highlight six areas related to innovative and transformative business models that facilitate the necessary level of innovation: personalized product, closed-loop process, asset sharing, usage-based pricing, collaborative ecosystem, and agile and adaptive organization.

Additionally, it's important to distinguish between opportunity discovery and opportunity creation. Opportunity discovery entails identifying existing opportunities,

while opportunity creation suggests that opportunities are not waiting to be recognized; rather, entrepreneurs actively create them through their actions, often in the pursuit of exploring new products or services (Ojala, 2016). This differentiation sheds light on the proactive role of entrepreneurs in shaping and creating opportunities rather than merely recognizing them.

## **2.5 Design Methods of DT Framework**

Developing frameworks and paradigms for digital transformation requires a deep understanding of the underlying components and their inherent characteristics. Adams, (2005) emphasizes the importance of comprehending these components and their hidden facets to effectively adapt them to specific needs. To initiate digital transformation, organizations must analyze existing business models and focus on the organization's mindset, which reflects its belief in digital solutions and is embodied in its capabilities and qualities (Soule *et al.*, 2015). The "Digital Dexterity" framework explores the digital aspects of various organizations to identify key elements.

Design theory, as employed by (El Sawy and Pereira, 2013), is utilized to theoretically analyze existing designs and elements of business models. This analysis compares components across different models to assess their treatment of the digital space and ecosystems. Long-term analysis of a single organization, as seen in the development of the HINGE framework (Fenton, Fletcher and Griffiths, 2020), or short-term analysis across multiple organizations, as used by (Nwaiwu, 2018), reveals the success or failure aspects of organizational models.

Actual case analysis, spanning various industries, broadens the scope of analysis and can lead to the creation of a general concept or framework, as demonstrated by Bharadwaj, (Bharadwaj *et al.*, 2013) and (Gupta Sunil, 2018) in developing a digital leadership framework. Alternatively, focusing on a specific disruptive model helps identify the foundational concepts, as seen in the work of (Kavadias, Stelios; Ladas, Kostas; Loch, 2016).

## **3.0 Research Methodology**

The chosen methodology for this research is the design science methodology, which is well-suited to solving practical problems and creating usable artifacts. Design science focuses on developing and evaluating artifacts to address real-world issues, aligning with the research's goal of creating a practical digital transformation framework (Hevner *et al.*, 2004). Design theory plays a crucial role in this methodology, emphasizing "how to" solutions. The research aims to provide a utility in the form of a digital transformation framework for SMEs, directly aligning with design science's objectives of developing artifacts that offer utility (Hevner *et al.*, 2004).

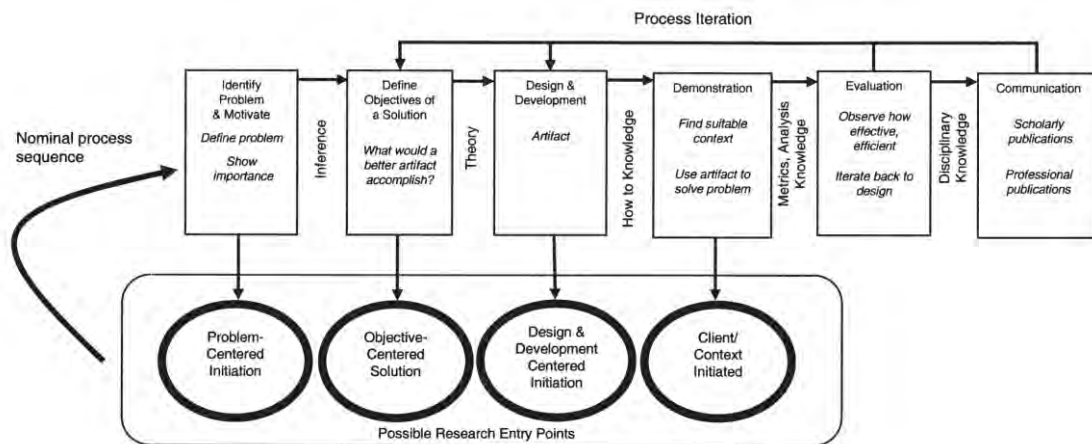


Figure 1 DSRM process model (Peppers et al., 2007)

Qualitative methods will be employed to gain a deep understanding of the problem and its context. The initial artifact construct will be formed through desk-based research, leveraging existing knowledge, frameworks, and models within the domain (Hevner *et al.*, 2004). The research process follows a structured sequence of activities, as proposed by (Peppers *et al.*, 2007). These activities include problem identification and motivation, defining solution objectives, design and development of the artifact, demonstration in a real-world context, evaluation of the artifact's performance, and communication of the findings. Importantly, this process allows for iteration and flexibility in the research approach (Peppers *et al.*, 2007).

The research will follow a problem-centered approach, as the research idea originates from observing the need to address digital transformation challenges in SMEs. This approach aligns with the methodology's problem-solving orientation (Peppers *et al.*, 2007). Overall, the design science methodology provides a structured and iterative

framework for developing a practical digital transformation framework that addresses the specific needs of SMEs. It emphasizes the creation of artifacts that offer real-world utility and aligns with the research's objectives and goals (Hevner et al., 2004; Peffers et al., 2007).

#### **4.0 Data Collection & Analysis**

The research proceeds with the data collection phase, which aims to gather primary research data from selected SMEs to explore their current digital transformation status, perceptions, challenges, and feedback on the developed theoretical framework. Two types of focus groups will be conducted:

1. **Exploratory Focus Groups:** These groups will gather data about the current status of the participating companies, their existing technologies, and their approach to digital transformation. The objective is to understand the organizational landscape in terms of technology utilization and gather insights into their current digital transformation journey.
2. **Confirmatory Focus Groups:** These groups will validate the importance of each proposed element within the digital transformation framework. Depending on the number of participants, there may be several sets of focus groups to gather comprehensive feedback.

The focus groups serves multiple purposes:

- Evaluate the practical applicability of the developed digital transformation framework (The DT Canvas).
- Collect real-world feedback on the framework's usability and effectiveness.
- Understand the dynamics and relationships between the framework components.
- Refine the theoretical framework based on real-world data and participant input.

During the focus groups, participants will be guided through discussions, and specific questions will be addressed to each participant to gather their insights. The data collected from these focus groups will be compared to the theoretical framework, allowing for adjustments and refinements as necessary. This iterative process aims to ensure that the developed framework aligns with the practical needs and challenges faced by SMEs in their digital transformation journeys.



#### 4.1 Exploratory data analysis

The exploratory data gathered from focus groups with SMEs in Oman provides several key insights:

- **Technology Utilization:** Most of the SMEs demonstrated good utilization of various technologies, including infrastructure, hardware, software (at a basic level), social media, messaging applications, and video collaboration tools. However, the average digital capability of these SMEs is rated at "5.6" on a scale of 10. This suggests that while they have a positive view of technology, their current utilization may not be sufficient to meet their business requirements and data operations.
- **Importance of Digital Transformation:** All participating companies acknowledged the importance of digital transformation and expressed a strong desire to digitally transform their businesses. However, they also expressed significant concerns related to technology investment, access to software services, training, technology knowledge, and other related factors.
- **Lack of Comprehensive Framework:** A major concern among SMEs is the absence of a practical and comprehensive framework to guide them through the process of digital transformation. While there is some support available from the government and the private sector for specific cases, SMEs believe that having a suitable framework is critical to the success of their digital transformation efforts.
- **Framework Need:** The SMEs emphasized the importance of having a digital transformation guiding framework. Some even considered it crucial for the success of their transformation initiatives. They expressed their willingness to participate in the transformation process if such a framework were made available.

These findings align with the research problem statement, highlighting the need for a comprehensive digital transformation framework for SMEs in Oman. The challenges identified in this sector directly correspond to the need for a solution in the form of a suitable framework. This framework should address various aspects of digital transformation, including the ecosystem, strategy, people, customer engagement, internal operations, technology utilization, innovation, revenue generation, and business continuity. Ultimately, the application of such a framework is expected to

benefit SMEs in Oman and help them navigate their digital transformation journey effectively. This insight addresses the research questions related to the need for a framework and the challenges faced by SMEs in Oman in their digital transformation efforts.

#### 4.2 Confirmatory data analysis

The confirmatory data collected from the primary research provides valuable insights into the applicability and relevance of the designed digital transformation framework.

Here are some key conclusions based on the data:

- **Component Importance:** The data shows that each component of the framework is considered highly important by the SMEs. This underscores the relevance and validity of the framework's design, as it addresses elements that are crucial for successful digital transformation.
- **Component Utilization:** While the importance of the framework components is recognized, the current utilization of these components within SMEs is rated as medium to low. This suggests that there is room for improvement in how these components are implemented and integrated into their digital transformation efforts.
- **Utilization Challenges:** The data also reveals that SMEs face challenges in fully utilizing these components. These challenges include the strength of the ecosystem and the support it provides, the guidance on implementing the digital strategy, the availability of training and digital skills, the ability to analyze technology investments, and understanding innovation in a practical manner.
- **Framework Relevance:** The data confirms the relevance and importance of the designed framework to real-world companies. The high importance attributed to each element demonstrates that the framework addresses critical aspects of digital transformation that are valued by SMEs.
- **Mitigating Challenges:** Some challenges identified may not be directly mapped to the framework, such as financial support or ecosystem maturity. However, the holistic nature of the framework assumes that it can partially or completely mitigate these challenges. For example, having a revenue analysis component in the framework can support SMEs in justifying their need for financial support. Additionally, considering the ecosystem as part of the transformational

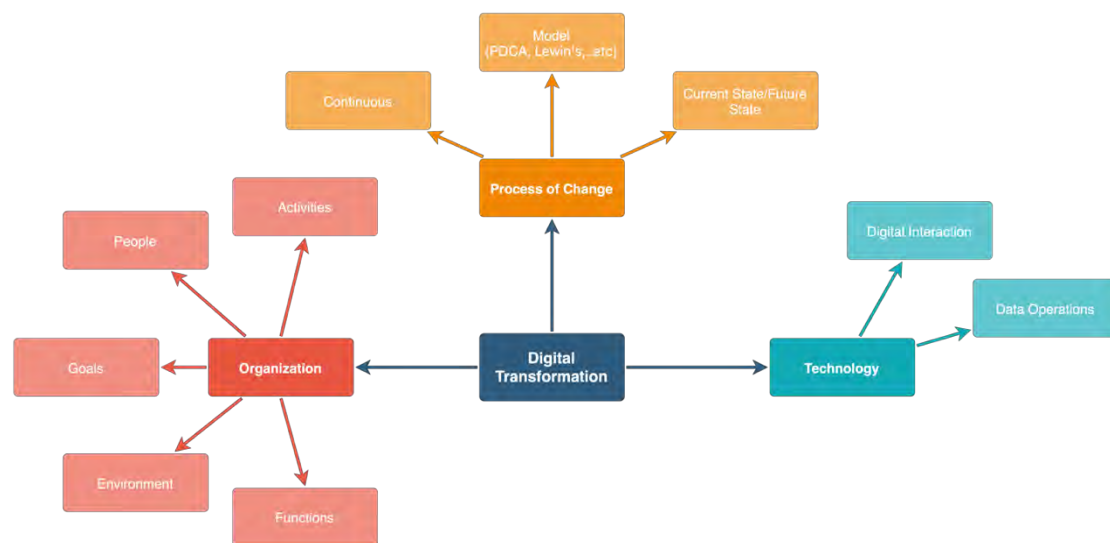
cycle enables companies to adapt to changes and leverage new technologies, which is essential for overcoming challenges related to ecosystem strength.

The confirmatory data supports the importance of the framework's components and their relevance to SMEs in Oman. While there are utilization challenges, the framework is seen as a valuable tool for SMEs to navigate their digital transformation journey and address key aspects of transformation, including strategy, ecosystem strength, continuity, skills development, customer engagement, revenue generation, innovation, and operational processes. The framework's holistic approach is expected to help SMEs overcome these challenges and drive successful digital transformations.

## 5.0 Design & Development

### 5.1 The digital interaction

By running a comparative analysis (Ragin, 2014) using a variation-finding strategy and on the existing definitions, three common terms are found to be used across all definitions, these terms can be unified into the following terms: Transformational Process, Organization, and Technology. Starting with Organization can be defined according to (Daft, 2010). as a social entity that is goal-directed, designed as deliberately structured and coordinated activity systems, and linked to the external environment. Organizations consist of people and their relations in order to achieve the functions, work activities, and goals while interacting with the external environment.

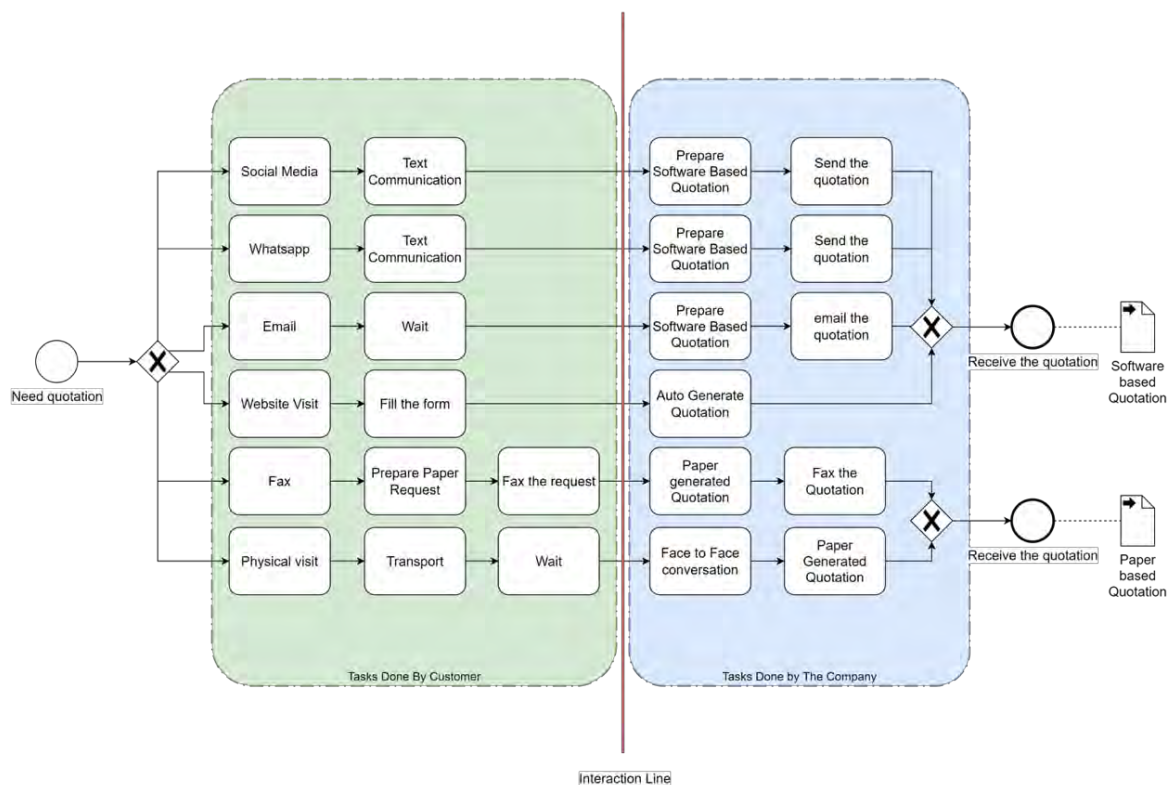


**Figure 2 Organizational, Process, and Technology Components**

Organizational change refers to the process of altering the way an organization operates. This can include changes to the company's structure, processes, technology, culture, or strategy. The goal of organizational change is often to improve efficiency, effectiveness, or competitiveness. Organizational change can be initiated by internal or external factors and can be planned or unplanned. It can be implemented through a variety of methods, such as restructuring, reengineering, or downsizing (Kotter, 2012), in short, it means creating a new way of doing business within the organizations. Change management is also defined by Moran and Brightman (2001, cited in Hussain et al. 2018) as 'the process of continually renewing an organization's direction, structure, and capabilities to serve the ever-changing needs of external and internal customers.

Technology is the application of knowledge for practical purposes ('Oxford English dictionary', 2000), Digital technology as a general definition is the use of digital mediums or devices in storing and exchanging information instead of physical, paper-based, or face-to-face communication. This means using digital devices such as computers, networks, smartphones, the software related to these devices, and the logical and processing power of the cohesive digital technology in data storing, searching, processing, analysis, visualization, and predictions. The application of digital technology is open to what can be digitalized within the technology roadmap since digital technology is the fastest-growing discipline (Kheyr *et al.*, 2019).

Based on the previous analysis, digital transformation is a process of organizational change using technology that creates digital interactions, this can be further simplified by rephrasing it into “a process of changing the current organizational interactions from its current state into a future digital state”, and by looking again into the previous map we can clearly define several interactions on the organizational level such as interactions within the organization and interactions between the organization and its environment. These interactions can be further detailed, and we can take one example of the organization interacting with its environment, a customer making a quotation request can be described in the following diagram:



**Figure 3. Interaction Analysis**

By breaking down the digital transformation into its components, understanding the nature of this phenomena as a continuous change process to the organizational components, and defining what are the important components that need to be visited.

By look into a deeper level of the transformational model using the output of the analysis that has reached the smallest building block of a digitalized organization, this block is the digital interaction. The organization, regardless of its components and structure, consists of interactions between its internal actors, external actors, internal to external, and external to internal, these interactions are part of processes, and the process itself might include digital and non-digital interactions. Digital interaction can be as simple as sending a request over the website, and as complicated as a fully automated machine learning and artificial intelligence sales interaction (Bharadwaj and Shipley, 2020)

Digital interaction is a fully digitalized exchange of information between two actors, and systems environments, that is required to achieve an important step of a process and participate in doing business in a digitalized manner. In a more simplified definition.

**5.2 Components and principles**

Several techniques are applied such as Categorizing the selected data, Identifying and naming concepts, Deconstructing and categorizing the concepts, integrating concepts, Synthesis, and resynthesis (Walls, Widmeyer and El Sawy, 1992), and running a comparative analysis (Johnston, 2014) on the existing frameworks, and on the components extracted from DT definitions we can reach a similar result.

Visor	Adam	M-PWR	Organizational DT resources	Harting	Digital transformation playground (DTP)	Wade	HINGE	Doukidis, spinellis and ebert, 2020
Value proposition	Vision	Digital operations	Strategy	IT landscape	Strategy orientation,	The business model	People	Customer experience

Interface	New perspectives	Data driven decisions	People	Internal business	Customer centricity	The internal structure	Data, hardware, software	Business processes
Service platform	Organization as an energy field	Collaborative learning	Organization	Enterprise architecture	ICT and process	The people	Process and communication	Business model
Operational model	Leadership	Technology experience	Customer	External driver of external business factors	Talent,	Processes		Organizational transformation
Revenue model	Human empowerment	Digital skills	Ecosystem		Capability and capacity strengthening,	ICT capability;		
	Performance excellence	High engagement	Technology		Innovation and culture	The offerings		
		Real-time customer data	Innovation		Organizational commitment	The engagement models		
		Integrated operation data						
		Collaborative tools						

**Table 1 Framework Components Based on Theoretical Analysis**

Strategic Response	Digital Culture	Revenue/Business model	Digital Technology	Continuous Processes
Employees Values	Customer Values	Profit	Asset Return	Innovative Business Building
Changing Economy	Digital	Complex Environment	Measuring Improvements	Drivers

**Table 2 Framework Components Based on Definition Analysis**

By analysing the grouped definitions, we can conclude the following elements

- Strategy group: This group includes elements representing the high-level decision of the organization such as Vision, Strategy, and Strategy orientation.
- The Customer group: This group includes several components such as new perspectives, customer, customer-centric...etc, from a digital transformation point, what matters is the interaction with the customer, and this involves products and services, the customer aspect is found to be a fundamental element of digital transformation across all existing frameworks, it is defined by the value proposition and the value created from the customer through services or products, it shows why a particular customer would value products or services and be willing to pay a premium price for them (el Sawy and Pereira, 2013), Customer also part of the changing organization forces and creating a new way of achieving the work according to (Adams, 2005),the emphasis on customer experience and expectation is also a fundamental part of the M-PWR model and the digital dexterity framework (Soule *et al.*, 2015). The customer element is

fundamental to all other studied frameworks and it is summarized in this research from a digital perspective as the digital interactions with the customer.

- Operations & Processes: This group includes elements such as Internal processes, Digital Operations, Internal business, and Internal structure.
- Technology: This group includes elements such as Interface, Platform, Data, Network, ICT, and Technology.
- Innovation Model: Includes innovation model, and culture.
- People: This theme includes elements such as Human Empowerment, Digital Skills...etc.
- Revenue: This theme includes the revenue model, organization commitment, and business model.
- Ecosystem: This theme includes all external factors such as echo systems, drivers, digital economy, and complex environment.
- Continuous Process: includes elements such as continuous process, and measuring improvements.

The anticipated framework will include elements of different nature and functionality, the analysis was able to define the building block of digital transformation which is digital interaction. In order to put them all together in one framework, we will study the following hypothetical situation. The concluded principles are:

Principle 1	Digital interaction is the exchange of information between actors that achieves data operations (Storing, processing, analysis.... etc.).
Principle 2	The more digital interactions the organization has, the more digitalized the organization will be, this can be used as a measuring tool for organizational digitalization maturity, it can also be used as a guide and reference toward achieving digital transformation within the organization.
Principle 3	There should be a digitalization maturity scale, the proposed scale consists of 5 stages of digitalization.
Principle 4	Digital transformation components can be extracted and analyzed based on existing knowledge, and they can be grouped based on their functionality.
Principle 5	Grouping elements of digital transformation according to their natural dynamics is very important to be reflected within the final framework.

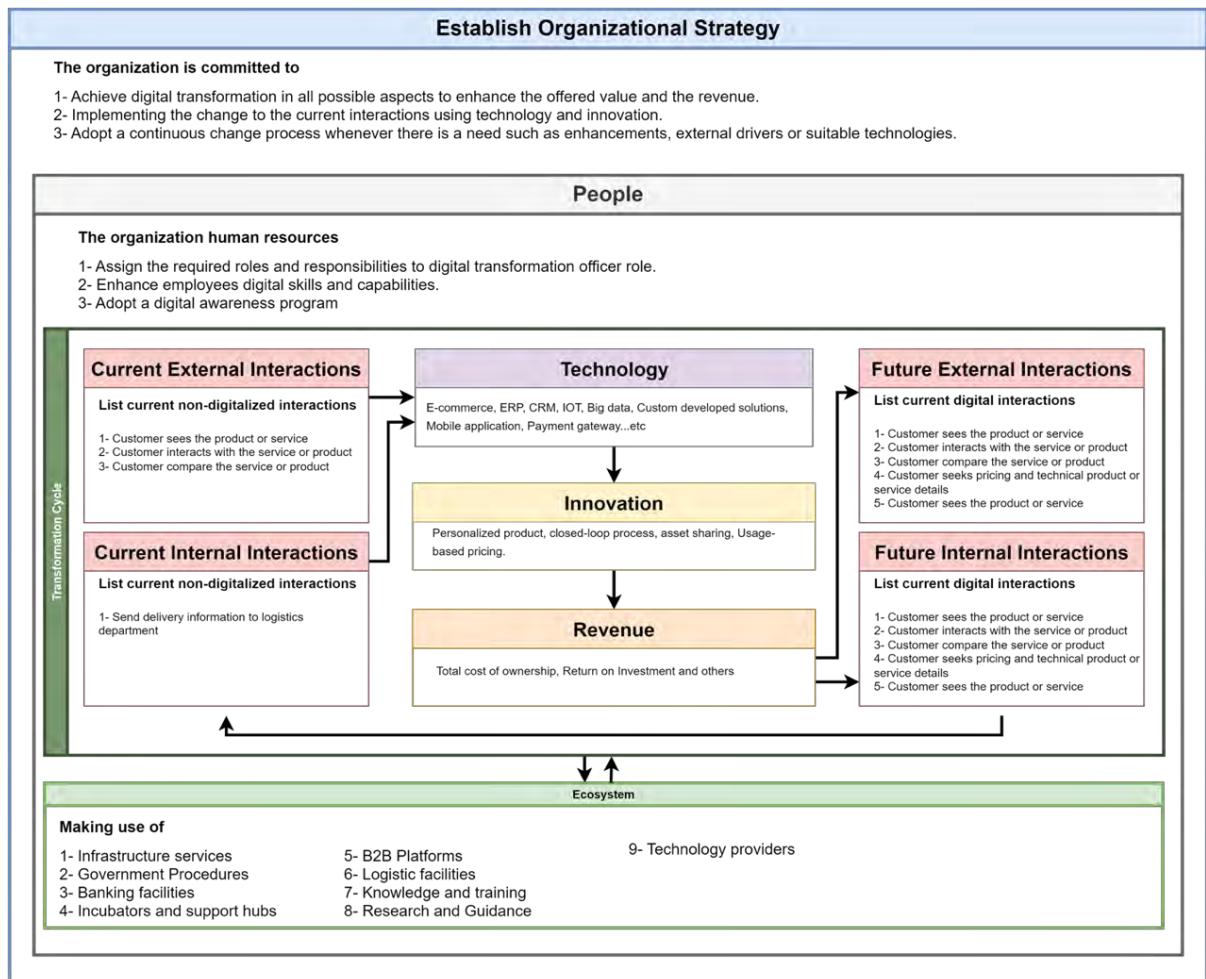
**Table 3 Principles Developed as a base for Framework Design**

### 5.3 The Canvas

By utilizing the previous design principles and by reflecting the nature of the components, the following design representation will reflect the framework



components, the contained components, the cycle, sequence, and interactions between components.



**Figure 3 Digital Transformation Canvas**

We can see that the strategy is the surrounding governance and the starting point, another surrounding component is the people skills and awareness, then we see the internal continuous cycle which starts with the current interaction and pass through three steps of technology, innovation, and revenue analysis to produce the digital interaction while consulting the ecosystem for any driver or available facilities.

## 6.0 Future Steps

To conduct a confirmatory case study and measure the extent to which the developed digital transformation framework can achieve digital transformation goals in SMEs, the research should proceed with the following future steps.

The Case Selection of a representative sample of SMEs in Oman who are willing to participate in the confirmatory case study. Ensure that the selected cases vary in terms of industry, size, and current digital transformation status to capture a diverse range of experiences (Eisenhardt, 1989). Before implementing the digital transformation framework, conduct a comprehensive assessment of each participating SME's current digital maturity, challenges, and objectives. This assessment should serve as a baseline measurement against which the impact of the framework can be evaluated (Kaplan & Norton, 1992).

The Framework Implementation starts with guiding the selected SMEs in implementing the digital transformation framework, with a focus on ensuring that each component of the framework is effectively applied according to the SME's unique context and needs. Provide necessary training and support to facilitate the implementation process (Hevner et al., 2004). During the implementation, both quantitative and qualitative data are gathered throughout the implementation phase and post-implementation phase. Data collection methods will include surveys, interviews, observations, and document analysis. Also the data collection will cover key performance indicators (KPIs) related to digital transformation goals, such as revenue growth, cost reduction, customer satisfaction, and innovation (Yin, 2018).

The next step is to measure the extent to which the SMEs achieve their digital transformation goals using the framework. Comparing the post-implementation KPIs with the baseline data collected in the pre-implementation assessment. Analyzing the changes and improvements in digital maturity and other relevant metrics (Eisenhardt, 1989). The research will Engage in ongoing discussions and feedback sessions with the participating SMEs to understand their experiences, challenges, and successes during the implementation. Use this feedback to make any necessary adjustments or refinements to the framework or its application (Hevner et al., 2004). The entire case study process will be documented, including the pre-implementation assessment, framework implementation, data collection, and analysis. Maintaining a detailed record of the SMEs' progress and the changes observed in digital transformation outcomes (Yin, 2018).

In the analysis and findings phase, the collected data will undergo a rigorous examination to evaluate the framework's contribution to the attainment of digital transformation objectives, with a focus on recognizing prevalent patterns, emerging trends, and noteworthy correlations within the dataset, as advocated by Eisenhardt (1989). A validation process will be executed to ensure the veracity and dependability of the research findings, and a careful assessment will be conducted to ascertain the extent to which these results can be extrapolated to other SMEs in Oman or analogous contexts, aligning with the principles of validation and generalization expounded by Yin (2018), and here triangulation is used as the data is collected from different sources of theoretical literature, the focus groups, the case study and the confirmatory focus group of experts.

## References

- Accenture (2021) Digital Transformation. Available at: <https://www.accenture.com/us-en/insights/digital/digital-transformation> (Accessed: 30 October 2022).
- Adams, J.D. (2005) Transforming work (2nd ed.). New York, NY: Miles River Press. Available at: <https://read.amazon.com/> (Accessed: 9 April 2020).
- Al-Abri, M.Y., Rahim, A.A. and Hussain, N.H. (2018) 'Entrepreneurial Ecosystem: An Exploration of the Entrepreneurship Model for SMEs in Sultanate of Oman', *Mediterranean Journal of Social Sciences*, 9(6), p. 193. Available at: <https://www.richtmann.org/journal/index.php/mjss/article/view/10344> (Accessed: 11 August 2021).
- Alqassabi, M.A. (2020) 'INSIGHTS ON SUSTAINABILITY OF SMALL AND MEDIUM ENTERPRISES IN OMAN: A CONCEPTUAL FRAMEWORK', *International journal of economics and financial issues*, 10(1), pp. 209–218. Available at: <https://doi.org/10.32479/ijefi.9063>.
- Bharadwaj, A. et al. (2013) 'Digital Business Strategy: Toward a Next Generation of Insights', *MIS quarterly*, 37(2), pp. 471–482. Available at: <https://doi.org/10.25300/MISQ/2013/37:2.3>.
- Bharadwaj, N. and Shipley, G.M. (2020) 'Salesperson communication effectiveness in a digital sales interaction 1', *Industrial Marketing Management*, 90. Available at: <https://doi.org/10.1016/j.indmarman.2020.07.002>.
- Bongiorno, G., Rizzo, D. and Vaia, G. (2018) *CIOs and the digital transformation: A new leadership role*. Cham: Cham: Springer. Available at: <https://doi.org/10.1007/978-3-319-31026-8>.
- Butt, J. (2020) 'A Conceptual Framework to Support Digital Transformation in Manufacturing Using an Integrated Business Process Management Approach', *Designs*, 4(3), p. 17. Available at: <https://doi.org/10.3390/designs4030017>.
- Capgemini (2020) Digital transformation a road map for billion dollar organizations. Available at: [https://www.capgemini.com/wpcontent/uploads/2017/07/Digital\\_Transformation\\_A\\_Road-Map\\_for\\_BillionDollar\\_Organizations.pdf](https://www.capgemini.com/wpcontent/uploads/2017/07/Digital_Transformation_A_Road-Map_for_BillionDollar_Organizations.pdf)
- CRN Team (2020) Five Challenges SMEs face in their Digital Transformation Journey - CRN - India. Available at: <https://www.crn.in/news/five-challenges-smes-face-in-their-digital-transformation-journey/> (Accessed: 9 April 2021).
- Daft, R.L. (2010) *Organization theory and design Tenth Edition*, VANDERBILT UNIVERSITY.
- Deloitte (2019) Digital transformation framework. Available at: <https://www2.deloitte.com/content/dam/Deloitte/fr/Documents/processes/digital-transformation-framework.pdf>
- Eisenhardt, K.M. (1989) 'Building Theories from Case Study Research', *The Academy of Management Review*, 14(4), pp. 532–550. Available at: <https://doi.org/10.2307/258557>.
- Fenton, A., Fletcher, G. and Griffiths, M. (2020) *Strategic Digital Transformation : A results-driven approach*. Routledge.
- Forrester (2020) Digital transformation framework. Available at: <https://go.forrester.com/research/forrester-s-digital-transformation-framework/>

- Gebayew, C. et al. (2018) 'A Systematic Literature Review on Digital Transformation', in 2018 International Conference on Information Technology Systems and Innovation (ICITSI), pp. 260–265.
- Gong, C. and Ribiere, V. (2020) 'Developing a unified definition of digital transformation', *Technovation*, p. 102217. Available at: <https://doi.org/https://doi.org/10.1016/j.technovation.2020.102217>.
- Gupta Sunil, 1958- author (2018) *Driving digital strategy : a guide to reimagining your business*.
- Hevner, A.R. et al. (2004) 'Design science in information systems research', *MIS Quarterly: Management Information Systems*, 28(1). Available at: <https://doi.org/10.2307/25148625>.
- Johnston, M. (2014) 'Secondary Data Analysis: A Method of Which the Time has Come', *Qualitative and Quantitative Methods in Libraries*, 3, pp. 619–626.
- Kavadias, Stelios; Ladas, Kostas; Loch, C. (2016) *The 6 Elements of Truly Transformative Business Models*, Hbr. Available at: <https://hbr.org/2016/10/the-transformative-business-model> (Accessed: 26 February 2021).
- Khezr, S. et al. (2019) 'Blockchain technology in healthcare: A comprehensive review and directions for future research', *Applied Sciences (Switzerland)*, 9(9). Available at: <https://doi.org/10.3390/app9091736>.
- Kotter, J.P. (2012) *Leading Change* - John P. Kotter - Google Books, 1996.
- Libarikian, A. (2019) *Future of Digital Transformation with McKinsey | CXOTalk*. Available at: <https://www.cxotalk.com/video/mckinsey-leap-beyond-digital-transformation> (Accessed: 25 February 2021).
- Muthuraman, S. et al. (2020) 'Framework for Oman's Economic Development through Small & Medium Enterprise', *International Journal of Business and Applied Social Science*, pp. 1–7. Available at: <https://doi.org/10.33642/ijbass.v6n6p1>.
- Nwaiwu, F. (2018) 'Review and Comparison of Conceptual Frameworks on Digital Business Transformation', *Journal of Competitiveness*, 10(3), pp. 86–100. Available at: <https://doi.org/10.7441/joc.2018.03.06>.
- Ojala, A. (2016) 'Business models and opportunity creation: How IT entrepreneurs create and develop business models under uncertainty: Business models and opportunity creation', *Information systems journal (Oxford, England)*, 26(5), pp. 451–476. Available at: <https://doi.org/10.1111/isj.12078>.
- Osterwalder, A. and Pigneur, Y. (2013) 'Designing Business Models and Similar Strategic Objects: The Contribution of IS', *Journal of the Association for Information Systems*, 14(5), pp. 237–244. Available at: <http://search.proquest.com/docview/1418397281/>.
- 'Oxford English dictionary' (2000). Oxford.
- Peppers, K. et al. (2007) 'A design science research methodology for information systems research', *Journal of Management Information Systems*, 24(3). Available at: <https://doi.org/10.2753/MIS0742-1222240302>.
- Perkin, N. and Abraham, P. (2017) *Building the agile business through digital transformation*. 1st Editio. Edited by P. Abraham 1960- author. London .
- Ragin, C.C. (2014) *The comparative method: moving beyond qualitative and quantitative strategies*. 1st edn. Berkeley: University of California Press. Available at: <https://doi.org/10.1525/j.ctt6wqbwk>.

- El Sawy, O.A. and Pereira, F. (2013) 'VISOR: A Unified Framework for Business Modeling in the Evolving Digital Space', in. Available at: [https://doi.org/10.1007/978-3-642-31765-1\\_3](https://doi.org/10.1007/978-3-642-31765-1_3).
- Soule, D.L. et al. (2015) 'Becoming a Digital Organization: The Journey to Digital Dexterity', SSRN Electronic Journal [Preprint]. Available at: <https://doi.org/10.2139/ssrn.2697688>.
- van Tonder, C. et al. (2020) 'A FRAMEWORK FOR DIGITAL TRANSFORMATION AND BUSINESS MODEL INNOVATION', *Management (Split, Croatia)*, 25(2), pp. 111–132. Available at: <https://doi.org/10.30924/mjcmi.25.2.6>.
- Ulas, D. (2019) 'Digital Transformation Process and SMEs', *Procedia computer science*, 158, pp. 662–671. Available at: <https://doi.org/10.1016/j.procs.2019.09.101>.
- Veronica, S. et al. (2020) 'International social SMEs in emerging countries: Do governments support their international growth?', *Journal of world business : JWB*, 55(5), p. 100995. Available at: <https://doi.org/10.1016/j.jwb.2019.05.002>.
- Walls, J.G., Widmeyer, G.R. and El Sawy, O.A. (1992) 'Building an information system design theory for vigilant EIS', *Information Systems Research*, 3(1). Available at: <https://doi.org/10.1287/isre.3.1.36>.
- Westerman, G. and McAfee, A. (2012) A major research initiative at the MIT Sloan School of Management Research Brief The Digital Advantage: How Digital Leaders Outperform Their Peers in Every Industry i. Available at: <http://digitalcommunity.mit.edu/docs/DOC-1105>. (Accessed: 25 February 2021).

# Cybersecurity Investments Metrics using FAIR-ROSI

## Abstract

*Traditional return on security investment (ROSI) models often emphasize investment costs and anticipated returns but overlook risk-related factors and qualitative cybersecurity metrics. To address this oversight, this paper employs an aggregation strategy that integrates five selected qualitative and quantitative metrics with the Factor Analysis of Information Risk (FAIR) model for risk analysis and quantification. The study pioneers the fusion of FAIR-ROSI models, combining practical qualitative and quantitative indicators to enhance the granularity of the traditional ROSI model. A case study is utilized to evaluate the proposed metrics. Empirical data from pre- and post-control measures reveal a narrow margin between actual and projected loss values and a significantly higher ROI compared to total security expenditure. The integration of FAIR model and ROSI model addressed the limitations found in traditional ROSI models concerning risk assessment. Such integration fosters a holistic approach to ROI and risk management, thereby facilitating informed decision-making.*

**Keywords:** Risk Assessment, Return on Security Investment (ROSI), Factor Analysis of Information Risk (FAIR), FAIR-ROSI model

## 1.0 Introduction

### 1.1 Current State of Cyber Security Investment

In light of the prevailing state of network security, investment in cyber security is imperative. Global reports indicate that the worldwide cybersecurity market has already surpassed 180 billion U.S. dollars. Network security investments serve multifaceted functions, which include but are not limited to protecting user data, mitigating financial disruptions, rectifying vulnerabilities, and preserving organisational reputation. These investments enable organisations to adhere to local policies and regulations, mitigating the risk of penalties and closures. By maintaining consistent and sustainable network security investments, organisations can proactively address evolving network challenges and secure both network and asset integrity (Fielder et al., 2016; Zamani et al., 2020).

Nonetheless, global underinvestment in cybersecurity persists as a pressing issue. Many Chief Information Security Officers (CISOs) are concerned about inadequate or unstable budgets for cybersecurity investments. Investment in cyber security is

usually lacking mainly because it does not directly generate revenue (Fedele & Roner, 2022; Gordon et al., 2018; Lee, 2021). Their purpose is to mitigate the losses stemming from security incidents (He et al., 2022; Janicke et al., 2021; Smith et al., 2021). This is compounded by the lack of understanding among investors, as well as the challenges in quantifying return on investment (ROI) (Armenia et al., 2021; Fedele & Roner, 2022; He et al., 2022; Loft et al., 2022). Consequently, security investments have become a critical issue for investors and CISOs across diverse organisations.

## **1.2 Limitations of the ROSI model and complementarity of the FAIR model**

To quantify cost-effectiveness, the European Network and Information Security Agency (ENISA) (2012) proposed the Return on Security Investment (ROSI) model, which provides an exhaustive quantitative risk assessment of the financial impact arising from security incidents that influence ROSI. While it has an economic lens, emphasizing the importance of reputation protection and ensuring regulatory compliance, the ROSI model overlooks non-economic dimensions, which narrows its viewpoint. Furthermore, the ROSI model, has an imbalanced focus on quantitative metrics, lacking qualitative security metrics that yield significant influence over cybersecurity investment choices. Relying solely on ROI can lead to an inadequate grasp of the complexities and challenges within an organisation's cybersecurity landscape. For example, according to Kesswan & Kumar (2015), using the ROSI model and focusing solely on Single Loss Expectancy (SLE), Annual Rate of Occurrence (ARO), and Annual Loss Expectancy (ALE) does not provide a robust and comprehensive assessment. The ROSI model, while simpler and useful for risk assessment, covers only a portion of the risk landscape. It does not account for interactions between metrics or the influence of other threatening events in the organisation, and it lacks the detailed information needed to quantify the likelihood of various risk events and losses.

This paper addresses the previously mentioned challenges by integrating the FAIR model with the ROSI model. The FAIR model offers a clearer framework for categorizing cybersecurity risks into quantifiable risk factors. In particular, it considers multiple risk factors and translates both qualitative and quantitative metrics into quantitative loss figures. Furthermore, the FAIR model calculates the level of risk that can be reduced by adopting a given countermeasure, as well as the loss figures



that will be incurred thereafter. When employed alongside ROSI, the FAIR model can be used as a reference, ensuring that the conclusions drawn from ROSI are both rigorous and objective. By combining the FAIR model and ROSI model, the proposed methodology provides a systematic and comprehensive strategy for risk assessment and assisting cyber security investment decision-makers in making more precise and well-informed decisions.

## **2.0 Related Work**

### **2.1 ROSI and FAIR model**

Cybersecurity investment, while not yielding direct returns, primarily aims to mitigate losses from cybersecurity issues. Consequently, it has yet to garner significant attention from investors. Due to the complexities of accurately evaluating and quantifying myriad cyber risks, leveraging quantitative outcomes for investment decisions proves to be more convincing (He et al., 2023; Loft et al., 2021). Kesswan & Kumar (2015) employed a cost-benefit analysis in cybersecurity decision-making. They elucidated the computation of return on investment (ROSI) in cybersecurity through real-world instances, meticulously accounting for variables impacting ROSI calculation, such as Single Loss Expectancy (SLE), Annual Rate of Occurrence (ARO), Annual Loss Expectancy (ALE), and various approaches like Annual Compliance Benefit (ACB), New Business Benefits (NBB), Goodwill Loss Expectancy (GLE), and Shared Reduced Loss (SRL). This approach, although focusing on ROI and loss-contributing variables, needs more integration and specific quantification of other qualitative and quantitative metrics.

The model proposed by Gordon & Loeb (2002), a pivotal outcome in information security economics, underscores organisations concentrating investments primarily on moderate information vulnerability levels during normal conditions. Through analysis of two vulnerability probability function types, it highlights that optimal security investment expenditure can either increase with vulnerability level or eventually decrease proportionally. Typically, most organisations' cybersecurity investments remain below 37% of anticipated loss, often much less. Consequently, targeted investments safeguard information assets with moderate threat levels. This paper can be employed alongside the model's investment cost-to-expected loss ratio, further

reinforcing the validity of indicator and model selection by horizontally comparing ratios generated using this paper's risk assessment model and identified indicators. However, the above research, while focusing on cost-benefit analysis, overlooks non-economic influences and other qualitative and quantitative indicators, thus falling short of achieving a comprehensive organisational risk assessment. In this context, a risk assessment model that integrates multiple influencing factors is imperative. Wang et al. (2020) introduced one of the more widely recognized quantitative risk assessment models—the Factor Analysis of Information Risk (FAIR) model. The FAIR model offers a clearer framework for categorizing cybersecurity risks into a set of quantifiable risk factors. Within this, risk (R) is derived from the multiplication of loss event frequency (LEF) and loss magnitude (LM). Moreover, the primary factors influencing loss event frequency (LEF) are derived from the threat event frequency (TEF) and vulnerability (V), where vulnerability (V) is further constituted by both threat capability (TC) and control strength (RS). In terms of Loss Magnitude (LM), it comprises both Primary Loss Magnitude (PLM) and Secondary Loss Magnitude (SLM). This paper has opted to select the FAIR model and integrate it with the ROSI model, fostering a comprehensive assessment approach that simultaneously accounts for both organisational risk factors and cost-effectiveness.

## **2.2 Cybersecurity Qualitative Metrics and Assessment**

Catota et al. (2018) enhanced threat response capabilities by proposing the Computer Security Incident Response Teams (CSIRTs) and promoting information sharing. They emphasized the significance of qualitative metrics such as threat responsiveness. However, they also noted a lack of comprehensive analysis regarding critical infrastructures, especially in quantifying the impact of security responsiveness metrics on an organisation's security capabilities. Additionally, Naseer et al. (2021) observed that the analytical information utilized in cybersecurity incident response (CSIR) is largely application-specific (e.g., analyzing advanced persistent threat traffic and examining server logs). They outlined four categories of analytical insights—real-time, forensic, predictive, and descriptive—alongside mechanisms like threat hunting, anomaly detection, and continuous monitoring. Proactive CSIR facilitates comprehension of adversary behavior and timely threat intelligence generation. The study offered methods for obtaining threat information and four criteria for analyzing such information. However, it did not specify particular sub-indicators to focus on

during the cybersecurity response process, such as organisational response time and recovery time.

Georgiadou et al. (2022) evaluated the qualitative indicator of inherent threat and established a security culture framework with two levels (individual and organisational) across nine dimensions. The framework analyzed diverse domains to derive corresponding inherent threat factors. Given the subjective nature of human-induced intrinsic threat's specific impact on qualitative indicators (threat level) and the challenge of precise quantification, this paper refrains from delving into the exact quantification of intrinsic threat. Nonetheless, Zadeh et al. (2020) comprehensively delineated the current cybersecurity threats organisations confront, categorizing them into four groups: physical threats, human threats, communication and data threats, and operational threats. The analysis highlighted the higher sensitivity of the IT and financial sectors to cybersecurity concerns, with physical threats being relatively less significant in these industries. The study provided a lens to emphasize the degree of cyber threats in the IT and financial sectors. It also depicted the four threat categories under Microsoft's Threat Model (STRIDE) and the NIST SP 800-30 standard, contributing to this paper's identification and reference of qualitative metrics.

Numerous organisations currently adopt a qualitative risk matrix for cybersecurity based on the NIST 800-30 global standard. However, relying solely on such an approach essentially ties cyber risk assessment to professionals' future predictions regarding specific attacks. Allodi & Massacci (2017) extend this perspective by introducing a quantitative assessment approach, evaluating the likelihood of non-targeted attacks through endpoint defence and periodic vulnerability assessment exercises. This method quantifies the likelihood of an attack more precisely. While this quantitative analysis offers a more scientific risk assessment method, concentrating solely on the number of vulnerabilities as a metric falls short in practice. Furthermore, Ghani et al. (2013) quantitatively evaluated software vulnerability qualitative indicators by leveraging the Common Vulnerability Scoring System (CVSS). This quantitative assessment methodology aids in prioritizing security investments, thereby minimizing losses from security vulnerabilities and optimizing the utility of security investments. Nevertheless, Houmb et al. (2010) scrutinized the qualitative impact indicators of CVSS and determined that focusing solely on the number of vulnerabilities results in only a 4% reduction in security risk, necessitating concentration on thousands of vulnerabilities. The collective findings of

these studies imply a need for introducing not only qualitative metrics analysis but also an emphasis on the impact of multiple quantitative metrics.

### **2.3 Cybersecurity Quantitative Metrics and Assessment**

As for the cybersecurity metrics, Ma (2021) combined qualitative and quantitative metrics to generate quantifiable data for structural comparisons. When assessing network vulnerability, it looks into the attack process, passive detection and active detection. The paper introduces several common network vulnerability assessments, including qualitative and quantitative approaches. The assessment of system security risk emphasizes assets and threat rate, but the evaluation of threat loss remains insufficient. Regarding the specific analysis and calculation of quantitative metrics, Kim (2019) introduced a decision-making method for estimating malware risk indices. The study analyzes the probability of malware and malicious activities (MAs) using a decision model that incorporates static and dynamic analyses to detect, identify, and classify various malicious activities and threat sources. By utilizing hierarchical analysis, it quantitatively assesses and quantifies the malware threat indices and subsequently examines the probability of malware and MAs. The primary emphasis is on quantitative metrics, probability of malware, probability of MAs.

Wang (2020) conducted an in-depth analysis of the vulnerability detection technique—the fuzzing technique. The findings from this study can provide improved detection tools for organisations using historical metrics statistics. CRISTEA (2021) introduced five significant security events: ransomware, malware, advanced persistent threats (APT), third-party threats, and external actor sabotage. The study proposed a more practical risk management model primarily centered on analyzing the frequency of these five types of security events, commonly referred to as the frequency of the threat. Although Allodi & Massacci (2017) acquired a risk metric by adopting quantitative risk analysis in conjunction with traditional qualitative analysis, organisations seeking a more intuitive organisational risk value need to estimate threat probability and subsequently establish a targeted baseline based on the calculated risk value. In this context, a suitable approach to assessing threat frequency involves initially comprehending the nature of the security event and then estimating the associated threat probability.

### 3.0 Methodology

This section introduces the proposed FAIR-ROSI model. This model aims to evaluate the rationality of cyber security investment decisions through the results of ROSI. The calculation of ROSI is based on the components and quantification methods in the FAIR model. In particular, we map a series of threat response properties, including threat level, threat frequency, threat response capability, and an added potential risk component, to FAIR components. This mapping is inspired by Qamar, S. at al. (2017)'s work, we map the threat response properties to FAIR components that are semantically similar and comparable. The threat response properties serve as a bridge between evaluating FAIR components and calculating ROSI, as these are the key properties used by both FAIR and ROSI. In this section, we first introduce the components and quantification methods of the FAIR model, then explain how the qualitative and quantitative metrics of the threat response properties map to FAIR components, and finally introduce the baseline settings of the qualitative and quantitative metrics. Figure 1 shows the high-level proposed FAIR-ROSI model with specific qualitative and quantitative metrics elaborated in Section 3.2.

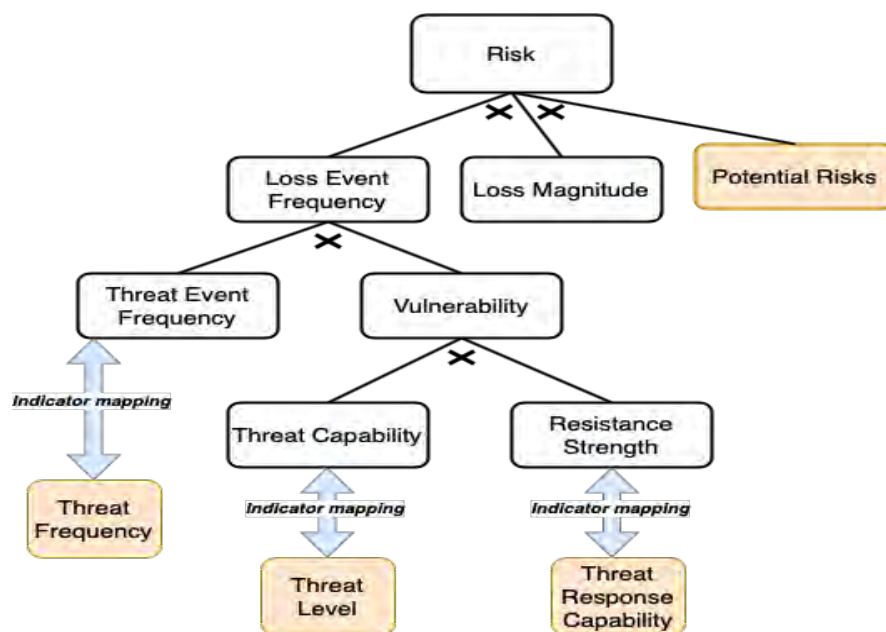


Figure 1. High Level FAIR-ROSI Model

#### 3.1 The components and the quantification in the FAIR model

As shown in section 2.1 and Figure 1, the FAIR model has eight key components. Loss Event Frequency (LEF) refers to the frequency of loss events occurrence. LEF is

affected by two sub-metrics: Threat Event Frequency (TEF) and Vulnerability (V). TEF refers to the likelihood of a threat event, while V denotes the susceptibility of an organisation's assets to a threat event. TEF can be further influenced by Contact Frequency (CF) and Probability of Action (PoA), and V is impacted by Resistance Strength (RS) and Threat Capability (TC), respectively. CF refers to the frequency at which a threat endeavors to exploit an asset, and PoA refers to the likelihood that a threat will execute a detrimental action after accessing an asset. RS assesses the effectiveness of organisations' security controls, while TC reflects the capability of a threat. The precise formulas are outlined as follows:

$$LEF=TEF*V \quad (1)$$

$$TEF=CF*PoA \quad (2)$$

$$V=RS*TC \quad (3)$$

Loss magnitude (LM) represents the magnitude of the loss expected at the time of the loss event. It comprises a primary loss magnitude (PLM) and a secondary loss magnitude (SLM). Among them, PLM is the loss directly caused by the loss event, and SLM is the further loss due to the primary loss. In this study, we will focus on the effect of PLM on FAIR modeling. The formula is outlined as follows:

$$LM=PLM+SLM \quad (4)$$

Risk (R) is a combination of loss event frequency and loss magnitude. The formula is outlined as follows:

$$R=LEF*LM \quad (5)$$

### **3.2 The Qualitative and Quantitative Metrics**

This section introduces and expands the qualitative and quantitative metrics as shown in Figure 1.

The Qualitative Metrics are listed below,

**Threat Responsiveness Capacity** assesses the level of response in the face of cyber threats, and consists of three sub-metrics: **response time**, indicating the time required to identify and respond to a threat event; **recovery time**, which signifies the duration to restore normal operations; and the **effectiveness of control measures**, quantifies the extent to which damage from cyber threats can be mitigated through effective control measures. These three metrics together cover the before (control measures), during (response time), and after (recovery time) stages of a cyber threat event, providing a holistic assessment of the organization's cybersecurity posture, making them a good measure of Threat Responsiveness Capacity.

Threat Responsiveness Capacity could be used to evaluate RS in FAIR model. It provides a more comprehensive view of an organization's ability to respond to cyber threats, which not only considers RS, but also the speed of response and recovery, which are critical aspects of managing cyber threats.

**Threat Level** refers to the specifics of cyber threats encountered by organisations. The specific quantification of the threat level is expanded to three sub-metrics: **threat potential** refers to how likely a specific threat is to occur; **severity of consequences** indicating the potential damage a threat could cause; and the **scope of threat**, which is the breadth of impact a threat could have. They collectively provide a detailed picture of the threat level.

In FAIR-ROSI model, threat level could map to TC, since according to the definition of TC, threat potential could be seen as the likelihood of a threat agent applying force against an asset. The severity of consequences and scope of threat further detail the potential impact of this force, which aligns with the FAIR model's focus on quantifying risk. This model will quantify the likelihood, severity, and scope of impact using a 0-1 scoring system. Given that TC is subject to change, the quantification of threat level ensures that the organisations routinely reassess the magnitude of TC. This quantification approach aligns with the requisites of TC and contributes to increased operability and dynamic adaptability.

**Potential Risk** is used to evaluate various future cyber threats an organisation could face. The two sub-metrics are **future threat likelihood**, **severity of future threat**. By incorporating Potential Risk, the FAIR model will be capable of accounting for risks that may emerge in the future. This model uses a 0-1 scoring system to quantify the likelihood and severity of the consequences of potential risks.

The Quantitative Metrics are listed below,

**Cybersecurity Return on Investment (ROSI)** is an economic metric used to assess the effectiveness of an investment by measuring the return generated. ROSI is calculated by utilizing Loss Event Frequency (LEF) and Loss Magnitude (LM) determined in the FAIR model, and investment guidance is provided based on the ROSI calculation results. A positive result indicates a favorable return on the investment, while a negative result suggests that the investment expectations have not been met.

**Threat frequency** is the actual number of times a specific threat event occurs within a designated time frame. Given that threat frequency relies on observed actual data for determination and can be tracked and updated through logs, monitoring systems, and other data sources, mapping threat frequency to TEF allows the abstract concept to be translated into concrete observed threat frequency data.

### 3.3 Baseline Setting

Each organisation establishes its specific acceptable values for the metrics based on its network security environment and security capacity. These values represent the organisation's "normal" state for network performance metrics, and any metric deviations from this baseline, particularly those falling below it, can be deemed anomalies or risks. This model integrates the principles of NIST Special Publication 800-53 Rev. 4 for determining high and medium-low impact security systems. It primarily centers on the cybersecurity investment objectives of the organisation, utilizing the historical security data of each organisation as a foundation. The analysis will be based on the average value of each metric across organisations.

#### 3.3.1 Baseline Setting for Qualitative Metrics

To establish the value for **Threat Responsiveness**, the assumption is that organisations can automatically compute values of the three sub-metrics, **response time**, **recovery time**, and **effectiveness of control measures**. This computation can be accomplished by utilizing the log and event monitoring system, the security event management system (SIEM), and the fault ticketing system during the acquisition of threat responsiveness data. Threat responsiveness assessment is based on the arithmetic mean of the response time, recovery time, and effectiveness of control measures.

To calculate the average response time (ART) of the sub-metrics, let's denote  $n_R$  as the number of response times in the organisation, and  $R_i$  as the  $i^{\text{th}}$  response time. Then, the formula for calculating ART is as follows:



$$ART = (1/n_R) * \sum R_i \quad (6)$$

Calculate the Mean Recovery Time (MRT) for the sub-metric, where  $n_M$  in the organisation denotes the number of recovery times and  $M_i$  denotes the  $i^{th}$  recovery time, then the MRT is:

$$MRT = (1/n_M) * \sum M_i \quad (7)$$

Calculate the Average Control Effectiveness (ACE) of the submetrics, the organisation where  $n_C$  denotes the number of controls and  $E_i$  denotes the effectiveness of the  $i^{th}$  control, then ACE is:

$$ACE = (1/n_C) * \sum E_i \quad (8)$$

The qualitative metrics of *Threat Level* are assessed qualitatively based on *threat potential, severity of consequences, and scope of threat*, with assessment levels of low, medium, and high, corresponding to a specific range of values of 0-0.33, 0.34-0.66, and 0.67-1.0. Threat level assessment is based on the arithmetic mean of the three sub-metrics. This assessment is combined with the organisation's historical data, including analysis of threat intelligence, past incidents, security logs, and other pertinent information. The details are shown in Table 1.

Threat Level assessment form			
Threat potential	Severity of consequences	Scope of threat	Numerical range
Highly unlikely	Minor impact	Affects a small part	0.0 - 0.33
Likely	Medium impact	Affects a part of the region	0.34 - 0.66
Highly likely	High impact	Affects entire organisation	0.67 - 1.0

**Table 1. Qualitative threat level assessment form**

The qualitative metrics of *Potential Risk* is assessed qualitatively based on the *future threat likelihood, severity of future threat*. Potential risk assessment is based on the arithmetic mean of the future threat likelihood and the severity of future threat. The specifics are as follows in Table 2.

Potential Risks Assessment Form		
Future threat likelihood	Severity of future threat	Numerical range

Highly unlikely	Minor impact	0.0 - 0.33
Likely	Medium impact	0.34 - 0.66
Highly likely	High impact	0.67 - 1.0

**Table 2. Qualitative Assessment of Potential Risks Form**

### 3.3.2 Baseline Setting for Quantitative Metrics

In this approach, the baseline of *ROSI* is set to 0 with the following considerations: a ROSI baseline set to 0 is the lowest investment baseline for the organisation. It is in line with the principle of zero return on investment in economics because the purpose of investment is to obtain greater benefits. When the ROSI is less than 0, it means that the investment does not have any value in the analysis of the benefits of the investment. Additionally, the indicator is easy to compare and analyze. By setting the baseline to 0, it provides a greater incentive for the organisation, and the members of the organisation can easily achieve a ROSI greater than 0.

*Threat Frequency* (TF) can usually be tracked and updated based on logs, monitoring systems, and other data sources. TF can be calculated using Loss Magnitude, Cost of Each Breach (CEB) and Number of Total Breaches (TB),

$$TF = LM/CEB/NTB. \quad (9)$$

The pseudocode in Section 4.3 shows in detail the process of how to use qualitative and quantitative metrics to evaluate ROSI through the quantitative method of the FAIR model.

## 4.0 Case Study

### 4.1 Case Introduction

Due to the limited accessibility of organisations network, obtaining accurate and sensitive network security data is challenging. This paper relies on statistics from the 2023 Cost of a Data Breach Report by IBM and supplements this data with reasonable assumptions. The report provides insights into the cost of a data breach across industries and sizes, considering various attack vectors, the average detection and containment time for different vulnerabilities or threats, and the magnitude of impact associated with various data breaches. Additionally, the report examines the financial

consequences of a data breach and offers insights into the factors influencing these costs. The case hypotheses in this paper will be based on specific data from the report. Company A is a professional network security service provider, and Company B is a new e-commerce platform with a good reputation. Companies A and B are both hypothetical companies based on real data from the IBM 2023 data breach report. Due to the rapid expansion of Company B's business in recent years, it has experienced an increase in phishing incidents, leading to substantial user data leakage. It has resulted in business interruptions, affecting the organisation's normal operations and causing significant losses.

#### **4.2 Case Statistics**

Company A reviews Company B's basic situation and discovers that Company B operates in the network technology industry with a workforce of 1,000 to 5,000 employees. The primary challenge faced by Company B is the substantial data leakage of customer, employee, and anonymous data due to numerous phishing incidents. The statistics reveal that the total cost of losses amounts to \$5,360,000, corresponding to the loss magnitude (LM). The main data leakage stems from a significant number of phishing incidents, accounting for \$5,360,000. Among these, the leakage of customer information constitutes a major portion, with each data leakage incident costing \$183.

Company B then implemented a range of security measures, including Intrusion Detection Systems (IDS), Intrusion Prevention Systems (IPS), and Multi-Factor Authentication (MFA), as well as the incorporation of security artificial intelligence and automated detection. Following the implementation of these control measures, the new loss magnitude (NLM) is \$2,650,000. During a specific time period for statistics on security metrics, the following results are obtained: a response time of 167 days, a recovery time of 47 days, and effectiveness of control measures at 95%. The threat likelihood metric is recorded at 0.4, the threat consequence severity is 0.6, and the threat impact range is 0.65. The potential risk likelihood metric is measured at 0.33, and the potential risk consequence severity is 0.6. Additionally, the threat frequency is found to be 27 times per year.

In this context, Company A needs to assess the situation, compares it with the baseline and advise on the return on the cyber security investment.

### 4.3 Baseline Setting

#### 4.3.1 Setting Qualitative Baseline

According to the “Cost of a Data Breach Report 2023” (IBM Security, 2023), IBM conducted statistical analysis to determine the averages of the two sub-metrics within the qualitative metrics for *threat response capability*. These are an average *response time* of 217 days and an average *recovery time* of 76 days. Furthermore, it’s assumed that the *average effectiveness* of the organisation's current control measures is 85%. As for the Threat Level, we assume the average values, derived from statistical analysis and assessment, for the organisation's *threat potential*, *severity of consequences*, and *scope of threat* to be 0.5, 0.8, and 0.8, respectively. For *Potential Risks*, we assume that the average value of the likelihood of potential risk values in the historical data of the organisation is 0.4, while the average value of the severity of potential risk consequences is 0.8.

#### 4.3.2 Setting Quantitative Baseline

As mentioned in Section 3.3.2, the baseline of *ROSI* is set to 0. It is in line with the principle of zero return on investment in economics because the purpose of investment is to obtain greater benefits. When the ROSI is less than 0, it means that the investment does not have any value in the analysis of the benefits of the investment.

As for the *Threat Frequency*, from the report, the cost of data leakage is \$5360000 per year. The main data breaches involve customer information leakage problems, and the cost of each data breach is \$183. In this regard, it is estimated that the number of data breaches caused by the number of data breaches per year is  $5360000/183=29290$ . In response, there are a total of 29,290 data breaches per year, so, taking into account the size of the organisation and the cost of a data breach, assuming that each threat event (e.g., phishing event) results in an average of about 800 data breaches, the annual threat frequency is  $29290/800=37$ .

Table 3 presents the data means of each metric in the study. It should be noted that some of the metric means are hypothetical values based on objective and real data.

Overall cybersecurity metrics	sub-metric	Mean value
Threat Response Capability	Response Time (days)	217
	Recovery Time (days)	76

	Effectiveness of Control Measures (%)	85
Threat level	Threat Likelihood	0.5
	Threat consequence severity	0.8
	Scope of threat impact	0.8
Potential risks	Likelihood of potential risk	0.4
	Severity of consequences of potential risks	0.8
Threat Frequency	Threat frequency (times/year)	37
ROSI	ROSI	0

**Table 3. Mean and baseline size of qualitative quantitative indicators**

#### 4.4 Use FAIR-ROSI Model to Analyze Return on Security Investment

Considering the statistics from the report (see Section 4.2) as inputs as well as reasonable assumptions, we calculated the return on investments after the organisation implemented security measures. The specific analysis process is outlined as follows using pseudocode.

The Algorithm CalculateNLEFandROSI(CompanyData)

Input: CompanyData containing information on baseline metrics and impact scores

Output: NLEF (New Loss Event Frequency), RiskMitigationValue, ROSI (Return on Security Investment)

1. Initialize LEF to 1, ROSI to 0, and RiskMitigationValue to \$2710000

// Calculate composite scores for threat response capability

2.  $NewResponseTimeScore = BaselineResponseTime / NewResponseTime$

3.  $NewRecoveryTimeScore = BaselineRecoveryTime / NewRecoveryTime$

4.  $NewControlEffectivenessScore = NewControlEffectiveness / BaselineControlEffectiveness$

5.  $CompositeResponseScore = (NewResponseTimeScore + NewRecoveryTimeScore + NewControlEffectivenessScore) / 3$

// Calculate composite scores for threat severity

6.  $NewThreatLikelihoodScore = BaselineThreatLikelihood / NewThreatLikelihood$

7.  $NewThreatDetectionTimeScore = BaselineThreatDetectionTime / NewThreatDetectionTime$

8.  $NewThreatImpactScore = BaselineThreatImpact / NewThreatImpact$

9.  $CompositeSeverityScore = (NewThreatLikelihoodScore + NewThreatDetectionTimeScore + NewThreatImpactScore) / 3$

// Calculate composite scores for potential risks

10.  $NewPotentialThreatLikelihoodScore = BaselinePotentialThreatLikelihood / NewPotentialThreatLikelihood$

11.  $\text{NewPotentialThreatImpactScore} = \text{BaselinePotentialThreatImpact} / \text{NewPotentialThreatImpact}$

12.  $\text{CompositePotentialRiskScore} = (\text{NewPotentialThreatLikelihoodScore} + \text{NewPotentialThreatImpactScore}) / 2$

// Calculate composite scores for threat frequency

13.  $\text{NewThreatFrequencyScore} = \text{BaselineThreatFrequency} / \text{NewThreatFrequency}$

// Calculate NLEF

14.  $\text{NTEF} = 1 / \text{NewThreatFrequency}$  /

15.  $\text{NPR} = 1 / \text{CompositePotentialRiskScore}$

16.  $\text{NV} = (1 / \text{CompositeResponseScore}) * (1 / \text{CompositeSeverityScore})$

17.  $\text{NLEF} = \text{NTEF} * \text{NV} * \text{NPR}$

// Calculate RiskMitigationValue

18.  $\text{RiskMitigationValue} = \text{BaselineLossMagnitude} - \text{NLEF}$

// Calculate ROSI

19.  $\text{ALEBefore} = \text{BaselineAnnualLossExpectancy}$

20.  $\text{ALEAfter} = \text{ALEBefore} - \text{RiskMitigationValue}$

21.  $\text{C} = \text{TotalAnnualSecurityCost}$

22.  $\text{ROSI} = (\text{ALEBefore} - \text{ALEAfter}) / \text{C}$

23. RETURN NLEF, RiskMitigationValue, ROSI

The analysis process includes the calculation of composite scores for threat response capability, composite scores for threat severity, composite scores for potential risk, risk mitigation value and ROSI. We aim to compare the loss before and after applying the new security measures which is detailed in Section 5.0.

## 5.0 Results Analysis

The results are reported in Table 4. Based on the results of IBM Security (2023) actual statistics, Company B reduced the actual loss margin (the actual value at risk) from \$5,360,000 to \$2,650,000 by implementing a series of security measures. When using the research method and research metrics identified in this paper, the calculated expected risk loss value was \$2,551,701 per year. A comparison revealed that the actual risk loss value is closer to the expected risk loss value. The expected risk loss value is relatively close. Additionally, when comparing with Company B's costs for

detection and notification upgrades totaling \$1,950,000 per year, it becomes more evident that the degree of risk mitigation (i.e., the reduction in risk value) significantly exceeds the costs incurred by Company B.

Furthermore, to better visualise the benefits of this investment, this paper uses the risk mitigation value to calculate the Return on Security Investment (ROSI), resulting in a ROSI calculation of 1.44. A comparison indicates that this value (1.44) is significantly greater than the baseline ROSI of 0, signifying a very high benefit from the investment. Of course, before investing, Company B can establish its risk tolerance level according to its specific circumstances and organisational fundamentals. It can also choose not to implement security measures if the calculated risk value remains below the established threshold.

Overall cybersecurity metrics	Metrics	Before implementation of security measures (Company B baseline)	After implementation of security measures
Threat Response Capability	Composite Score	1	1.21
	Response time (days)	217	167
	Recovery time (days)	76	47
	Effectiveness of control measures (%)	85	95%
Threat level	Composite Score	1	1.15
	Threat Likelihood	0.5	0.4
	Threat consequence severity	0.8	0.6
	Scope of threat impact	0.8	0.65
Potential risks	Composite Score	1	1.15
	Potential Risk Likelihood	0.4	0.33
	Potential Risk Consequence Severity	0.8	0.6
Threat Frequency	Composite Score	1	1.32
	Threat Frequency (times/year)	37	27
Other	ROSI	0	1.44
	Actual Loss Value	\$5,360,000	\$2,650,000

**Table 4. Changes in indicators and composite scores for qualitative and quantitative indicators before and after control measures**

## 6.0 Discussion

This paper is the first attempt to combine the FAIR model and the ROSI model. It uses the FAIR model to calculate the risk mitigation value and directly incorporates it into the ROSI model for calculation. With the FAIR model's quantitative risk analysis, this research connects the value of risk with both Loss Event Frequency (LEF) and Loss Magnitude (LM). It fully addresses the multiple variables impacting LEF,

including Threat Event Frequency (TEF) and Vulnerability (V), capturing the maximal influence that these factors exert on the value of risk. Through the utilization of the FAIR model to quantify a plethora of qualitative and quantitative metrics, it amends and enhances the LEF, offering a precise basis for organisational risk assessment. Combining these models effectively leverages the strengths of quantitative and qualitative analyses. The FAIR model furnishes precise values, while the ROSI model integrates economic considerations. This comprehensive analysis empowers organisations to gain a comprehensive understanding of the impact of diverse threats and risks on the organisation. Furthermore, the FAIR model provides quantitative risk scenarios, and the ROSI model measures the relationship between risk mitigation values and costs. This paper's fusion of the FAIR and ROSI models serves not only to quantify risk and assess the effectiveness of risk mitigation but also to merge risk management with cost-effectiveness. The result is a more holistic, precise, and targeted risk management analysis within organisations. This combination facilitates a superior grasp and measurement of risk, leading to more informed decision-making and the formulation of optimal risk management strategies, particularly when dealing with limited resources.

Although the FAIR-ROSI model offers valuable insights, one constraint is that it is applied to a hypothetical case derived from the IBM 2023 Data Breach Report (IBM Security, 2023), and not to an actual organization with real-life data. Applying the model to empirical data from an actual case study would greatly strengthen its validity and enhance its practical contribution. This step would allow for a more nuanced understanding and validation of the model in a practical context. However, the extensive and complex nature of conducting such an empirical study necessitated its exclusion from this current research phase. Looking forward, we plan to address this gap in a comprehensive, separate work, intended for future journal publication. The next phase of our research will focus on a detailed exploration and analysis of the model using empirical data and real-world cases, aligning with the rigorous standards required for a thorough empirical investigation. This future work promises to not only validate but also potentially refine our model, offering robust, practical insights applicable in real-world scenarios.

## **7.0 Conclusion**



In summary, based on the selected qualitative and quantitative metrics, this paper represents the first attempt to combine the FAIR model and ROSI model to quantify various risks, calculate the expected risk value for the organisation, and compare it to the actual risk value. It provides a more intuitive measure of the validity of metric selection and baseline setting. Furthermore, through risk mitigation value and subsequent calculation of ROSI, the use of risk value and ROSI for guiding investment decision-making is explored. This study offers a method for risk estimation and quantitative analysis of return on investment, serving as a more scientific approach to supporting investment decision-making. Of course, this study has several limitations, and in the future, efforts will focus on applying the framework to actual cases and continually optimizing the metrics and models. The aim is to provide a comprehensive, objective, and current cybersecurity decision support guide.

## References

- Allodi, L., & Massacci, F. (2017). Security Events and Vulnerability Data for Cybersecurity Risk Estimation. *Risk Analysis*, 37(8), 1606–1627. <https://doi.org/10.1111/risa.12864>
- Armenia, S., Angelini, M., Nonino, F., Palombi, G., & Schlitzer, M. F. (2021). A dynamic simulation approach to support the evaluation of cyber risks and security investments in SMEs. *Decision Support Systems*, 147(June 2020), 113580. <https://doi.org/10.1016/j.dss.2021.113580>
- Catota, F. E., Granger Morgan, M., & Sicker, D. C. (2018). Cybersecurity incident response capabilities in the Ecuadorian financial sector. *Journal of Cybersecurity*, 4(1), 1–20. <https://doi.org/10.1093/cybsec/tyy002>
- CRISTEA, L. M. (2021). Risks Associated with Threats Related to Disruptive Technologies in the Current Financial Systems Context. *Audit Financiar*, 19(161), 119–129. <https://doi.org/10.20869/auditf/2021/161/002>
- Fedele, A., & Roner, C. (2022). Dangerous games: A literature review on cybersecurity investments. *Journal of Economic Surveys*, 36(1), 157–187. <https://doi.org/10.1111/joes.12456>
- Fielder, A., Panaousis, E., Malacaria, P., Hankin, C., & Smeraldi, F. (2016). Decision support approaches for cyber security investment. *Decision Support Systems*, 86, 13–23. <https://doi.org/10.1016/j.dss.2016.02.012>
- Georgiadou, A., Mouzakitis, S., & Askounis, D. (2022). Detecting Insider Threat via a Cyber-Security Culture Framework. *Journal of Computer Information Systems*, 62(4), 706–716. <https://doi.org/10.1080/08874417.2021.1903367>
- Ghani, H., Luna, J., & Suri, N. (2013). Quantitative assessment of software vulnerabilities based on economic-driven security metrics. *2013 International Conference on Risks and Security of Internet and Systems, CRiSIS 2013*, 1–8. <https://doi.org/10.1109/CRiSIS.2013.6766361>

- Gordon, L. A., & Loeb, M. P. (2002). The Economics of Information Security Investment. *ACM Transactions on Information and System Security*, 5(4), 438–457. <https://doi.org/10.1145/581271.581274>
- Gordon, L. A., Loeb, M. P., Lucyshyn, W., & Zhou, L. (2018). Empirical Evidence on the Determinants of Cybersecurity Investments in Private Sector Firms. *Journal of Information Security*, 09(02), 133–153. <https://doi.org/10.4236/jis.2018.92010>
- He, Y., Zamani, E. D., Lloyd, S., & Luo, C. (2022). Agile incident response (AIR): Improving the incident response process in healthcare. *International Journal of Information Management*, 62(October 2021), 102435. <https://doi.org/10.1016/j.ijinfomgt.2021.102435>
- He, Y., Zamani, E., Yevseyeva, I., & Luo, C. (2023). Artificial Intelligence–Based Ethical Hacking for Health Information Systems: Simulation Study. *Journal of Medical Internet Research*, 25, 1–16. <https://doi.org/10.2196/41748>
- Houmb, S. H., Franqueira, V. N. L., & Engum, E. A. (2010). Quantifying security risk level from CVSS estimates of frequency and impact. *Journal of Systems and Software*, 83(9), 1622–1634. <https://doi.org/10.1016/j.jss.2009.08.023>
- IBM Security. (2023). *Cost of a Data Breach Report 2023*. [https://doi.org/10.1016/s1361-3723\(21\)00082-8](https://doi.org/10.1016/s1361-3723(21)00082-8)
- Janicke, H., Smith, R., Maglaras, L., Cook, A., & He, Y. (2021). Agile Incident Response in Industrial Control Environments. *Cybersecurity Issues in Emerging Technologies*, CRC press, 43–58.
- Kesswani, N., & Kumar, S. (2015). Maintaining cyber security: Implications, cost and returns. *SIGMIS-CPR 2015 - Proceedings of the 2015 ACM SIGMIS Conference on Computers and People Research*, 161–164. <https://doi.org/10.1145/2751957.2751976>
- Kim, D. (2019). Decision-making method for estimating malware risk index. *Applied Sciences (Switzerland)*, 9(22). <https://doi.org/10.3390/APP9224943>
- Lee, I. (2021). Cybersecurity: Risk management framework and investment cost analysis. *Business Horizons*, 64(5), 659–671. <https://doi.org/10.1016/j.bushor.2021.02.022>
- Loft, P., He, Y., Janicke, H., & Wagner, I. (2021). Dying of a hundred good symptoms: why good security can still fail - a literature review and analysis. *Enterprise Information Systems*, 15(4), 448–473. <https://doi.org/10.1080/17517575.2019.1605000>
- Loft, P., He, Y., Yevseyeva, I., & Wagner, I. (2022). CAESAR8: An agile enterprise architecture approach to managing information security risks. *Computers and Security*, 122, 102877. <https://doi.org/10.1016/j.cose.2022.102877>
- Ma, W. (2021). Research on network vulnerability assessment based on attack graph and security metrics. *Journal of Physics: Conference Series*, 1774(1). <https://doi.org/10.1088/1742-6596/1774/1/012070>
- Naseer, H., Maynard, S. B., & Desouza, K. C. (2021). Demystifying analytical information processing capability: The case of cybersecurity incident response. *Decision Support Systems*, 143(December 2020), 113476. <https://doi.org/10.1016/j.dss.2020.113476>
- Qamar, S., Anwar, Z., Rahman, M. A., Al-Shaer, E., & Chu, B. T. (2017). Data-driven analytics for cyber-threat intelligence and information sharing. *Computers & Security*, 67, 35-58.
- Smith, R., Janicke, H., He, Y., Ferra, F., & Albakri, A. (2021). The Agile Incident Response for Industrial Control Systems (AIR4ICS) framework. *Computers and Security*, 109, 102398. <https://doi.org/10.1016/j.cose.2021.102398>

- Wang, J., Neil, M., & Fenton, N. (2020). A Bayesian network approach for cybersecurity risk assessment implementing and extending the FAIR model. *Computers and Security*, 89, 101659. <https://doi.org/10.1016/j.cose.2019.101659>
- Zadeh, Jeyaraj, & Biros. (2020). Characterizing Cybersecurity Threats to Organizations in Support of Risk Mitigation Decisions. *E-Service Journal*, 12(2), 1. <https://doi.org/10.2979/eservicej.12.2.01>
- Zamani, E., He, Y., & Phillips, M. (2020). On the Security Risks of the Blockchain. *Journal of Computer Information Systems*, 60(6), 495–506. <https://doi.org/10.1080/08874417.2018.1538709>

# Digital inclusion network building: a network weaving analysis

*Sharon Wagg*

*University of Sheffield*

*Sara Vannini*

*University of Sheffield*

*Efpraxia D. Zamani*

*Durham University*

*Maira Klyshbekova*

*University of Sheffield*

*Bethany Aylward*

*University of Sheffield*

*(Completed research)*

## **Abstract**

*This paper aims to investigate digital inclusion network building as a mechanism for reducing digital poverty. Analysing a rural digital inclusion network in the UK, and drawing on Network Weaving Theory (Holley, 2013), this case study analyses how people's roles and places play a big part in both the construction and growth of the network, as well as in the advancements of its initiatives. The contribution of this study is significant. Theoretically, this research builds on the literature on contextual conditions to digital access and adoption and proposes a novel theoretical framework to unpack the complexity of digital inclusion network building - based on Holley's theory. Results will inform UK regional organisations' practices for establishing effective digital inclusion networks in post-pandemic, 'digital by default', and cost of living crisis times in the country. It will also provide recommendations for national policies to strengthen the resilience and sustainability of digital inclusion provision.*

**Keywords:** digital inclusion, network building, network weaving, digital poverty

## **1.0 Introduction**

Efforts to reduce digital poverty through digital inclusion have long preoccupied scholars and policymakers, but the task of implementing workable digital inclusion initiatives is complex due to a multitude of factors which contribute to digital exclusion (Wagg and Simeonova, 2022). Arguably this complexity has been exacerbated as a consequence of the Covid-19 pandemic and the consequent dramatic movement to 'online only' products and services in the UK. This almost overnight movement further exposed digital poverty inequities and the vulnerabilities created through the 'digital by default' culture. We now have the opportunity to revisit how

digital inclusion is being delivered in the community with a transformative perspective for social change.

During the pandemic, a myriad of emergency digital inclusion initiatives were spurred into action in the country, some with more positive results than others. Communities and organisations came together as they sought innovative solutions to reduce digital poverty (Holmes and Burgess, 2022), highlighting also how much communities were struggling in this sense. Specific solutions gaining attention - and the focus of this paper - include the recent growth of digital inclusion network building. Here, we are not talking about technological digital infrastructure networks such as fibre, broadband or mobile. Our focus is on networks of organisations that provide social and community support through digital inclusion activities. For many of these organisations, digital inclusion is one of many activities they perform.

Network building has grabbed the attention of funding bodies and civil societies, who recognise the need to find new ways to look at problems which have not been solved via solutions such as better infrastructure or more skills-based training. These organisations are now sponsoring the creation of mechanisms to support networks of local government authorities, community partners and organisations from the voluntary, community and social enterprise (VCSE) sector that are addressing digital inequality (The British Academy, 2023).

Indeed, despite the UK being widely considered as a well-connected country with a robust digital economy, several intractable digital divides have persisted over many years that exclude a proportion of society. Scholars argue such digital divides have been exacerbated by the drive to digitise government services and press forward with a 'digital-by-default' agenda and the closure of face-to-face services (Holmes and Burgess, 2022). As a result, the last decade has witnessed a shift from digital being a 'nice to have' to a 'necessity' as more and more products and services are now only accessible online. This shift some would argue has led to the creation of an era of 'compulsory computing' (Allmann and Blank, 2021) or 'digital enforcement' (Díaz Andrade and Techatassanasoontorn, 2021), where everyone is expected to not only have access to digital technologies, but also have the confidence and know-how of

using ICTs for delicate needs such as applying for welfare benefits, online banking, or discussing health-related issues.

Those particularly affected are often people experiencing different levels of marginalisation, as they are more likely to experience ‘digital poverty’ and interact with and use the online world fully (Allmann, 2022). Scholars argue that digital inclusion policies continue to struggle to address significant inequality issues (Diaz Andrade and Techatassanasoontorn, 2021). They also highlight a scarcity of literature exploring digital inclusion through a theoretical lens (Wagg and Simeonova, 2022; Al-Muwil et al., 2019).

In this paper, we argue the use of theory is fruitful in guiding digital inclusion research and recommend the use of Network Weaving (Holley, 2013) as a theoretical framework to understand the complexity of digital inclusion network building and to explore its potential to complement and advance existing understanding of digital inclusion practices. We aim to gain a better understanding of digital inclusion network building as a mechanism for reducing digital poverty, and posit the following research questions: How do people’s roles and places play a big part in the construction and growth of a digital inclusion network? How does this reflect in the advancements of its initiatives? How can the Network Weaving theoretical framework help us unpack the complexity of digital inclusion network building?

The contributions of this paper include: i) new insights into the understanding of digital inclusion network building; ii) builds on the literature on contextual conditions to digital access and adoption; and iii) proposes a theoretical framework to unpack the complexity of digital inclusion network building using Network Weaving

## **2.0 Literature review**

### **2.1 The issue with ‘digital-by-default’ policies**

Existing studies on digital inclusion/exclusion have highlighted the need to better understand whether and to what extent local networks and place-based approaches can address inequities in terms of digitalising core aspects of our everyday lives (e.g.,

Smith et al. 2010; Park et al, 2019). Policies and agendas that have been prioritising ‘digital-by-default’ (often also referred to as ‘digital first’), such as the ones recently adopted in the UK, have resulted in more and more products and services being delivered primarily and often exclusively in digital forms. Often, such policies are framed as being progressive, more efficient, and more productive (Schou and Hjelholt, 2018).

However, experience shows that, despite digital technologies becoming more ubiquitous, their access and usability remain unequal and problematic. The root cause for this most of the time is a combination of other inequalities that compound each other, further exacerbating the negative impacts of accelerated digitalisation among the most vulnerable. For example, while there are numerous affordable devices (e.g., low-cost smartphones), the increasing cost of living poses significant challenges as to who can go and remain online (Nathanial-Ayodele and McGrath, 2023). Equally, however, even for those households that may be financially better off, other factors might come into play, including living in less-resourced areas. Indeed, geography plays a crucial role as where one lives influences whether one can connect via broadband or mobile connectivity. In more detail, rural and remote, hard-to-reach areas are often characterised by little to no broadband infrastructure (Valentin-Sivico et al, 2022; Philip and Williams, 2019; Wagg and Simeonova, 2022).

Further, the COVID-19 pandemic and its implications have deepened and accentuated structural inequalities. During and after the pandemic, digitally-enabled solutions to deliver services to confront and overcome the pandemic have accelerated, e.g., the NHS COVID-19 contact tracing app (Eom and Lee, 2022) and similar solutions. Yet, despite that such initiatives allowed for business continuity (Zamani et al., 2022), their wide adoption exacerbated digital inequalities, and further drew attention to the fact that a lot more people and communities can be digitally excluded in more complex ways than what conceptualised by previous simplistic notions and divisions of ‘haves’ and have nots’ (Zheng and Walsham, 2021).

## **2.2 Digital Poverty**

The complexity of digital poverty is influenced by a multitude of factors. Zamani and Vannini (2022) suggest that digital poverty arises from a combination of deprivations

in areas such as education, employment, and health, as well as disparities in access to broadband infrastructure and various demographic characteristics. This complexity extends beyond simple binary divisions, like gender or age, and instead encompasses a spectrum of characteristics.

These intricate factors, combined with the tendency for digital inclusion projects to receive short-term funding, present a formidable challenge for policymakers when it comes to implementing effective digital inclusion initiatives (Bach et al., 2013). Short-term projects often falter because they fail to empower the community and don't allow for the cultivation of a necessary digital culture within the community, let alone considering the scalability of such initiatives (Hemerling et al., 2018).

The persistence of digital poverty in the UK has led to a recent policy focus through evidence-based reviews by the Digital Poverty Alliance (Allmann, 2022) and The British Academy (2022). These reviews, and other research during and post the pandemic, brought to light the socio-economic implications of digital poverty (Deloitte, 2023) and the increasing issue of 'data poverty' defined by Lucas et al. (2020) as "those individuals, households or communities who cannot afford sufficient, private and secure mobile or broadband data to meet their essential needs".

Other attention has been drawn to the need to move away from existing policy measures of digital exclusion based on access to or use of internet services to establish a benchmark for a minimum digital standard of living (MDSL) for households that complements existing indices but also captures the real issues faced by society. This led to the development of the MDSL which "includes, but is more than, having accessible internet, adequate equipment, and the skills, knowledge and support people need. It is about being able to communicate, connect and engage with opportunities safely and with confidence" (Blackwell et al, 2023, p. 3).

### **2.3 Digital Inclusion Approaches**

Digital inclusion initiatives historically have been dominated by the installation and provision of digital infrastructure and devices. However, approaches to digital inclusion also include a strong social aspect where efforts involve providing digital skills training and social support (Asmar et al., 2020). These social initiatives are carried out by a range of organisations, including public libraries, advisory centres,



service providers, adult education institutions, housing associations, learning centres, and, at times, financial institutions, and telecommunications companies (Al-Muwil et al., 2019; Reisdorf and Rhinesmith, 2020). Digital inclusion initiatives have often relied upon what could be termed as community infrastructure organisations (Cook et al., 2023) or intermediary organisations (Wagg, 2021) reaching out, engaging and supporting digitally excluded individuals (Torrecillas et al., 2014). How such initiatives are delivered varies, from targeted approaches, targeting specific groups of individuals, to more community asset-based (or strength-based) approaches (Reisdorf and Rhinesmith, 2018). Scholars highlight how there has been an overreliance on such organisations which are often poorly funded and under-resourced (Wagg and Simeonova, 2022), following years of disinvestment and disenchantment in policy and practice, creating a fragmented landscape of provision of civil society infrastructure (Macmillan, 2021). But it is the hyperlocality and trusted nature that make such organisations crucial for digital inclusion initiatives.

Indeed, policymakers worldwide have recognized the importance of digital inclusion (Diaz Andrade and Techatassanasoontorn, 2021; Faith, Hernandez, and Beecher, 2022). In the UK, this commitment is evident through the government's Digital Inclusion strategy (Cabinet Office, 2014) and the proliferation of digital inclusion initiatives (Mervyn et al., 2014, Wagg, 2021). The number of initiatives surged during the pandemic to help tackle digital poverty. As well as focusing on access to technology and digital skills, these initiatives involved device gifting and donations, temporary removal of caps on broadband use, agreed between the Government and some internet providers, to allow individuals who could not afford to buy more data or wi-fi provision to continue to use the internet (Holmes and Burgess, 2022).

However, it's important to note that not all digital inclusion programmes have been successful (Madon et al., 2009; Helsper and Reisdorf, 2017; Davies et al., 2017). Despite calls for deeper insights into digital inclusion, existing literature primarily focuses on the perspective of a 'digital skills deficit,' with limited attention given to the contextual factors that influence it (Vannini et al., 2017; Lythreathis et al., 2022). Notable exceptions which take a more place-based perspective include Park et al., 2019; Reisdorf and Rhinesmith, 2020; and Guenther et al., 2020.

The 'deficit' perspective is also dominant in current policy initiatives, such as the 2022 UK Digital Strategy, which regards digital skills as central to the nation's growth, competitiveness, and long-term prosperity (UK Digital Strategy, 2022). Considering however what we know about digital poverty, its underlying causes, and the far-reaching consequences it carries, the 'deficit' approach in both research and policy, raises questions about the effectiveness of current discussions on digital inclusion and the necessary changes to effectively support communities toward reducing inequalities.

#### **2.4 Digital inclusion networks**

Digital inclusion networks are not new. National networks in existence include the National Digital Inclusion Network (formerly the UK Online Centres Network) and the Digital Inclusion Network (operated by the National Digital Inclusion Alliance in the US). However, a recent emerging trend that seems to be driving new approaches to digital inclusion involves network building and mapping (Wagg, 2021; Mason et al., 2022).

Furthermore, as part of the broader effort to comprehend digital poverty, there is a growing inclination towards regional mapping exercises. These exercises aim to pinpoint areas where digital inclusion initiatives are needed, particularly in regions marked by multiple deprivation indicators, inadequate digital connectivity, and gaps in training provision. Projects that have undertaken such activities include the LOTI Digital Inclusion Innovation Programme, (LOTI, 2022) and the Greater Manchester Digital Skills map (GMCA, 2023). Similarly, the University of Sheffield carried out a mapping exercise to understand digital poverty in the South Yorkshire region (Zamani & Vannini, 2022).

It's worth noting that the utility of such mapping exercises is limited in the short term because the circumstances of communities and individuals can and often do change. Local digital inclusion provision may not always be accurately reflected, and most mapping efforts rely on datasets that capture conditions at a specific moment in time. Nevertheless, these exercises can be valuable tools for identifying areas with greater need and, consequently, for developing place-based interventions and supporting new approaches to digital inclusion provision.

Despite this drive towards digital inclusion network building and mapping and calls from scholars highlighting this need, digital inclusion networks remain an under-researched area (Wagg et al., 2024). This therefore provides an opportunity to problematise and challenge taken-for-granted assumptions about existing digital inclusion provisions and reveal the complexity of building digital inclusion networks and principles that can be applied to other local contexts.

### **3.0 Network Weaving**

Recognising the importance of theory to understand research, as well as the need for more theory in digital inclusion scholarship (Gomez & Pather, 2012; Gomez, 2013), we adopt Network Weaving (NW) as a theoretical framework to underline not only the importance of on the ground networks to address communities' digital inclusion needs, but also to understand how digital networks should be set up and work to meet their communities' needs.

As a theory, NW was developed by June Holley (2013) to assist low-income entrepreneurs in one of the most economically disadvantaged areas in the United States, and with the primary aim to facilitate the establishment of networks that would catalyse transformative change within these underserved communities.

Holley (2013) emphasises how a focus on networks enables societal transformations via rectifying disparities in power dynamics among people, communities, and institutions. Networks, in fact, function as the foundational support structure of social systems, ensuring their stability. Therefore, exposing and understanding the networks that maintain the status quo is essential to disrupting it and achieving systemic change. This focus enables people and communities to deliberately nurture new networks, focusing on people experiencing patterns of marginalisation, and dispersing power. According to NW, facilitating transformational change involves repositioning oneself within networks and nurturing grassroots ones.

Holley (ibid.) proposes that NW can help these transformations by deliberately fostering relationships among peers who acknowledge and appreciate the respective contributions, and by regarding each individual as a potential leader, capable of forging connections and instigating collaborative endeavours. The consequent

redistribution of power among the members of the network, as well as the involvement of all stakeholders' perspectives, are emphasised.

Holley theorises distinct and coexisting types of networks and centres her focus on their roles, as well as on the roles of the people involved. Network Weavers can assume four different roles, prioritise the development of the network (Connector Catalyst and Guardian), or take action (Self-organised Project Coordinator and Facilitator). Additionally, primarily functioning at the micro- (Connector Catalyst and Project Coordinator), or at the macro-level (Network Guardian and Facilitator). Their responsibilities encompass enhancing the overall systems of networks or aiding individuals in recognising the advantages associated with a network-oriented approach. What's common is how, in NW, leadership is characterised by its distributed nature, wherein every individual has the potential to assume a leadership role and leadership skills do not need to be innate but can be developed. The innate capacity of networks to self-organise is seen as a catalyst for proactive engagement that fosters a sense of ownership and active participation. Holley's emphasis here is also on the support needed for network leaders to succeed (ibid). Less attention is posed in the theory to the role that places and spaces can play in NW.

We recognise that the lens of NW, although little used in Information Systems, can lead to fruitful conversations in the field (see also: Marais & Vannini, 2021). Analysing digital inclusion networks through NW introduces a novel perspective to research in the field, and it helps us place a strong emphasis on the dynamic roles of individuals, recognising their distributed nature, and appreciating the unique strengths that each person brings to the table. This perspective underscores the importance of committing to support as a cornerstone for both individual and network growth, fostering a culture of "supporting the supporters". Furthermore, through NW, the significance of relationships and the necessity of nurturing and caring for one another within the network is acknowledged, emphasising community-building and shifting the conventional focus of digital inclusion from a mere infrastructure and skills problem to a holistic, community-centred place-based approach. Finally, NW explores the concept of redistributing access to power through people. It advocates for the necessity of contextual and co-created solutions, recognizing that change doesn't always equate to uniform accessibility. In doing so, it paves the way for a transformative approach to network dynamics that embraces diversity and inclusivity

as its guiding principles. Elsewhere, NW has indeed been suggested as a transformational, feminist approach to social change (Marais & Vannini, 2021).

## **4.0 Methodology**

This study adopted a case study approach guided by the following research questions:

1. How do people's roles and places play a big part in the construction and growth of a digital inclusion network?
2. How does this reflect in the advancements of its initiatives?
3. How can the Network Weaving theoretical framework help us unpack the complexity of digital inclusion network building?

This case study was selected specifically for its uniqueness rather than its potential for generalisability. As per case study methodology, individual cases can convey principles that, although not universally applicable, can be learned from, adapted, and applied in diverse contexts (Yin, 2009). This specific case, centred on a local digital inclusion network established to address digital poverty in a rural area of the UK, was therefore chosen due to its potential in offering valuable lessons on the issue of digital inclusion.

To refer to the network we analyse, we will use the pseudonym Dedicated Digital Inclusion Network (DDIN) and Rural Business & Community Foundation (RBCF) as the organisation that set it up. We conducted semi-structured interviews with DDIN network participants at multiple levels, using a combination of snowball sampling and purposeful sampling. We proceeded by advertising our study with the network within meetings and using the mailing list set up by the main network convenors. Interviews were conducted with individuals who volunteered initially, and additional participants were identified through recommendations from the initial interviewees. Subsequently, we monitored the number of interviews conducted with stakeholders at various network levels (as outlined below). To ensure comprehensive coverage, we sent targeted communication to organisations at the remaining levels requiring interviews, with continued coordination through the primary network convenor. Our aim was not to gather a similar amount of interviews per network level. Rather, we aimed to ensure the representation of as many different voices within the network. This is why

interviewees at Level 3 constitute the highest number in the corpus. In total, we gathered 21 interviews with:

- Level 1 (L1) Key stakeholders managing the development of the DDIN network (2 interviewees);
- Level 2 (L2) Key stakeholders involved in the mapping exercise of digital poverty and digital inclusion activities in Derbyshire (1 interviewee);
- Level 3 (L3) Organisations who have joined the network and /or collaborate with the network (15 interviewees);
- Level 4 (L4) Individuals from organisations that have been involved in digital inclusion networks and mapping digital poverty/ inclusion activities in other locations in the UK (3 interviewees).

We also conducted a total of five non-intrusive observations (four were in-person, and one online): we did not participate nor interrupt the participants and activities being held during the observations (Creswell, 2014). We followed a semi-structured observational protocol, making written fieldnotes and including criteria such as: local infrastructure, layout of the place, description of activities, capturing the essence of the workplace, description of digital inclusion activities, and description of volunteers' or workers' duties. We also took reflexive notes after each observation. The activities observed differed: one involved a workplace, three involved digital training sessions, and one was an online steering group meeting.

The two data sets from semi-structured interviews and observations were coupled together in the analysis step for this paper. We used Dedoose for thematic analysis purposes (Braun & Clarke, 2006). We applied abductive coding, following a predefined codebook as well as creating new codes starting from the data. The finalised codes were then organised into themes, through which the results were described. In this paper, we present findings from the two themes 'roles' and 'place'.

## **5.0 Findings**

### **5.1 The importance of roles**

One of the core dimensions that emerged during data collection and analysis, was that of 'roles'. This is about the roles performed by actors involved in the network; the

motivations and enablers of those actors performing those roles, the barriers faced by those actors; and the emotions of those actors undertaking such roles. The participants of this study were involved in a wide range of organisations in which they occupied an equally wide range of roles, and these correspond with the levels of stakeholders outlined earlier in this paper.

The existence of the DDIN was enabled by funding to specifically tackle digital inequality in the county (Interviewee 3, L1) and for those weaving the network, the witnessing of connections being made, best practices being shared, and provision being improved is a key motivator. This feeds into the role of positive reinforcement in maintaining momentum within the DDIN; the sharing of success stories between network members is not only valuable in terms of skills and knowledge sharing but reminds members of the impact their work has on people's lives:

*The benefits of a person who works as a facilitator of a network where you start to see how the sharing of information and the getting that sense of the bigger picture is really, really valuable. (Interviewee 3, L1)*

*So, there's one that's really, I think, quite significant. It was a gentleman, only an individual, but somebody who was profoundly deaf, and the service that's providing him with support has been able to make such a difference for him. (Interviewee 2, L1)*

The importance of this positive feedback loop came across strongly in both interviews and observation. The quote above refers to a story shared by a network member in a DDIN meeting observed by the research team, and the quote below captures the joy and encouragement that connections between organisations foster:

*It is that knock-on effect that is so inspiring, not just for me but for the people who have done it. And I think the news that I had last week from Community Charity that they had helped 405 individual people to understand computers – that is a fantastic achievement. (Interviewee 8, L3)*

Participants delivering digital inclusion activities spoke of their desire to help people, with many pinpointing the fear of technology as a significant obstacle to minimising digital inequality. In one case a participant referred to their own experiences of

witnessing the impact digital exclusion can have in a society where 'digital-by-default' makes it increasingly difficult to maintain a good quality of life, coupled with their confidence in using ICTs as a key motivator for volunteering:

*I've got some spare time at the moment, I've got a lot of experience of computers, phones, etc., so it's really an opportunity to give something back. And also, I see from relatives the exclusion that people are experiencing. Often, they're being forced into it because of modern life. (Interviewee 9, L3)*

It was also suggested that the individual character of those involved in delivering digital inclusion activities is an important consideration. There has to be a level of compassion towards users who experience technophobia and an understanding that skills that are perhaps intuitive to some are a steep learning curve for others:

*I wouldn't like to think people were frightened of something which is relatively easy to understand. I don't mind. I spent 6 hours trying to teach a lady how to use a mouse, so I've got the patience. (Interviewee 14, L3)*

The time volunteers are willing to devote to the users of their digital inclusion activities is mirrored by the time DDIN members are willing to devote to feeding back to the network and supporting the work of others:

*What struck me is the amount of goodwill there is amongst projects and organisations to communicate with each other and help each other out when they can. For instance, the surveys that we've been doing with our network, it's totally on their goodwill; we're not providing them with funding, or anything. We've had some really comprehensive answers and people have taken a lot of time, and some of the case studies that they've provided have given us such a personal insight into the difference that it makes. (Interviewee 2, L1)*

Just as user participation in digital inclusion activities was found to have social drivers, this research also found that those delivering services were motivated by the relationships that are built through providing support to others:

*I mean, it's another reason I like doing it, actually, because you get to chat to some of the older residents of the Town and find out what the town was like*



*and what the railway was like when it was running, and all sorts of things like that.* (Interviewee 9, L3)

Considering the high proportion of volunteers involved in the provision of digital inclusion activities it is unsurprising that one of the key challenges is time:

*They're often very busy. And with paid staff and with volunteers, they're often doing multiple things and multitasking, and so time is very valuable.*  
(Interviewee 2, L1)

Linked with the barrier of time is the need to prioritise the distribution of resources to the most impactful activities, and this comes at a cost. One participant volunteering in the delivery of digital inclusion activities acknowledged that services need to be publicised better to increase their impact, but their organisation does not have the capacity to dedicate time to outreach:

*It is more a matter of getting more publicity. I'm not at the moment – and I don't think Bill or anybody else is either – in a position to do an awful lot.*  
(Interviewee 12, L3)

Directing an organisation's resources to the running of digital inclusion sessions only makes sense when those sessions are well-publicised and therefore well attended. Multiple participants, especially those who volunteer their time to deliver the sessions, acknowledged a general feeling of frustration with the poor attendance:

*If we go and there is nobody there needing help you feel what am I doing here, it is a bit of a waste of time... If you haven't got anybody to deal with, it gets a bit boring sometimes.* (Interviewee 10, L3)

Priorities can also be dictated by the terms stated by funding bodies, and it is in the organisation's best interest to adhere to the stipulations of the funding contract if they are to be successful in future bids:

*People are time poor and resource poor, and they're always having to prioritise whatever they're being funded to deliver, whoever they've got a contract with or a grant with, they will prioritise delivering against those contracts.* (Interviewee 25, L4)

Whilst the scale of the DDIN and the enthusiasm of its members to share best practices is a triumph, an L1 participant involved in the development and coordination of DDIN found this role overwhelming at times:

*It's been exhausting at times. It has been challenging because... it's taken me right out of my comfort zone at times because there's so many different themes where digital overlaps and they're not necessarily where you have any expertise in or any standing. Whereas everyone seems to be an expert in their own particular field so it's quite hard. (Interviewee 3, L1)*

In addition to feeling overstretched and sometimes underqualified, multiple participants expressed their frustration that the efforts of voluntary organisations and charities within the network are overlooked by local governments. The DDIN was established to fill gaps in digital inclusion provision across the county, a service which some participants argue should be within the remit of the local government. For them, digital inclusion would be a statutory service, eliminating the need for a network and the organisations and charities it contains, the fact that this is not the reality has led Interviewee 11 to believe that the importance of this work is not recognised:

*I would like to see a bit more acknowledgement of what they have done from the district council and certainly the county council because when we first set this up I called a meeting... and we had a senior county councillor come and there was a retort if you like, 'Stop trying to reinvent the wheel'. And what I would like to say to him now, 'I am not reinventing the wheel, I have added a few more spokes'. (Interviewee 8, L3)*

*This is something that should sit with a statutory service really, something like a digital inclusion officer at the council. That should be where something like this sits, eventually. Because it does need to be considered a statutory thing, I think. It needs to be considered that important and I think at the moment it is not considered that important. (Interviewee 11, L3)*

Fortunately, there is no shortage of recognition within the network itself, the coordinators of the DDIN were forthcoming in their gratitude towards their members and praise of the vital work they do:

*We're just here to facilitate it. They're the people that are making it happen and making it work, really. They're the important ones.* (Interviewee 2, L1)

L1 participants downplayed their value as 'guardians' and 'facilitators' of the DDIN, building and maintaining the network at a macro level (Marais & Vannini, 2021). However, in line with network weaving theory (Holley, 2013; Marais & Vannini, 2021), Interviewee 8 - who acts as a connector catalyst, forging links and promoting participation in a hyperlocal context - was clear that it is the interplay of different but equally important roles within the network that are key to its positive impact in the area:

*All I am doing is joining the dots, but I think joining the dots is as important as being part of a service.* (Interviewee 8, L3)

## **5.2 The importance of place**

A second core dimension that emerged during data was that of 'place'. Place in terms of rurality, the venues and locations where members of the network delivered their digital inclusion activities, and in how the digital inclusion network convened and communicated.

Our findings revealed that throughout our time engaging with the digital inclusion network, activities were being undertaken to grow and promote the network and cultivate relationships between the organisations. This was done by L1 using regular email communication and newsletters, online events, and the encouragement of knowledge sharing between organisations within and beyond the network, thus encouraging 3-way communication. L1 also collaborated with L2 to survey members of the digital inclusion network and completed a mapping exercise of digital poverty and digital inclusion activities within Derbyshire. Our qualitative insights enable a deeper understanding of the network.

The locations where organisations and volunteers set up their digital inclusion initiatives are fairly diverse, as are the range of services and activities related to digital inclusion they offer. This is reflected also in their being present and occupying spaces, and in the design of how these spaces would be and operate - or the inability of

organisations to have a say on this design. While some organisations operate within dedicated venues, specifically set up for digital inclusion activities, others have to use multipurpose rooms within other institutions, which are, at times, less ideal - but they make do. And finally, others adapt spaces they already use to offer other services, offering digital inclusion as an extension to these, rather than a distinct, dedicated space.

This is partly reflected also in their long-term vision for the offered services, and in the choice of technologies that are out at public disposal. For example, the people managing the network at L1 level mentioned more than once the example of food pantries, which have seen the necessity of providing their community with some digital access services in the last few years:

*So you might have one that has people coming in to use the pantry and they've got a little laptop at the side because they don't have much space, and they just help them in the corner, to another one that's got an awful lot of space and it's maybe a little bit more formal in terms of the help that they provide [...] We do have a digital pantry in [an urban town] and they have no internet, so they're running it all with dongles. They're very resourceful. (Interviewee 2, L1)*

In our observations, this heterogeneity of places and spaces became apparent. We visited training sessions held within the premises of the local village hall, which also had a café providing high-quality coffee and freshly prepared meals. Both trainers and attendees usually meet before the training session to have coffee or tea and have a chat. Another organisation regularly hosts a digital café within the premises of a modern recreational building which was built with funding raised from the organisation itself during the Covid-19 pandemic. The facilities are also used by the community for sports and other community activities. The atmosphere is vibrant and the space is humming with activity. Upon entering, we were greeted with a spacious room with floor-to-ceiling windows on one wall, views of hills and playing fields, and French doors that lead out onto a big patio:

*There is a group doing yoga on a Monday morning [...] And then Tuesday, we've got a number of district council health and wellbeing sessions. And also*

*the facility is the start and end point for a community walk. So they'll walk around the grounds and through the park and back to the Recreation Building and then have tea, coffee, etc., [...]. Wednesday, that's another fairly full day because we've got more health and wellbeing sessions and Tai Chi. And then Thursday, Digital Café, which is a community helpline for anyone who has computer or mobile phone problems. And then Friday morning is a parent and toddlers group. (Interviewee 6, L3)*

Irrespective of this heterogeneity, the place and space where organisations gather, as well as the way organisations set them up to meet communities' needs, seem to play an integral part in fostering community, providing educational experiences, and advancing digital inclusion. So much that barriers to accessing and organising suitable spaces for their activities are often felt as a barrier to participation from the community and the effectiveness of their digital inclusion efforts:

*I prefer the library myself. When I first started they got a side room where there were refreshments, 20p [...] So, about halfway through, everybody get together and get away from the screens and have a good old chat and social. [...] It was nice but since this pandemic, the room has been shut off and now there's a fee of £14 to hire it. You've got to hire it, so I think that's the reason [we do not meet there anymore], I'm not sure, but it was really popular. People used to come with loads and loads of people, used to do what they'd got to do on the computer and then go for the [social] and then go back to the computer afterwards, yes. (Interviewee 1, L3)*

Barriers to the establishment of the desired environment and infrastructure are usually due to financial constraints:

*We are limited by the space we have got and everything. We are lucky in that the council let us use that for free, the library and where you went, the Community House. [...] It is free, I mean that is the thing, the whole thing is free to people because we don't get charged for the space. (Interviewee 10, L3)*

Interviewees also insist on the importance of socialisation as a fundamental component of their activities. The opportunity of meeting and interacting with other people is a motivator for people to join and participate in the sessions, and a way for them not to feel alone and different as they need help to figure out how to access the digital environment. Organisations feel that fostering community and socialisation is another way they can contribute not only to more efficient digital inclusion services but also more vastly to their communities' well-being and development.

*For the people that we supported, it was as much about the face-to-face contact and seeing someone as it was about what they were actually learning [...] And the main thing, the main motivation behind those has been about connections and helping people keep in touch with each other and have fun together of one thing or another. (Interviewee 13, L3)*

*I think one of the things going back to the Digital Café, yes, they get a cup of tea and a biscuit, and yes they have the technology. But when they are doing either they are able to talk socially. [...] People have got the ability to talk about what their concerns are and how they can get the help, where they can get it and how much they need that help [...] I know some people had help with getting access to help for their own needs in their own home in their own time, without it being seen as an embarrassment, a dignity factor that people in older age don't want to apply for things. They think they are seen as scroungers and they are not. And that is the sort of thing, that is the impact that it is having on this social interaction and I am quite genuinely sure that some people just come for the cup of tea and a chat, as well as a little bit of information on computer work. It is that connectivity that has created such an almost bold, balanced organisation, a well-balanced input into the community. It is no push and shove – it is easy, easy and very, very well-received. (Interviewee 8, L3)*

L1 organisations also emphasise the significance of this element of socialisation for the volunteers themselves, who are finding in the network a space to support each other across organisations:

*People can advise each other on the best way that they've found to recruit volunteers or ... Something that we've had feed back on is that people have found the [online] sessions are more likely to be attended if they're put across as being quite informal, not like teachy, teachy. And almost it's the social aspect that's emphasised. And we've been able to advise another project, who has set up sessions but isn't having the footfall that they wanted, that this is what other projects have found, that if you play it down almost but emphasise the social side and make it sound quite informal, people are more likely to come to you. So it's shared experiences, I guess, a lot of the time. (Interviewee 2, L1)*

This is clearly an example of network weaving when L1 organisations operate as conduits putting organisations in touch with one another, e.g when one organisation needs to find a supplier of second-hand, refurbished digital devices to support one of their clients and L1 immediately puts them in touch with another organisation - solving the problem immediately, ultimately saving time for the organisation making the enquiry who had no idea where to find such refurbished but also cost as such devices as cheaper.

## **6.0 Discussion & Conclusion**

This paper aimed to investigate digital inclusion network building as a mechanism for reducing digital poverty. The results of this research have evidenced how the importance of 'role' and 'place' are significant contributing factors when it comes to digital inclusion network building concerning understanding the context of where people reside, the localities and venues of where organisations are situated that deliver digital inclusion activities, and the roles of actors involved in building and sustaining the network. A limitation of this research was the short length of the project resulting in a short time period to collect data. Despite that, this research is able to offer the following conclusions and contributions as outlined below.

The results of this study align with previous research in which scholars have identified the importance of place-based solutions to digital inclusion by working collaboratively with the people who live and work locally to gain an understanding of

the needs of the community, but also the assets, resources and geography of a locality (Reisdorf and Rhinesmith, 2018; Park et al., 2019). Furthermore, the importance of facilities and the appeal of places and spaces are revealed to be crucial in digital inclusion initiative provision. As a result, the findings of this research conclude that the design of spaces should be prioritised to facilitate interpersonal interactions, fostering socialisation, promoting a sense of inclusivity, and ensuring overall comfort for individuals. Dedicated spaces should be intentionally designed and funded as part of digital inclusion initiatives and networks. These outcomes align with previous research that examined public access to information technology venues from a perspective of development and social change, which emphasised the importance of the “cool factor” (venues and spaces that are attractive, cosy, friendly, reliable) and social interaction in addressing digital access. The concept of ‘coolness’ and its influence on the access and interaction with technologies by social groups at risk of digital exclusion is still very little considered by the IS literature and certainly needs to be further investigated (Gould and Gomez, 2010; Vannini et al. 2015).

The use of volunteers revealed in this study also aligns with previous research on how they are deeply embedded within digital inclusion. Indeed, this study evidences how the network brought people together (both paid staff and volunteers). However, this study also revealed what could be argued as an overreliance on volunteers, as noted by previous research (Casselden et al., 2019). Such reliance makes the sustainability of digital inclusion activities fragile, as the availability of volunteers changes, particularly during a cost-of-living crisis, where they may no longer be able to afford to volunteer. A significant barrier revealed by both paid staff and volunteers was the lack of recognition by national and local policymakers of the digital inclusion work being provided. Such views align with previous research which highlights the disconnect between digital inclusion policy and practice (Wagg and Simeonova, 2022).

Finally, this study used NW to make sense of the data. It enabled us to reveal nuances in how digital inclusion networks are developed and operate and evidenced the role individuals and organisations play as part of that network. NW enabled our study to reveal a granular understanding of how different organisations rely on actors who possess a variety of motivations and experiences (as per Holley, 2013, everyone is a



network weaver), some operating through other networks, revealing the potential impact of how established networks operating within networks at local, regional and national and policy level can provide beneficial outcomes and generate capacity building. As a result, this paper reveals opportunities for change in the provision of digital inclusion initiatives through network building with an emphasis on ‘roles’ and ‘place’ that have implications for future digital inclusion delivery and policy and practice.

## 7.0 References

Allmann, K. and Blank, G. (2021). Rethinking digital skills in the era of compulsory computing: methods, measurement, policy and theory, *Information, Communication & Society*, 24(5), 633-648

Allmann, K. (2022). UK Digital Poverty Evidence Review 2022, Digital Poverty Alliance. Available at:  
<https://digitalpovertyalliance.org/uk-digital-poverty-evidence-review-2022/executive-summary/>

Al-Muwil, A., Weerakkody, V., El-haddadeh, R. and Dwivdei, Y. (2019). Balancing digital-by-default with inclusion: a study of the factors influencing E-inclusion in the UK, *Information Systems Frontiers*, 21(3), 635-659.

Asmar, A., Van Audenhove, L. and Mariën, I. (2020) Social Support for Digital Inclusion: Towards a Typology of Social Support Patterns, *Social Inclusion* 8,(2), 138–150.

Bach, A., Shaffer, G. and Wolfson, T. (2013). Digital human capital: Developing a framework for understanding the economic impact of digital exclusion in low-income communities, *Journal of Information Policy*, 3, 247-266.

Bowen, G. A. (2009). Document Analysis as a Qualitative Research Method. *Qualitative Research Journal*, 9(2), 27-40.

Blackwell, C., Davis, A., Hill, K., Padley, M., Yates., and Stone, E. (2023). A UK Minimum Digital Living Standard for Households with Children. Available at:

[https://www.lboro.ac.uk/media/wwwlboroacuk/content/crsp/downloads/reports/MDLS%20UK%20report\\_Final.pdf](https://www.lboro.ac.uk/media/wwwlboroacuk/content/crsp/downloads/reports/MDLS%20UK%20report_Final.pdf)

Braun, V., Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.

Cabinet Office (2014). Government Digital Inclusion Strategy. Available at <https://www.gov.uk/government/publications/government-digital-inclusion-strategy> (accessed 13 January 2023).

Casselden, B., Pickard, A., Walton, G. and McLeod, J. (2019). Keeping the doors open in an age of austerity? Qualitative analysis of stakeholder views on volunteers in public libraries. *Journal of Librarianship and Information Science*, 51. 869-883.

Creswell, J. (2014). *Educational Research: Planning, conducting and evaluating quantitative and qualitative research*. Essex, UK: Pearson.

Cook, J., Thiery, H., & Burchell, J. (2023). No Longer 'Waiting for the Great Leap Forwards'? Advances in Local State-Voluntary and Community Sector Relationships During Covid-19. *Journal of Social Policy*, 1-20.

Davies, H.C., Eynon, R. and Wilkin, S. (2017). Neoliberal gremlins? How a scheme to help disadvantaged young people thrive online fell short of its ambitions, *Information, Communication and Society*, 20(6), 860-875.

Diaz Andrade, A. and Techatassanasoontorn, A.A. (2021). Digital enforcement: rethinking the pursuit of a digitally-enabled society, *Information Systems Journal*, 31(1), 184-197.

Deloitte (2023). Digital Poverty in the UK: A socio-economic assessment of the implications of digital poverty in the UK, September 2023, on behalf of Digital Poverty Alliance. Available at: <https://digitalpovertyalliance.org/digital-poverty-in-the-uk-a-socio-economic-assessment-of-the-implications-of-digital-poverty-in-the-uk/>

Eom, S.J., and Lee, J. (2022). Digital government transformation in turbulent times: Responses, challenges, and future direction, *Government Information Quarterly*, 39(2), 1-9.

Faith, B., Henrnandez, K., Beecher, J. (2022). Digital Poverty in the UK. IDS Policy Briefing 202. Institute of Development Studies, accessed via:  
<https://www.ids.ac.uk/publications/digital-poverty-in-the-uk-accessible-version/>

GMCA (2023). Greater Manchester Digital Skills map, available at  
<https://www.greatermanchester-ca.gov.uk/news/get-online-greater-manchester-digital-skills-map/>

Gómez, R. & Gould, E. (2010). The “Cool Factor” of Public Access to ICT: Users’ Perceptions of Trust in Libraries, Telecentres and Cybercafés in Developing Countries. *Information Technology & People*, 23, 3, 247-264.

Gomez, R. (2013). The changing field of ICTD: Growth and maturation of the field, 2000–2010. *The Electronic Journal of Information Systems in Developing Countries*, 58(1), 1-21.

Gomez, R., & Pather, S. (2012). ICT Evaluation: are we asking the right questions?. *The Electronic Journal of Information Systems in Developing Countries*, 50(1), 1-14.

Guenther, J., Smede, B., and Young, M. (2020). Digital inclusion in central Australia: what is it and what makes it different? *Rural Society*, 29(3), 154-170

Hemerling, J., Kilmann, J., Danoesastro, M., Stutts, L., and Ahern, C. (2018). *It's not a digital transformation without a digital culture* Boston Consult. Group, 1–11.

Helsper, E. and Reidorf, C. (2017). The emergence of a ‘digital underclass’ in Great Britain and Sweden: changing reasons for digital exclusion, *New Media and Society*, 19(8), 1253-1270.

Holley, J. (2013) *Introduction to Network Weaving*. Network Weaver Publishing, Athens, Ohio, USA

Holmes, H. and Burgess, G. (2022) Digital exclusion and poverty in the UK: How structural inequality shapes experiences of getting online, *Digital Geography and Society*, 3, 1-9

LOTI (2022) Digital Inclusion Innovation Programme, London Office of Technology and Innovation available at [https://loti.london/wp-content/uploads/2022/02/LOTI-Digital\\_Inclusion\\_Report.pdf](https://loti.london/wp-content/uploads/2022/02/LOTI-Digital_Inclusion_Report.pdf)

Lucas, P.J., Robinson, R. & Treacy, L. (2020). What is Data Poverty? London, Nesta. Available at [https://media.nesta.org.uk/documents/What\\_is\\_Data\\_Poverty.pdf](https://media.nesta.org.uk/documents/What_is_Data_Poverty.pdf)

Lythreatis, S., Singh, S.K. and El-Kassar, A.B. (2022). A Digital divide: A review and future research agenda, *Technological Forecasting & Social Change*, 175.

Macmillan, R. (2021). A surprising turn of events - episodes towards a renaissance of civil society infrastructure in England. *People, Place and Policy*, 15(2), 57-71

Madon, S., Reinhard, N., Roode, D. and Walsham, G. (2009). Digital inclusion projects in developing countries: processes of institutionalization, *Information Technology for Development*, 15(2), 95-107

Marais, M., Vannini, S. (2021) Network Weaving to foster Resilience and Sustainability in ICTD. At IFIP WG 9.4 Virtual Conference “Resilient ICT4D”. 26-28 May 2021.

Mason, K., Wagg, S., Ge, B., Harrison, B., Hayes, N., Perez, D., Walker, T. and Wilkes, M. (2022). Digital poverty transformation: accessing digital services in rural northwest communities. Lancaster University. Available at <https://www.thebritishacademy.ac.uk/publications/digital-poverty-transformation-accessing-digital-services-in-rural-northwest-communities/> (Accessed 13 January 2023).

Mervyn, K., Simon, A. and Allen, D.K. (2014). Digital inclusion and social inclusion: a tale of two cities, *Information, Communication and Society*, 17(9), 1086-1104.

Nathaniel-Ayodele, S., McGrath, T. (2023) Internet access: essential utility or human right? Data Poverty Lab, available at <https://www.goodthingsfoundation.org/wp-content/uploads/2023/04/Internet-Access-Report-Final.pdf>

Park, S., Freeman, J., & Middleton, C. (2019). Intersections between connectivity and digital inclusion in rural communities. *Communication Research and Practice*, 5(2), 139-155.

Philip, L. and Williams, F. (2019). Remote rural home-based businesses and digital inequalities: Understanding needs and expectations in a digitally underserved community, *Journal of Rural Studies*, 68, 306-318.

Reisdorf, C. and Rhinesmith, C. (2018), Digital inclusion in international perspective—an asset-based approach to digital inclusion research in the US context, in Ragnedda, M. and Mutsvairo, B. (Eds), *Digital Inclusion an International Comparative Analysis*, Rowman and Littlefield, Lanham, MD.

Reisdorf, B.C. and Rhinesmith, C. (2020). Digital Inclusion as a Core Component of Social Inclusion, *Social Inclusion*, 8(2), 132–137

Schou, J., & Hjelholt, M. (2018). Digital citizenship and neoliberalization: governing digital citizens in Denmark. *Citizenship Studies*, 22(5), 507-522.

The British Academy (2022) *Understanding Digital Poverty and Inequality*. The British Academy, London. Available at <https://www.thebritishacademy.ac.uk/publications/understanding-digital-poverty-and-inequality-in-the-uk/>

The British Academy (2023) *Digital Technology and Inequality: Policy Brief*, February 2023. Available at: <https://www.thebritishacademy.ac.uk/publications/digital-technology-and-inequality-policy-brief/>

Torrecillas, C., Centeno, C. and Misuraca, G. (2014). Measuring the impact of inclusion intermediary actors: Characterisation and mapping of Inclusion intermediary Actors in the EU27, European Commission, JRC-IPTS Scientific and Political Report.

Smith, S., Bellaby, P., & Lindsay, S. (2010). Social inclusion at different scales in the urban environment: Locating the community to empower. *Urban studies*, 47(7), 1439-1457.

UK Digital Strategy (2022). Policy Paper, Department of Digital, Culture, Media & Sport, accessed via: <https://www.gov.uk/government/publications/uks-digital-strategy>

Vannini, S., Rega, I., Sala, S., Cantoni, L., (2015). Using Photo-elicitation to Explore Social Representations of Community Multimedia Centres in Mozambique. *The Electronic Journal of Information Systems in Developing Countries (EJISDC)*, 67,8.

Vannini, S., Nemer, D., Halabi, A., Sabiescu, A.G., David, S. (2017). Critical Incidents Analysis: mismatching expectations and reconciling visions in intercultural encounters. *Journal of Community Informatics*, 13(2), Special Issue: CIRN Conference, 2016.

Valentín-Sívico, J., Canfield, C., & Egbue, O. (2022). Push them forward: Challenges in intergovernmental organizations' influence on rural broadband infrastructure expansion. *Government Information Quarterly*, 39(4), 101752.

Wagg, S., Vannini, S., Zamani, E., Klyshbekova, M., Jia, X. and Aylward, B. (2024) Digital inclusion network development: A case study in Derbyshire. The British Academy. [in press]

Wagg, S. and Simeonova, B. (2022). A policy-level perspective to tackle rural digital inclusion, *Information Technology & People*, 35(7), 1884-1911.

Wagg, S. (2021) An Investigation of Digital Inclusion in UK Rural Communities. PhD, Loughborough University.

Yin, R. K. (2009). *Case Study Research: Design and Methods* (4th ed.). Thousand Oaks, CA: Sage Publications.

Zamani, E. D., Vannini, S. (2022). Understanding Digital Poverty in South Yorkshire. Partnerships and Regional Engagement (PRE), South Yorkshire Office for Data Analytics.

Zamani, E. D., Griva, A., & Conboy, K. (2022). Using business analytics for SME business model transformation under pandemic time pressure. *Information Systems Frontiers*, 24(4), 1145-1166.

Zheng, Y., & Walsham, G. (2021). Inequality of what? An intersectional approach to digital inequality under Covid-19. *Information and Organization*, 31(1), 100341.

---

# Conceptualising the Artificial Intelligence Divide: A Systematic Literature Review and Research Agenda

## Authors

Aishatu Mohammed Lawan  
Mohammed Ali  
Maria Kutar

## Affiliations

University of Salford, Manchester UK

## *Completed Research*

### Abstract

*As artificial intelligence (AI) increasingly shapes our world, the AI divide, comprising access, skills, and outcomes, becomes a critical issue. The AI divide encompasses disparities in access, skills, and ethical outcomes related to AI technology. This paper employs a socio-technical framework to comprehensively understand and address the multifaceted challenge of the AI divide. A systematic literature review was conducted, analysing 46 studies that examined various aspects of the AI divide. Three levels of the AI divide emerged: access, skills, and outcomes. Geographical disparities, algorithmic literacy, digital skills, and algorithmic bias were identified as key factors influencing the AI divide. Accessible AI technology, digital and algorithmic literacy, unbiased AI outcomes, and effective regulation are essential for a more equitable AI landscape. We conclude that a socio-technical approach is crucial for bridging the AI divide, ensuring that the benefits of AI are accessible to all in our rapidly evolving technological landscape.*

**Keywords:** AI Divide, Artificial Intelligence, Digital Divide, Socio-Technical, Systematic Literature Review

## 1.0 Introduction

The rapid advancement of artificial intelligence (AI) applications has ushered in an era where inequalities in access to AI technologies among different countries and organisations have started to emerge. Much like the well-documented digital divide related to disparities in internet access (Van Dijk, 2020), a new phenomenon, the "AI divide" or "algorithmic divide," is taking shape. This divide is hindering a significant portion of the global population, spanning both advanced and developing nations, from fully realising the potential of machine learning and artificial intelligence technologies (Yu, 2020). Therefore, this discourse is essential to understand how to prevent the widening of the existing digital divide, which efforts have long sought to bridge. Although there is emerging discussion in the area, few studies have systematically and comprehensively reviewed the AI divide. (Dwivedi et al., 2021) argued for more



research on AI societal impacts, and as we witness the rapid increase in AI technologies and their adoption, addressing this issue has become essential.

The primary purpose of the paper is to conduct a systematic literature review (SLR) to uncover the breadth and depth of previous research efforts related to the concepts and contexts of the AI divide. Additionally, the researchers aim to construct a conceptual model based on the findings. As a corollary, the researchers will also identify gaps in the existing literature, thereby providing a research direction that will guide future researchers interested in studying the AI divide.

To gain a comprehensive understanding of prior research, data extracted from multiple databases is analysed and synthesised thematically using the three levels of the digital divide framework identified by (Ragnedda & Muschert, 2017; J. Van Dijk, 2020). The adoption of this three-level digital divide literature framework is instrumental in achieving a well-structured analysis and synthesis of data obtained from the SLR. This framework categorises digital disparities into three distinct levels: access, skills, and outcomes. By organising our themes based on these levels, we establish a logical and systematic approach to understanding the multifaceted aspects of digital inequality, thereby placing the exploration of AI inequality on a solid foundation based on the three-level digital divide. The integration of this framework enhances the depth and clarity of our analysis, allowing us to examine how disparities in digital access and skills contribute to varying outcomes. The objective of this systematic literature review is threefold:

- To discover the breadth and depth of the AI divide literature.
- To identify the gaps in the literature and propose new research directions.
- To inform the development of a conceptual model of the AI Divide.

## **1.1 Digital Divide**

The term "digital divide" was first used by Rogers, (1962) to describe the difference in digital access between current and potential users. Rogers' approach, however, merely emphasises how user requirements affect accessibility. Other researchers continued to study the digital divide; studies on the digital divide have concentrated on a range of environments, including mobile devices, e-government, the Internet, and financial technology.

According to the OECD, (2001), the term "digital divide" refers to the disparity in access to and use of information and communication technologies (ICTs) and the Internet for a wide range of purposes between individuals, households, businesses, and geographic areas at different socio-economic levels. This statement suggests that the concept of the digital divide pertains to the disparity between individuals who possess internet connections and those who lack such access. Early researchers in the field of the digital divide examined the disparity in individuals' access to technological resources, such as computers and the Internet (Van Dijk, 2006). However, as technology develops over time, the definition of the "digital divide" expands to include more than just internet access, necessitating the adoption of a comprehensive perspective.

To define the digital divide, Hilbert, (2011) identified four factors: 1) The individual who accesses the technology (the agent); 2) the attributes or traits possessed by the individual (the agent's characteristics); 3) the way the individual establishes a connection (the method of connection); and 4) the specific systems to which the individual establishes connections (the target systems). The categorisation of the digital divide has been undertaken by researchers, who have identified three main levels. For example, Wei et al., (2011) cited in Carter et al., (2020) identified three levels of the digital divide (p. 170):

*“The digital access divide (the first-level digital divide) is the inequality of access to information technology (IT) in homes and schools. The digital capability divide (the second-level digital divide) is the inequality of the capability to exploit IT arising from the first-level digital divide and other contextual factors. The digital outcome divide (the third-level digital divide) is the inequality of outcomes (e.g., learning and productivity) of exploiting IT arising from the second-level digital divide and other contextual factors.”*

While prior studies primarily concentrated on the digital divide's first and second levels, third-level digital divides, which emphasise ICT's concrete impact and involvement, are proposed as being as important as the first- and second-level digital divides (Scheerder et al., 2017). In other words, having access to and using technologies does not always translate into a positive outcome (Carter et al., 2020). This research references the conceptual framework proposed by Hargittai, (2002) and Van Deursen

& Helsper, (2015) to analyse the first (access), second (usage), and third-level (outcome) digital divides (Lutz, 2019). According to Lutz, (2019), this technique provides a straightforward approach to organising the literature.

Heeks, (2022) argued that over the past decade, the digital divide has encompassed more than one digital divide but multiple digital divides, such as technology divides like AI. As the world advances in the age of Industry 4.0, with AI at its core, it is important to study the digital divide from the perspective of not only Internet inequalities but also inequalities in AI. This leads to what is being referred to as the "AI divide."

## **1.2 Artificial intelligence (AI) Divide**

The field of artificial intelligence (AI) has rapidly evolved since John McCarthy's initial use of the term in 1956 (McCarthy 2011 cited in Bjola, 2022). AI encompasses a wide array of disciplines, methodologies, and applications, seeking to replicate human cognition and learning through advanced algorithms (Bjola, 2022). In a broader sense, AI approximates human or animal cognition using machines and algorithmic systems to mimic learning. With its roots in fields as diverse as robotics, electrical engineering, economics, and philosophy, AI has become an integral part of various sectors, including business. In the business domain, the implementation of AI has been studied extensively for its potential to enhance productivity and decision-making (Al-Surmi et al., 2022). Recent research has shown that AI, particularly generative AI, can generate substantial value and improve organisational performance, contributing billions to global corporate profits (MGR, 2023). As AI's capabilities continue to expand, it is increasingly integrated into daily life, reflecting its significant influence (Elliott, 2019).

However, as AI's influence grows, concerns about the unequal distribution of its benefits have emerged, leading to the concept of "AI divides" (Yu, 2020.). Yu introduced the term "algorithmic divide," which encompasses various aspects such as awareness, access, affordability, availability, and adaptability (Yu, 2020). This means the AI divide encompasses disparities in access, skills, and outcomes related to AI technologies across individuals, organisations, and nations. This divide gives rise to algorithmic deprivation, discrimination, and distortion issues that impact both individuals and organisations.

At the individual level, the AI divide predicts that employment opportunities involving low-level digital abilities may decline while those requiring high-level AI skills will increase (Carter et al., 2020). The demand for advanced digital skills has translated into higher wages for those possessing AI literacy, creating disparities among individuals. At the organisational level, industries equipped with advanced AI technologies exhibit greater productivity, giving them a competitive advantage (Gmyrek P et al., 2023). Schwab, (2020) predicts significant job displacement and the creation of new opportunities by 2025, requiring a shifting division of labour between humans, machines, and algorithms. In developing countries, the potential for AI-driven job augmentation exists, but inadequate infrastructure, such as broadband connectivity and electricity supply, limits practical benefits (Gmyrek et al., 2023). Nations adopting AI technologies gain economic advantages, leading to disparities in AI technology gains.

As we navigate the age of AI, it is imperative for individuals in both developed and developing nations to cultivate knowledge about the positive and negative aspects of AI technologies. While AI offers significant advantages, not all individuals have access to AI-enhanced technology products and services. This lack of access contributes to a state of "algorithmic deprivation" (Yu, 2020), leading to disparities in access, utilisation, and outcomes. To bridge the AI divide and prevent the further widening of pre-existing digital divides, understanding and conceptualising this multifaceted challenge is vital.

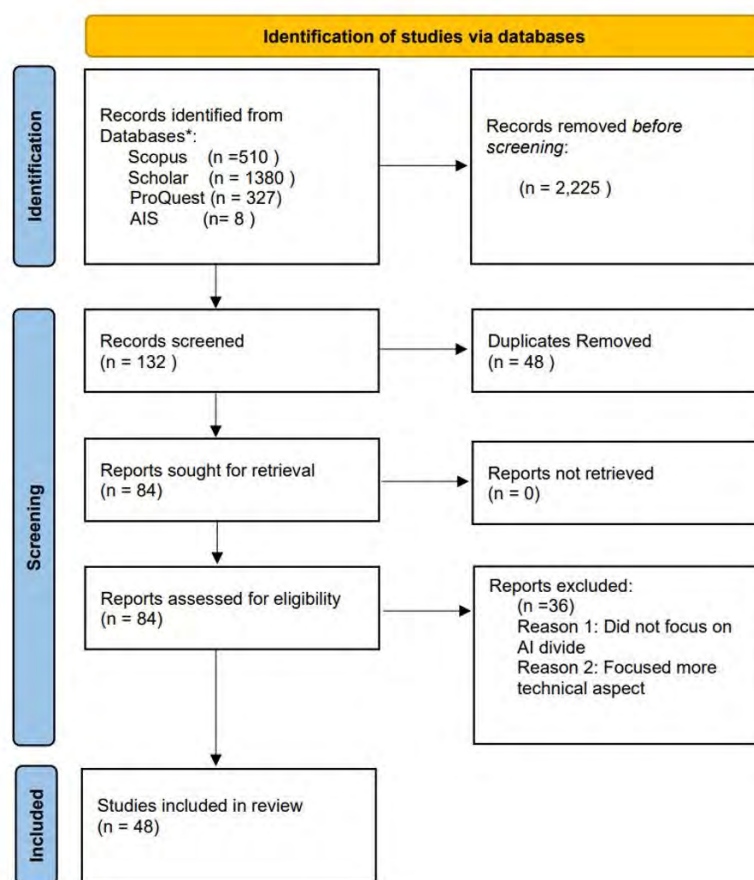
## **2.0 Methodology**

This paper revolves around a systematic literature review (SLR), which offers a comprehensive overview of academic literature. We conduct a thematic analysis using the three levels of the digital divide—access, skills, and outcomes—which form the basis of our conceptual framework. Throughout the analysis, we identify gaps in the literature that point to areas for future research. Our paper begins by presenting the SLR methodology, followed by the analysis and conceptual framework.

### **2.1 SLR Methodology**

The study employed a systematic literature review (SLR) to assess the current academic literature regarding the AI divide, identify research gaps, and delineate critical areas for further investigation. SLR is a rigorous and recognised approach for systematically collecting, organising, and analysing relevant research papers on a specific topic (Fink,

2019; Okoli, 2015; Petticrew & Roberts, 2008). Given the emerging nature of the study area, SLR offers a structured method for synthesising prior research findings (Webster & Watson, 2002) and serves as a valuable reference point for future scholarly investigations (Kitchenham et al., 2011; Paré et al., 2015). Borges et al., (2021) highlight the thorough examination and comprehensive understanding of various facets of the topic that SLR allows, further emphasising the role of systematic reviews in evaluating the breadth and diversity of research efforts within a field (Paré et al., 2015). Hence, our systematic review adhered to the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework proposed by (Liberati et al., 2009) (see Figure 1).



**Figure 1. Prisma Flow Diagram.**

PRISMA provides a comprehensive explanation and elaboration of the reporting checklist items derived from PRISMA standards, which are widely accepted in the social sciences for conducting systematic reviews. Applying the PRISMA 2020 criteria was pivotal in selecting the publications for this research. According to Page et al., (2021), the use of PRISMA enhances the value of systematic reviews for both users and

authors by facilitating the preparation of transparent, comprehensive, and accurate documentation of the included papers, enabling the review to fulfil various essential roles.

### **2.1.1 Method of Data collection**

While some systematic reviews (Hanelt et al., 2021) rely on a single database as their primary data source, using multiple databases is an established strategy to enhance the comprehensiveness of the research sample. This methodology is commonly employed in systematic reviews (e.g Scheerder et al., 2017)). The present study followed this methodology by utilising four databases, including SCOPUS, ProQuest, AIS, and Google Scholar.

### **2.1.2 Inclusion & Exclusion Criteria**

The inclusion criteria for the review involve selecting articles and conference papers published within the last decade (2013–2023), written in English, and related to the fields of business, management and accounting, decision sciences, social sciences, arts, and humanities. While the concept of AI is not entirely new(OECD, 2019), the implications of the AI divide have gained significant attention in the past decade(Lewicki et al., 2023). These studies must also include relevant search terms in their title, abstract, or keywords and undergo peer review.

Exclusion criteria include the exclusion of languages other than English. No exclusions were made based on methodology, allowing for the inclusion of empirical, theoretical, and conceptual studies. Additionally, research that did not specifically address the AI divide was excluded from the review. For example, papers discussing the digital divide were excluded because they have been extensively discussed in the literature. By the early 20th century, there were over 14,000 papers dedicated to examining the issue of the digital divide ((L. Yu, 2006).

### **2.1.3 Search Strategy**

The databases were examined using the search string provided in Table 1 for the conducted searches. The next step involved the process of crafting and refining the search. During this step, the search keywords were based on the main concepts of the review questions as well as the inclusion and exclusion criteria outlined in the study plan. These keywords were used to search the titles or abstracts of the database items.

The third phase encompassed the execution and documentation of the search process. Due to variations in database structure and indexing strategies, a different search strategy was required for each database.

Search String	Keywords
<b>1st Search</b>	("algorithmic inclusion" OR "algorithmic divide" OR "algorithmic fairness" OR "Generative AI" OR "artificial-intelligence" OR "artificial intelligence" OR "machine learning" OR "emerging technologies" OR "digital transformation") AND ("socio-technical implications" OR "social implications" OR "ethical implications" OR "organisational change" OR "socio-technical perspective")
<b>2nd Search</b>	("socio-technical implications" OR "social implications" OR "socio-technical perspective*" OR "socio technical systems") AND ("algorithmic inclusion" OR "algorithmic divide" OR "algorithmic decision making" OR "algorithmic fairness" OR "artificial intelligence" OR "machine learning" OR "AI")
<b>3rd Search</b>	"algorithmic divide" OR "AI divide"

**Table 1. Search Strings.**

The fourth step involved managing the search results. Searches were conducted using three distinct search strings across multiple databases to comprehensively investigate the AI divide. The first search string was designed to explore the intersection of AI and socio-technical implications. In ProQuest, this initial search yielded 297 results ( $n = 297$ ), of which 25 were screened. In Scopus, the same search criteria generated a larger pool of 1,201 results ( $n = 1,201$ ). Following rigorous inclusion and exclusion criteria, which focused on a 10-year timeframe (2013–2023), articles, and conference papers, the pool was narrowed down to 510 papers ( $n = 510$ ). In Google Scholar, the initial search string, "Exact phrase of string 2013-2023," produced 1,380 results ( $n = 1,380$ ), with 12 papers selected for further review. In the AIS database, the first search string produced 285 results ( $n = 285$ ), out of which 25 were removed due to duplicate studies. The second search string aimed to delve into the socio-technical perspectives of the AI divide, considering algorithmic inclusion, decision-making, and the implications of AI. In ProQuest, the second search generated 212 results ( $n = 212$ ), with five papers making it through the screening process. Scopus provided a more extensive result pool of 633 papers ( $n = 633$ ), of which 9 were selected for further analysis. In AIS, the second string yielded 461 results, and 25 were selected. The third search string, "algorithmic divide" OR "AI divide," was conducted to specifically address the AI divide aspect. In ProQuest, 20 results were found ( $n = 20$ ), with 3 papers screened, while Scopus yielded 14 results, with 1 paper making it through the screening process. In Google Scholar, the

third search string generated 349 results ( $n = 349$ ), with 20 papers screened. In AIS, the third string revealed only 6 results, with 8 studies selected after removing duplicates. In total, 82 papers were pooled ( $n = 82$ ), and 36 were removed as duplicates. Overall, a total of 46 papers ( $n = 46$ ) were screened across all databases as part of this systematic review, allowing for a comprehensive examination of the AI divide within the specified research parameters.

#### 2.1.4 Assessment of Literature

The researcher developed data extraction forms that included essential article details, such as title, authors, publication year, journal, study setting, data type, and research method. Following the approach recommended by Tranfield et al., (2003), this method helps reduce bias and human error while providing data for further analysis. The forms also included the level of the digital divide addressed, ensuring a comprehensive data collection process. Some papers, despite being identified by the three search strings, did not address any level of the digital divide and are not reported here. These papers fell into two broad categories: the impact of AI on organisations and the challenges and opportunities of AI.

#### 2.1.5 Categorisation of Reviewed Studies

To systematically categorise the studies, this research employed a framework based on the first, second, and third levels of the digital divide, adapting it to the context of the AI divide (see Table 2).

Level	Description
First Level: Access	All studies that address disparities in access to AI technologies were categorised at this level. These studies primarily focus on the initial stage of AI utilisation, emphasising accessibility issues. For example, Algorithmic awareness and Physical access to AI technologies are grouped as First Level.
Second Level: Usage	Studies examining inequalities in the utilisation of AI were grouped under the second level of the AI divide. These investigations delve into how AI technologies are employed, highlighting variations in usage patterns e.g Algorithmic skills to use AI
Third Level: Outcome	Studies that delve into the outcomes of AI implementation and their impact on individuals, organisations, and nations were classified under the third level of the AI divide. This level concerns itself with the consequences and effects of AI engagement such as Ethics, Bias, and fairness,

**Table 2. Study Categorisation.**



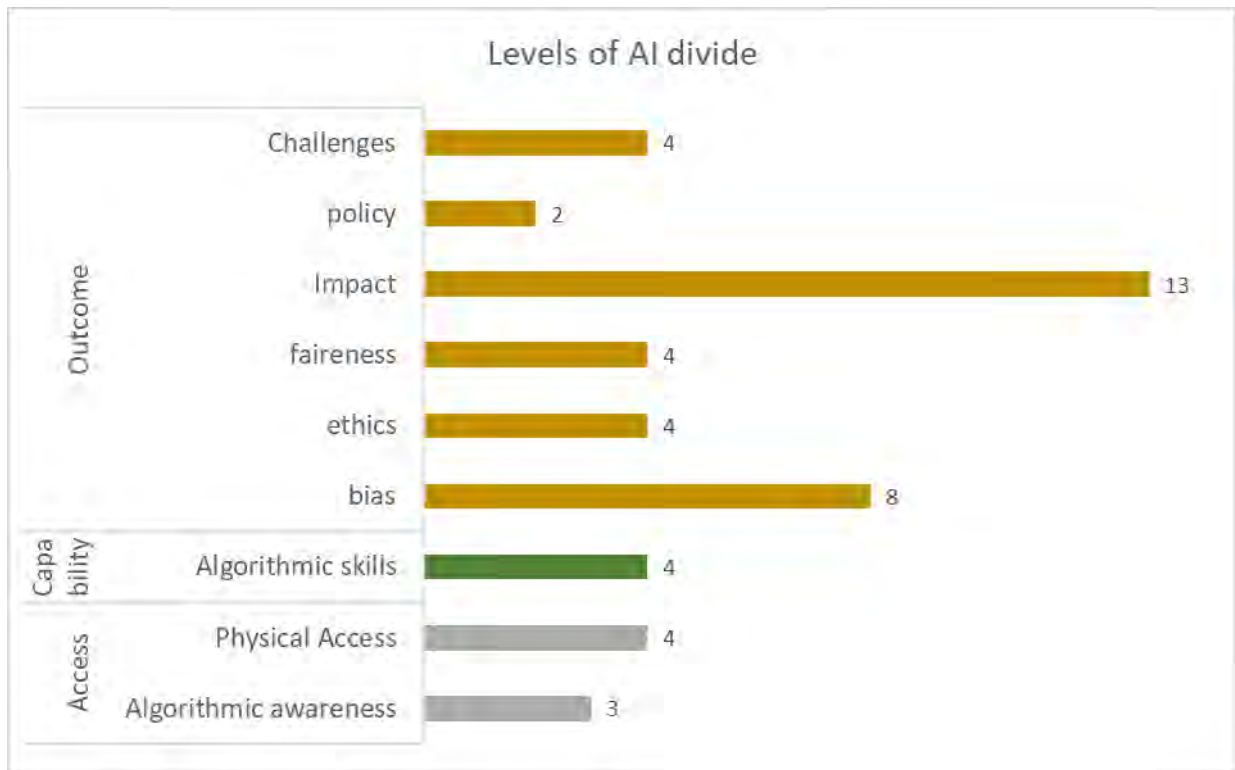
By adopting this categorisation framework, this research seeks to provide a structured and comprehensive analysis of the AI divide across its various dimensions, enhancing our understanding of this evolving phenomenon.

### **3.0 Development of the Conceptual Framework**

Utilising the categorised and reviewed studies, a conceptual model has been formulated to conceptualise and measure the AI divide, which will be elaborated upon in the following sections. This process consists of two parts: First, an **initial analysis of the SLR according to the three levels of the AI divide** and the **identification of the drivers for the divide**. Second, a detailed discussion and synthesis of the **social, technical, and socio-technical factors influencing the digital divide** is provided.

#### **3.1 Conceptualising the AI Divide: The 3 levels of AI Divide**

This paper will present a thematic analysis of the AI divide literature, structured around the established digital divide levels (Carter et al., 2020; Lutz, 2019). Carter et al. (2020) introduced the concept of the AI divide, which consists of three key elements: the access gap, capacity gap, and result gap. The authors also proposed theoretical frameworks encompassing the conceptualisation, modelling, and analysis of the divide, which serve as the foundational structure for understanding the digital divide within the realm of artificial intelligence (AI). These frameworks lay the groundwork for a socio-technical research agenda that addresses the evolving role of AI in the context of the digital divide. Figure 2 illustrates the themes with their corresponding number of studies.



**Figure 2. Quantified Studies on the Levels of AI divide.**

### **3.1.1 First Level of AI divide: Access Level**

The concept of the 'first-level digital divide', which pertains to limited access to digital resources, has been the subject of extensive research. Considering the reviewed literature, Carter et al., (2020) advocated for an exploration of the AI divide through the lens of the first level of the digital divide. In this context, the themes of algorithmic awareness and physical access to AI technologies correspond to the first level of the AI divide. These themes align with the discussion of the first level of the digital divide, where motivation and physical access were identified as the main components. Notably, various studies out of the 46 papers included in our review underscored physical access to AI technologies and algorithmic awareness as significant contributors to the AI divide ( $n = 7$ ; 15.2%).

#### **3.1.1.1 Algorithmic Awareness**

Several studies ( $n = 3$ ; 6.5%) have presented arguments on algorithmic awareness as a major cause of the AI divide (Lauterbach, 2019; Lythreatis et al., 2022; Zarouali et al., 2021).

Zarouali et al., (2021) argue that the literature regarding algorithmic awareness is characterised by several key factors: 1) limited scope of research; 2) inconclusive

conclusions derived from these studies; and 3) a predominant focus on algorithms within specific mediated contexts, limiting the generalizability of their findings. This analysis reveals potential indications of a significant deficiency in algorithmic knowledge within the general populace, especially among specific vulnerable demographic cohorts. The present findings prompt a discussion about the potential for AI technologies to exacerbate rather than alleviate digital disparities. Ensuring that individuals possess equitable knowledge is crucial for maximising the advantages derived from algorithmic systems. At the very least, it is imperative to provide equal opportunities for individuals to cultivate these knowledge. Failure to do so may result in the emergence of the AI divide, as examined previously.

According to Lauterbach, (2019), policymakers must demonstrate leadership to address the challenges associated with artificial intelligence (AI) and promote the importance of self-education. Policymakers should provide information regarding the impact of AI at regional and industrial levels, as well as identify the distinct advantages that different countries possess to capitalise on the opportunities presented by AI.

Lythreathis et al., (2022) proposed that the examination of algorithmic awareness and data disparities gives rise to a new digital divide. This phenomenon offers a significant opportunity to contribute to the digital divide literature. If there is a lack of effort in clarifying AI technology and fostering a discourse among many stakeholders in society, there is a possibility of experiencing an exacerbation of inequality, prejudice, and marginalisation(Lauterbach, 2019).

*Research Gap 1: What is the extent and impact of the AI awareness gap?*

### **3.1.1.2 Physical Access**

Several studies ( $n = 4$ ; 8.7%) have discussed physical access as a component of the AI divide (Ade-Ibijola & Okonkwo, 2023; Haber, 2020; Wu, 2022; Yu, 2020.). Physical access encompasses not only access to AI technologies but also their peripherals. Although AI has the potential to provide significant advantages for individuals and society, not all individuals have access to AI technologies.

At the domestic level, individuals may face exclusion due to various factors, such as financial constraints, limited availability in the local market, or technical limitations, including inadequate skills to use these technologies. At the global level, the issue of

access has become increasingly severe, especially when considering the restricted availability of computing resources, internet connectivity, and advanced communication technologies in underdeveloped nations (Yu, 2020). For instance, the limited level of internet penetration in Africa can be attributed to infrastructural challenges, such as inadequate access to key internet infrastructure components like fibre optic cables, cell towers, and base stations (Ade-Ibijola & Okonkwo, 2023). Internet penetration rates in Japan, the United States, and the United Kingdom exceed 90%, while in countries like Burundi, the Central African Republic, Eritrea, and Western Sahara, these rates are approximately 5% or lower. As a result, the availability of algorithm-enhanced technological products and services cannot be assumed in developing nations (Yu 2020).

Wu, (2022) found that while the disparity in Internet accessibility has decreased, there is a significant gap in the accessibility of AI-enabled technologies. This means that the extent to which individuals may benefit from machine learning and artificial intelligence is inversely proportional to their level of availability and affordability. The absence of such access will result in individuals situated on the disadvantaged end of the spectrum being deprived of numerous political, social, economic, cultural, educational, and career prospects made available using AI. Moreover, the absence of access to these technologies will initiate a detrimental cycle in which individuals with greater access to technology will experience increased economic growth, exacerbating the disparity between those who possess and those who lack such resources (Yu, 2020). It might be argued that to bridge the algorithmic divide, it is essential to comprehend the possible long-term consequences of AI accessibility and the legal structure within which these technologies will function (Haber, 2020).

*Research Gap 2: How do disparities in physical access to AI technologies and variations in algorithmic awareness contribute to the emerging AI divide?*

### **3.1.2 Second Level of AI divide: Capability Level**

Hargittai (2001 cited in (Lutz, 2019) introduced the concept of the second-level digital divide to distinguish between disparities in Internet access, and disparities in skills and utilisation, referred to as the second-level gap. This means the focus of research on digital inequalities has progressively transitioned from examining access to investigating skills and usage. Like the digital divide, it is imperative to consider the AI divide with respect to using AI technology. This section examines the second tier of the digital divide, specifically focusing on individuals' skills in utilising AI technologies and their level of algorithmic literacy. Notably, various studies out of the 46 papers included in our review underscored the second level digital divide pertaining to algorithmic literacy ( $n = 4$ ; 8.7%).

#### **3.1.2.1 The Knowledge Divide: Algorithmic Literacy**

Several studies ( $n = 4$ ; 8.7%) have discussed Algorithmic literacy as a component of the AI divide ((Cameron et al., 2023; Chowdhury et al., 2022; Haber, 2020.; Zarouali et al., 2021)

In the context of this research, AI literacy, or algorithmic literacy, pertains to the ability to use AI technologies, encompassing theoretical and practical competencies for AI system creation and utilisation. As AI integration grows, understanding algorithmic functions becomes crucial due to a lack of user awareness (Cameron et al., 2023). Algorithmic literacy focuses on the societal dimensions of technology, including economic, political, and social aspects.

Algorithmic literacy is essential in today's technology-driven society, where algorithms shape digital platforms. It's vital to recognise that algorithms are not impartial computations but have significant societal implications (Zarouali et al., 2021). This study extends beyond algorithmic literacy to encompass a broader understanding of AI and its applications, enabling informed choices in AI and algorithmic systems.

Haber, (2020.) suggest the possible ways to address the algorithmic divide, such as increasing algorithmic literacy, creating ethical standards, ensuring transparency and accountability, and fostering public awareness and education.

Corporations recognise AI's potential advantages but underutilise it due to low knowledge and awareness of AI capabilities, limitations, and integration within the workforce (Chowdhury et al., 2022).

*Research gap 3: How does algorithmic literacy and the knowledge gap related to artificial intelligence (AI) capabilities impact the usage of AI technologies?*

### **3.1.3 Third Level of AI divide: Outcome level**

The conceptualisation of the digital divide has evolved beyond mere access to the Internet and proficiency in its usage. Researchers argue for a more comprehensive approach by considering the outcomes of Internet usage (Lythreatis et al., 2022; Van Deursen & Helsper, 2015). The emergence of the third-level digital gap, building upon the first- and second-level divides, is a recent scholarly development (Lutz, 2019; Scheerder et al., 2017; Wei et al., 2011). Third-level digital divides relate to disparities in the benefits derived from Internet usage, especially when access and usage patterns are similar.

Lutz, (2019) stress the need to examine both the advantages and disadvantages of Internet usage within third-level digital divide research. Lutz's (2019) analysis underscores the shift in research focus from access to skills and uses to outcomes, emphasising the importance of investigating this aspect. Carter et al., (2020) argue for a nuanced exploration of the elements contributing to the AI divide, particularly at the outcome level of the digital divide. In this study, the "outcome-level" AI divide pertains to inequalities arising from AI technology usage.

Ethics, algorithmic bias, and fairness shape AI outcomes. Biased ethics may perpetuate disparities, while algorithmic bias can lead to unfair outcomes, disproportionately affecting certain groups. Ensuring algorithmic fairness is crucial to preventing discrimination and mitigating the widening of the AI divide.

Notably, various studies out of the 46 papers included in our review underscored the third level digital divide pertaining to privacy, ethics, bias, fairness, legislation, and regulations related to AI usage. ( $n = 35$ ; 76% out of the total papers only 17 specifically discuss AI outcome).

### **3.1.3.1 Ethics**

Several studies ( $n = 4$ ; 8.7%) have discussed Ethics as a component of the AI divide (Ashok et al., 2022; Mbuy & Ortolani, 2022; Munoko et al., 2020; van Bruxvoort & van Keulen, 2021). The role of ethics in the AI divide is of utmost importance, as it encompasses the moral considerations surrounding the development, deployment, and regulation of AI. These ethical considerations ultimately influence how societies navigate the potential benefits and risks associated with artificial intelligence. Decisions guided by ethical principles have the potential to either narrow or widen the AI divide, depending on whether AI technologies are developed and utilised in a manner that promotes fairness, accountability, transparency, and the overall well-being of all individuals and communities ((Munoko et al., 2020)).

The challenges arising from algorithms, including artificial intelligence (AI), extend beyond the confines of technical aspects. To ensure the efficient implementation and utilisation of algorithms, it is imperative to contextualise them within a socio-technical framework encompassing regulations, rules, and organisational structures (van Bruxvoort & van Keulen, 2021). The importance of addressing ethical concerns is underscored by the necessity for continuous monitoring, which involves the implementation of feedback mechanisms among developers, adopting organisations, professionals, and regulators. The significance of considering ethics because of AI implementation is underscored by (Ashok et al., 2022). The significance of establishing governance models that encompass the evaluation of AI system outcomes and encompass the entire lifecycle from design and deployment to decommissioning is emphasised in (Mbuy & Ortolani, 2022).

### **3.1.3.2 Bias**

Several studies ( $n=8$ ; 17.39%) have discussed bias as a component of the AI divide (Hall & Ellis, 2023; Mehrabi et al., 2021; Mitchell et al., 2021; Ntoutsis et al., 2020; Sartori & Theodorou, 2022; Turner Lee, 2018; Vesnic-Alujevic et al., 2020; Yu et al., 2023).

Ntoutsis et al., (2020) define bias as the predisposition or prejudice within an AI system's decision-making process that favours or discriminates against individuals or groups unjustly.

Bias hidden in AI systems, often due to complex design, can amplify over time, resulting in inequitable outcomes (Yu et al., 2023). Most AI relies heavily on data, which can introduce biases learned by algorithms, ultimately leading to biased outcomes. Algorithms may exhibit bias due to design decisions, even without inherent data bias. Such bias can impact practical systems and accumulate further biased data for subsequent algorithms (Mehrabi et al., 2021), making it more complex than a technical issue (Hall & Ellis, 2023). Sartori & Theodorou, (2022) advocate for a socio-technical framework when examining AI, offering a comprehensive understanding from both social and technological perspectives. (Hall & Ellis, 2023) emphasise that viewing algorithmic bias only as a technical challenge and focusing on training data may oversimplify the issue, leading to discriminatory outcomes.

AI's rapid advancement is transforming various industries (Lauterbach, 2019), with the potential to amplify prejudices, heighten cybersecurity vulnerabilities, and impact employment opportunities. It's crucial that AI systems are designed inclusively, considering various dimensions like gender, education, and ethnicity (Sartori & Theodorou, 2022). Biased algorithmic outcomes can negatively affect users, reinforcing pre-existing biases (Mitchell et al., 2021). Various applications, such as chatbots, employment matching, aircraft routing, automated legal help for immigration algorithms, and search and advertising placement, exhibit inherent biases. Some social groups face disadvantages, leading to "institutional bias," where specific institutions favour some groups while disadvantaging others (Vesnic-Alujevic et al., 2020).

Equity is a priority for policymakers addressing algorithmic bias, though it remains a complex challenge. Computer and data scientists consistently weigh data accuracy against sufficiency, especially for specific populations or concerns (Lee, 2018). Inclusive teams in coding work can effectively address bias in data and algorithms (Lauterbach, 2019). To ensure AI's comprehensive impact on various aspects of human existence, designers must ensure accurate representation of all segments of society.

As AI integrates further into daily life, technology developers must acknowledge bias and discrimination, prioritising responsible AI use. Relying solely on technology is insufficient to address the wide range of AI biases and issues (Ntoutsi et al., 2020). Critics argue that AI exacerbates existing biases and inequities due to inherent biases in AI system architecture and the data they use (Johnson & Reyes, 2021).



### **3.1.3.3 Fairness and Transparency**

The current challenge facing AI-based systems pertains to their lack of transparency and explainability, which, in turn, erodes users' trust regarding AI outputs (Chowdhury et al., 2022). To address this challenge, Kusters et al., (2020) emphasise the need for AI research to incorporate decision explainability, transparency regarding dataset bias, and the establishment of evaluation procedures and regulatory bodies, all of which contribute to accountability.

Furthermore, AI practitioners should adopt measures like transparency, explainability, and inclusivity to mitigate biases. It is essential to prioritise the establishment of robustness, security, and data privacy protocols to maintain the integrity and reliability of AI systems (Kusters et al., 2020). Empirical data has affirmed the existence of divergent outcomes across various AI applications. (Hardt et al., 2016) highlights the lack of clarity in the concept of fairness concerning the cost of AI, prompting scholars to encourage future research into alternative notions of fairness and their correlation with users' subjective perceptions of fairness. It is crucial to recognise that fairness is fundamentally a socio-technical concept, implying that what qualifies as fair or unfair is not an inherent attribute of algorithms but rather a subjective decision made by programmers.

In this regard, (Yu, 2020) underscores the importance of creating an enabling environment that includes adequate preparation, legal support, and legislative backing for the effective implementation of algorithms to promote equitable use. This involves integrating transparency, accountability, and impact assessments throughout the algorithmic development process to uphold fairness and mitigate biases.

### **3.1.3.4 Policy, Governance, and Regulation**

Government support, data protection rules, and ethical AI policies can significantly impact the accessibility and fairness of AI technology. Stringent policies and laws in some nations can hinder the widespread adoption of AI technology, placing them at a disadvantage compared to countries with more permissive policies (Gao et al., 2023). Effective AI regulation requires a comprehensive understanding of its implications, economic effects, and complex regulatory landscape (Gao et al., 2023). The expanding use of artificial intelligence across various sectors requires policymakers to adopt an informed, evidence-based, and forward-looking perspective on the potential

implications of AI. Currently, there is a lack of a comprehensive framework for the design and governance of AI technology that includes all relevant stakeholders. The decisions made today regarding technology will have long-lasting effects on individuals and businesses worldwide. Without addressing issues of inequality, marginalisation, and discrimination through open dialogue among societal actors, these problems are likely to worsen in the future.

*Research Gap 4: How do privacy, ethics, bias, fairness, transparency, and policy factors impact the outcomes and equitable utilisation of artificial intelligence (AI) technology, and what regulatory and governance mechanisms can be put in place to ensure fair and responsible AI deployment?*

### 3.2 Conceptualising the AI divide

The analysis of the SLR demonstrated that the AI divide can be conceptualised as illustrated in Figure 3. This figure provides an overview of the three levels of the AI divide and their interrelationships, forming the foundation of our conceptual model. In the next section, we will review the existing literature to examine the factors driving each level of the AI divide and present the complete conceptual model that integrates the findings from the SLR analysis with the identified drivers.



Figure 3. Conceptualising the AI divide based on the SLR Analysis.

#### 3.2.1 Analysing the AI Divide: Factors driving the AI divide

The profound impact of artificial intelligence (AI) is indisputable, permeating through various layers of society, from the individual to the global scale. As previously discussed, it is crucial to recognise that the benefits of AI are not distributed equally, giving rise to the AI divide. This divide encompasses disparities in accessing and utilising AI-enabled technology, as well as the resulting societal impacts. The complex interplay of socio-technical factors, as evidenced by the findings of the SLR, shapes this divide. Examining these factors informs the development of our conceptual model

and provides a narrative that is relevant to understanding the research gaps identified by the SLR.

The socio-technical perspective advocates that a comprehensive understanding of information systems (IS) phenomena necessitates the simultaneous consideration of both social and technical dimensions, recognising them as integral components within a multifaceted system (Carter et al., 2020).

The AI divide encompasses inequalities related to accessing and utilising AI-enabled technology, along with its consequential impacts. When conceptualising and analysing this divide, it is imperative to identify and assess the socio-technical factors that drive it. AI, as a practice, inherently exists at the crossroads of various disciplines, making it all the more essential to adopt a socio-technical perspective (Sartori & Theodorou, 2022).

Achieving a brighter future in AI requires more than just technical advancements in the algorithms themselves; it also involves designing interactions between AI and humans that consider the respective strengths and weaknesses of both parties (Zhou et al., 2023)

It is evident that much of the discourse surrounding AI, particularly within the realm of business research, has gravitated predominantly towards the technical aspects (Nguyen et al., 2022) assert that AI should not be seen as simple, self-contained technological systems; instead, they should be viewed as complex, open, sociotechnical entities intricately interconnected with material infrastructure, social dynamics, politics, and economics. Hence, it becomes increasingly important to not only delve into the intricacies of the technology but also to grasp the broader social implications of AI. By harmonising the social and technical aspects of AI, we can pave the way for a more equitable and ethically sound integration of AI into our lives. In doing so, we can transcend the AI divide and harness the full potential of artificial intelligence for the benefit of all.

### **3.2.1.1 Social Factors**

In the realm of digital inequality, contemporary analyses point to age, gender, ethnicity, labour, education, and regional disparities as pivotal factors (Van Dijk, 2020). These factors can be categorised into two dimensions: personal and positional. Personal

factors encompass individual characteristics such as age, language, ethnicity, and gender, while positional factors relate to socio-economic standing, including income, education, and occupational status. This classification is fundamental to understanding digital inequality.

Extending this notion to the sphere of artificial intelligence, it becomes evident that AI should strive for inclusivity and equity. The foundation for building AI systems must rest on high-quality data that considers gender, education, ethnicity, and other socio-economic variables that often underscore disparities (Sartori & Theodorou, 2022). As observed by (Wu, 2022), wealth disparities in China contribute to a considerable AI access divide. Lutz, (2019) reinforces this by asserting that individuals with higher income, education, and occupational prestige tend to have greater access to emerging technologies like AI, reaping more extensive economic and social benefits. Conversely, older individuals often face difficulties accessing and utilizing AI, as well as adapting to the rapid pace of technological evolution (Lutz, 2019).

The concern surrounding the potential of new technologies exacerbating existing inequalities becomes particularly pronounced in contexts where multi-layered inequality, including race, ethnicity, and social background, is prevalent. Rizk (2020) highlights that disparities rooted in these factors are often exacerbated by emerging technologies. This issue, especially in the realm of AI, has the potential to perpetuate structural biases and discrimination, with algorithmic bias manifesting in domains such as criminal justice, healthcare, and education (Lutz 2019).

Addressing bias and unfairness emerges as a central challenge in AI, as pointed out by (Sartori & Theodorou, 2022). The AI technical community predominantly focuses on calling for transparency, explainability, and accountability. However, the need for human control, encompassing requirement definition, design, and development methodologies, is paramount to engendering a fairer AI. It is essential to recognise AI technology as a social practice situated within specific institutional contexts.

Furthermore, linguistic diversity, essential to cultural diversity, is a pertinent concern, particularly in the digital realm, where approximately 17% of the world's languages, many of which belong to Africa, are designated as "low resource languages." These

languages face marginalisation in the digital landscape, including AI deployments that predominantly emanate from the Global North (Ade-Ibijola & Okonkwo, 2023).

At the organisational level, Enholm et al., (2022) assert that culture exerts influence over employees' willingness and capacity to interact with and create AI applications. An innovative culture can bolster the acquisition of AI skills crucial for effective AI use. As AI matures, social inequalities stemming from the AI divide will continue to surface, intertwined with cultural and ethical concerns that impact the design, adoption, and utilisation of AI systems across diverse contexts and among various stakeholders (Hangl et al., 2022)

Regarding gender disparities, it is a well-established fact that women often confront data and algorithm biases that both reflect and amplify pre-existing inequities. AI algorithms, predominantly developed in the Global North, frequently draw on datasets that inadequately represent the realities of African contexts, thereby excluding certain communities, including women. Furthermore, women remain underrepresented in STEM fields and in the development and design of emerging technologies, as emphasised by Lutz (2019).

In the realm of AI, trust and acceptance pose considerable challenges (Hangl et al., 2022). Employees, customers, and other stakeholders often perceive AI systems as black boxes, potential sources of bias, or threats. Resistance to change, stemming from concerns about job displacement, accompanies a reluctance to acquire new skills. AI systems and employees may also face cognitive, relational, and structural challenges during integration, including issues related to trust, learning, ethics, and job design (Makarius et al., 2020).

The lack of AI legislation in Africa is primarily due to policymakers' limited technology expertise. Technical experts, particularly in economic policymaking, are interested in AI, but there's a lack of digital leadership in politics and institutions (*Artificial Intelligence for Economic Policymaking The Frontier of Africa's Economic Transformation*, 2023). Regulations also impact inequalities, for example, data privacy and security laws, transparency, and accountability in algorithmic decision-making affect trust and fairness. The absence of regulation and gaps in understanding AI risks contribute to these challenges (White & Lidskog, 2022).

These multifaceted dynamics further underscore the intricate nature of digital inequality and AI's role in shaping it. A comprehensive understanding of these multifaceted issues is essential to fostering a more equitable and inclusive AI landscape.

### **3.2.1.2 Technical Factors**

The influence of artificial intelligence (AI) on digital inequality is a multifaceted phenomenon, where its effect can either amplify or alleviate disparities depending on the context of its application. Central to this is the role of data, which is often proprietary and challenging to access. This data, whether collected and owned by private companies with privileged access or obtained through costly private research involving sensitive information, can serve as an exclusionary barrier for startups, small and medium-sized enterprises (SMEs), universities, and research organisations seeking to advance in AI domains (Kitsara, 2022).

Furthermore, the performance of an AI system is inextricably linked to the quality and quantity of data it receives. AI, much like humans, learns from the data at its disposal but necessitates a significantly larger dataset to discern patterns than a human does. The accuracy of the data input directly impacts the precision of the outcomes. For instance, SMEs that extensively utilise data must have access to high-quality data and AI tools to uphold their credibility (Szedlak et al., 2020)

Moreover, the foundation of AI systems lies in infrastructure, which must be reliable, adaptable, and compatible with AI applications. Challenges in this realm encompass issues like system integration, interoperability, maintenance, and governance. Addressing data usability and access issues is crucial, necessitating improvements in data quality, privacy, security, and ownership policies and practices (Gwagwa et al., 2020) .

At the core of AI systems are their algorithms, which demand accuracy, efficiency, robustness, and explainability for effective application. Algorithmic challenges include selecting appropriate algorithms, optimising them, validating their performance, testing, and ensuring interpretability (Lutz, 2019). The design and development of AI systems and algorithms can significantly impact digital inequalities. Choices regarding data sources, methodologies, and criteria in AI application development can

inadvertently introduce biases and errors, potentially disadvantaging specific groups (Hangl et al., 2022).

Furthermore, the limited state of internet penetration across the African continent can be attributed to infrastructure issues, characterised by inadequate access to electricity and limited investments in internet infrastructure such as fibre-optic cables, cell towers, and base stations. The World Bank reports that 80% of urban populations in sub-Saharan Africa have access to electricity, compared to a mere 28% in rural areas, highlighting the stark urban-rural divide in access and the challenges to effective adoption of AI (Gwagwa et al., 2020). Lutz (2019) further contends that access and utilisation of AI are contingent on the availability and quality of digital infrastructure, encompassing broadband networks, cloud computing, and data centres. These infrastructure resources are often unevenly distributed across regions and countries, leading to geographical disparities in digital opportunities and challenges, further perpetuating digital inequalities.

### **3.2.1.3 Socio Technical Factors**

The digital skills divide, which has long been a contributing factor to the overall digital divide, plays a crucial role in shaping the AI divide. According to (Van Dijk, 2020), as the digital divide evolves, disparities in technology-related skills significantly contribute to inequalities in technology usage. The emergence of algorithmic skills further exacerbates inequality in accessing and utilising AI technology.

Significantly, digital skills literacy stands as a major hurdle to the adoption and integration of artificial intelligence in Africa. Sub-Saharan Africa lags all other world regions, with the lowest percentage of citizens possessing digital skills, representing approximately half of the global by average (Madden P & Kanos D, 2020).

Considering these socio-technical factors, which encompass both social and technical dimensions, it becomes evident that digital skills and AI skills exert a substantial influence on the AI divide. Addressing these skill disparities is pivotal to ensuring a more equitable AI landscape. Hence, we present the three levels of the AI divide along with the socio-technical factors that underpin and drive this division (see Figure 4).

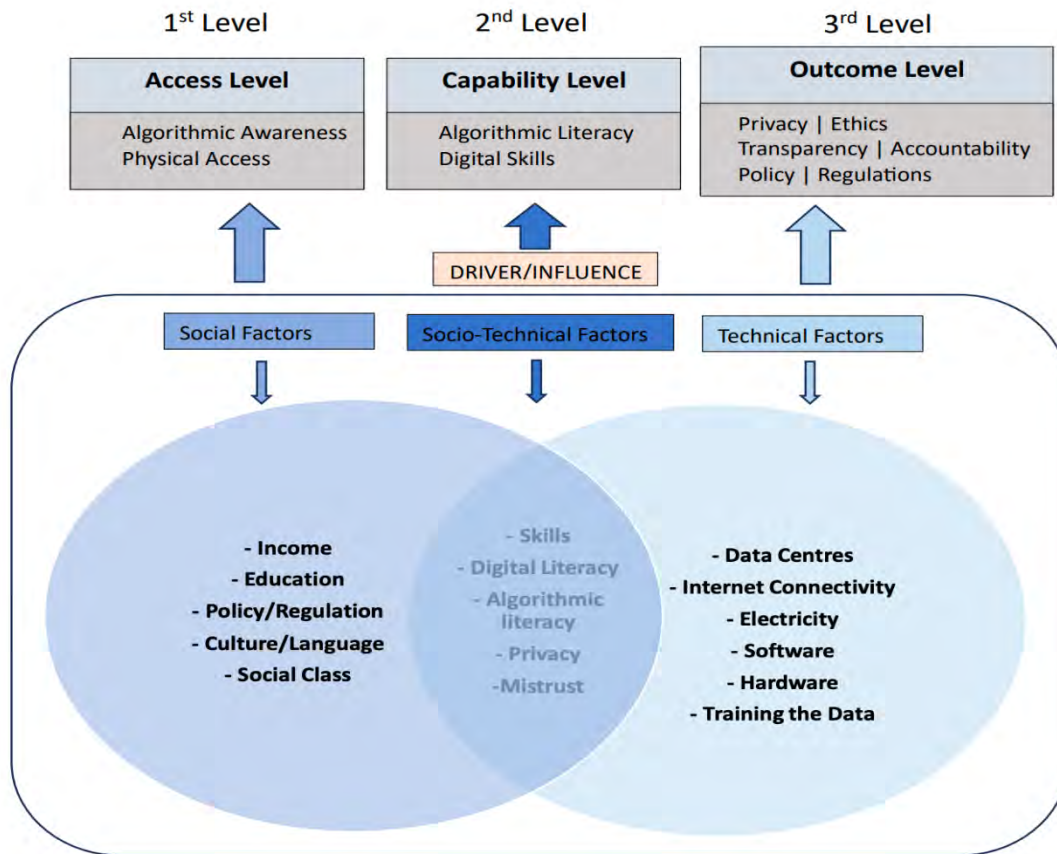


Figure 4. Conceptual Model of the Socio-technical Implications of the AI divide.

## 4.0 Conclusion

### 4.1 Summary of Findings

The synthesis reveals that individuals and groups lacking a fundamental understanding of AI or facing barriers to accessing the necessary hardware are more likely to experience disparities in AI adoption. While studies emphasise the existing algorithmic knowledge gap within the populace, they do not delve into the context-specific nature of these studies. The authors have identified vulnerable demographic cohorts, but they have not explored the social factors, such as culture, contributing to this gap. Studies on algorithmic awareness have a limited scope, and the conclusions derived from these studies are not entirely conclusive, as they are conducted within specific mediated contexts, making it difficult to generalise their findings.

While acknowledging the potential emergence of an algorithmic awareness divide, researchers have not delved deeply into the specific consequences of this divide. For instance, the impact of a lack of algorithmic awareness on an individual's ability to



make informed decisions or its effects on their participation in the workplace, particularly in an AI-driven job market, remains underexplored.

The authors provide critical insights into algorithmic awareness and physical access in the context of the AI divide. However, access to AI technologies is mainly enjoyed by large organisations, necessitating the democratisation of access to AI technologies to enable small businesses that may not have the financial means to acquire AI technology. Democratising access to AI is the best way to ensure equitable access to AI technologies, as even though small businesses may have access, their limited finances may still hinder their ability to afford AI technology.

At the second level of the AI divide, it becomes evident that AI literacy and algorithmic expertise play a crucial role in exacerbating this divide. The lack of skills at the outcome level hampers individuals' ability to participate in an AI-driven job market, make informed decisions about AI, and fully realise the potential benefits of AI. The authors successfully establish algorithmic literacy as a crucial skill in the era of AI. However, it would be beneficial to extend this discussion by addressing the potential consequences of lacking algorithmic literacy. For instance, how does the lack of algorithmic literacy impact an individual's ability to make informed decisions in the digital age?

The third level of the AI divide revolves around the outcomes derived from AI usage. The authors have identified several factors contributing to this stage, including algorithmic biases in AI systems that can perpetuate existing inequalities and lead to unequal outcomes. The authors focus on the concept of bias as a predisposition or prejudice in AI systems, and it is essential to emphasise that not all bias is intentional; many biases are unintentional and may result from historical data or algorithm design decisions. Addressing unintended bias is a critical aspect of mitigating the AI divide. While the authors emphasise the importance of addressing bias, they do not provide an extensive strategy to mitigate bias in AI systems. Exploring best practices, regulations, and ethical frameworks for reducing bias is important.

The authors have emphasised that AI bias has the potential to exacerbate the existing digital divide. AI can either contribute to reducing or increasing these disparities, and it is vital to consider how bias plays a role in this process.

Although the authors identify the idea that fairness is a subjective decision made by programmers, by expanding on this concept, they could explore how cultural, social, and individual factors influence these subjective notions of fairness and how they might lead to bias in AI systems. While the authors highlight the need for transparency and accountability, the concept of algorithmic auditing could be used to address fairness. Auditing AI systems involves systematically examining them for bias and unfairness.

Finally, regulations and laws have a significant impact at this level, shaping the degree to which artificial intelligence is utilised for the benefit of society or widening existing inequalities. Overall, the focus of AI implementation in business research has predominantly leaned towards the technical aspects of the spectrum.

#### **4.2 Literature Gaps and Future Research Direction**

This study adopts a socio-technical perspective to explore the AI divide and its societal implications. The focus has been on uncovering the unintended divisions arising from artificial intelligence. Despite extensive research and synthesis, certain gaps remain, which is unsurprising given the emerging nature of AI, especially in developing nations grappling with persistent technological and digital disparities.

One notable gap is the limited attention given to the context of developing economies. There is a lack of research exploring the specific measures required to create a conducive ecosystem for AI adoption. There is need for research on the "extent and impact of the AI awareness gap" or studies quantifying access disparities in that context. While studies identify digital and AI skills within the workforce, further investigation is needed to understand how the absence of AI skills affects their readiness to integrate AI into business processes.

Most studies have not utilised primary evidence to derive their findings, which suggests an avenue for future research. Also, exploring different thematic perspectives for dataset analysis, such as geographical or chronological dimensions, can offer fresh insights.

Addressing these dimensions will advance our understanding of the AI divide and contribute to a more equitable AI-driven world.

### 4.3 Final Remarks

In conclusion, this study conducted a comprehensive systematic literature review to examine the notion of the AI divide, encompassing disparities in access to AI technologies. The research employed the well-established three levels of the digital divide as a framework for analysing the results of this review. As an outcome of this investigation, a conceptual model has been constructed, providing a valuable foundation for future research endeavours. Moreover, this study successfully identified research gaps, which should guide and inspire further investigations into the AI divide, contributing to a deeper understanding of this critical area

### References

- Ade-Ibijola, A., & Okonkwo, C. (2023). *Artificial Intelligence in Africa: Emerging Challenges*. In *Responsible AI in Africa: Challenges and Opportunities*. Cham: Springer International Publishing.
- Al-Surmi, A., Bashiri, M., & Koliouisis, I. (2022). AI based decision making: combining strategies to improve operational performance. *International Journal of Production Research*, 60(14), 4464–4486.
- Artificial Intelligence for Economic Policymaking The Frontier of Africa's Economic Transformation*. (2023).
- Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for Artificial Intelligence and Digital technologies. In *International Journal of Information Management* (Vol. 62). Elsevier Ltd. <https://doi.org/10.1016/j.ijinfomgt.2021.102433>
- Bjola, C. (2022). AI for development: Implications for theory and practice. *Oxford Development Studies*, 50(1), 78–90.
- Borges, A. F., Laurindo, F. J., Spínola, M. M., Gonçalves, R. F., & Mattos, C. A. (2021). The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. *International Journal of Information Management*, 57, 102225.
- Cameron, L., Lamers, L., Leicht-Deobald, U., Lutz, C., Meijerink, J., & Möhlmann, M. (2023). Algorithmic Management: Its Implications for Information Systems Research. *Communications of the Association for Information Systems*, 52, 518–537. <https://doi.org/10.17705/1CAIS.05221>
- Carter, L., Liu, D., & Cantrell, C. (2020). Exploring the Intersection of the Digital Divide and Artificial Intelligence: A Hermeneutic Literature Review. *AIS*

- Transactions on Human-Computer Interaction*, 12(4), 253–275.  
<https://doi.org/10.17705/1thci.00138>
- Chowdhury, S., Budhwar, P., Dey, P. K., Joel-Edgar, S., & Abadie, A. (2022). AI-employee collaboration and business performance: Integrating knowledge-based view, socio-technical systems and organisational socialisation framework. *Journal of Business Research*, 144, 31–49. <https://doi.org/10.1016/j.jbusres.2022.01.069>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- Elliott A. (2019). *The culture of AI: Everyday life and the digital revolution*. Routledge.
- Enholt, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2022). Artificial intelligence and business value: A literature review. *Information Systems Frontiers*, 24(5), 1079–1734.
- Fink, A. (2019). *Conducting Research Literature Reviews: From the internet to paper*. Sage Publications.
- Gao, L., Li, G., Tsai, F., Gao, C., Zhu, M., & Qu, X. (2023). The impact of artificial intelligence stimuli on customer engagement and value co-creation: The moderating role of customer ability readiness. *Journal of Research in Interactive Marketing*, 17(2), 317–333.
- Gmyrek P, Berg J, & Bescond D. (2023). *Generative AI and jobs: A global analysis of potential effects on job quantity and quality*.
- Gwagwa, A., Kraemer-Mbula, E., Rizk, N. , Rutenberg, I. , & De Beer, J. (2020). Artificial Intelligence (AI) deployments in Africa: benefits, challenges and policy dimensions. *The African Journal of Information and Communication*, 28, 1–28.
- Haber, E. (n.d.). Algorithmic Inclusion' (2020-2023) 72 Florida Law Review Forum 94  
MLA 9th ed. Haber, Eldar. In *Florida Law Review Forum* (Vol. 94).  
<http://hdl.handle.net/10625/56949>.
- Hall, P., & Ellis, D. (n.d.). *A systematic review of socio-technical gender bias in AI algorithms*. <https://doi.org/10.1108/OIR>
- Hanelt, A., Bohnsack, R., Marz, D., & Antunes Marante, C. (2021). A systematic review of the literature on digital transformation: Insights and implications for

- strategy and organizational change. . *Journal of Management Studies*, 58(5), 1159–1197.
- Hangl, J., Behrens, V. J., & Krause, S. (2022). Barriers, Drivers, and Social Considerations for AI Adoption in Supply Chain Management: A Tertiary Study. *Logistics*, 6(3), 63. <https://doi.org/10.3390/logistics6030063>
- Hardt, M., Price, E., & Srebro, N. (2016). Equality of opportunity in supervised learning. *Advances in Neural Information Processing Systems* , 3315–3323.
- Hargittai, E. (2002). Second-Level Digital Divide: Differences in People’s Online Skills. . *First Monday*, 7(4).
- Heeks, R. (2022). Digital inequality beyond the digital divide: conceptualizing adverse digital incorporation in the global South. *Information Technology for Development*, 28(4), 688–704.
- Hilbert, M. (2011). The end justifies the definition: The manifold outlooks on the digital divide and their practical usefulness for policy-making. . *Telecommunications Policy*, 35(8), 715–736.
- Johnson, K. N., & Reyes, C. L. (2021). Exploring the Implications of Artificial Intelligence’ (2021) 8(2) *Journal of International and Comparative Law* 315 MLA. In *Journal of International and Comparative Law* (Vol. 8, Issue 2). <https://www.cdc.gov/media/releases/2020/p0121-novel-coronavirus-travel-case.html>
- Kitchenham, B. A., Budgen, D. , & Brereton, O. P. (2011). Using mapping studies as the basis for further research—a participant-observer case study. *Information and Software Technology*, 53(6), 638–651.
- Kitsara, I. (2022). Artificial intelligence and the digital divide: From an innovation perspective. In *Platforms and Artificial Intelligence: The Next Generation of Competences*, 245–265.
- Klaus Schwab. (2020). *Annual Report*.
- Kusters, R., Misevic, D., Berry, H., Cully, A., Le Cunff, Y., Dandoy, L., Díaz-Rodríguez, N., Ficher, M., Grizou, J., Othmani, A., Palpanas, T., Komorowski, M., Loiseau, P., Moulin Frier, C., Nanini, S., Quercia, D., Sebag, M., Soulié Fogelman, F., Taleb, S., ... Wehbi, F. (2020). Interdisciplinary Research in Artificial Intelligence: Challenges and Opportunities. *Frontiers in Big Data*, 3. <https://doi.org/10.3389/fdata.2020.577974>

- Lauterbach, A. (2019). Artificial intelligence and policy: quo vadis? *Digital Policy, Regulation and Governance*, 21(3), 238–263. <https://doi.org/10.1108/DPRG-09-2018-0054>
- Lewicki, K., Lee, M. S. A., Cobbe, J., & Singh, J. (2023). *Out of Context: Investigating the Bias and Fairness Concerns of “Artificial Intelligence as a Service.”* <https://doi.org/10.1145/3544548.3581463>
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of Internal Medicine*, 151(4), W-65.
- Lutz, C. (2019). Digital inequalities in the age of artificial intelligence and big data. *Human Behavior and Emerging Technologies*, 1(2), 141–148. <https://doi.org/10.1002/hbe2.140>
- Lythreathis, S. , Singh, S. K. , & El-Kassar, A. N. (2022). The digital divide: A review and future research agenda. *Technological Forecasting and Social Change*, 175, 121359.
- Madden P, & Kanos D. (2020). *Figures of the week: digital skills and the future of work in Africa*. Brookings Institution. <https://www.brookings.edu/articles/figures-of-the-week-digital-skills-and-the-future-of-work-in-africa/>
- Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. (2020). Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research*, 120, 262–273. <https://doi.org/10.1016/j.jbusres.2020.07.045>
- Mbuy, S. P., & Ortolani, M. (2022, July). *Algorithmic Impact Assessment for an Ethical Use of AI in SMEs*. <https://doi.org/10.14236/ewic/hci2022.34>
- McKinsey Global Report. (2023). *The economic potential of generative AI: The next productivity frontier*.
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A Survey on Bias and Fairness in Machine Learning. In *ACM Computing Surveys* (Vol. 54, Issue 6). Association for Computing Machinery. <https://doi.org/10.1145/3457607>
- Mitchell, S., Potash, E., Barocas, S., D’Amour, A., & Lum, K. (2021). Algorithmic fairness: Choices, assumptions, and definitions. *Annual Review of Statistics and Its Application*, 8, 141–163.

- Munoko, I., Brown-Libur, H. L., & Vasarhelyi, M. (2020). The Ethical Implications of Using Artificial Intelligence in Auditing. *Journal of Business Ethics*, 167(2), 209–234. <https://doi.org/10.1007/s10551-019-04407-1>
- Nguyen, Q. N., Sidorova, A., & Torres, R. (2022). Artificial Intelligence in Business: A Literature Review and Research Agenda. *Communications of the Association for Information Systems*, 50(1), 175–207. <https://doi.org/10.17705/1CAIS.05007>
- Ntoutsis, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdil, W., Vidal, M. E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., ... Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3). <https://doi.org/10.1002/widm.1356>
- OECD. (2001). *Understanding the Digital Divide*. <https://www-oecd-org.salford.idm.oclc.org/sti/1888451.pdf>
- OECD. (2019). *ARTIFICIAL INTELLIGENCE & RESPONSIBLE BUSINESS CONDUCT*.
- Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems*, 37.
- Page, M. J. , McKenzie, J. E. , Bossuyt, P. M., Hoffmann, T. C., Mulrow, C. D. , & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery*, 88, 105906.
- Paré, G. , Trudel, M. C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183–199.
- Petticrew, M., & Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. Wiley & Sons.
- Ragnedda, M., & Muschert, G. W. (2017). *Theorizing Digital Divides* (1st ed.). Routledge.
- Rogers, E. M. (1962). *Diffusion of Innovations*. Free Press of Glencoe.
- Sartori, L., & Theodorou, A. (2022). A sociotechnical perspective for the future of AI: narratives, inequalities, and human control. *Ethics and Information Technology*, 24(1). <https://doi.org/10.1007/s10676-022-09624-3>

- Scheerder, A., Van Deursen, A., & Van Dijk, J. (2017). Determinants of Internet skills, uses and outcomes. A systematic review of the second-and third-level digital divide. *Telematics and Informatics*, *34*(8), 1607-1624.
- Szedlak, C., Poetters, P. , & Leyendecker, B. (2020). Application of artificial intelligence in small and medium-sized enterprises. *Proceedings of the International Conference on Industrial Engineering and Operations Management*.
- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, *14*(3), 207–222.
- Turner Lee, N. (2018). Detecting racial bias in algorithms and machine learning. *Journal of Information, Communication and Ethics in Society*, *16*(3), 252–260. <https://doi.org/10.1108/JICES-06-2018-0056>
- van Bruxvoort, X., & van Keulen, M. (2021). Framework for assessing ethical aspects of algorithms and their encompassing socio-technical system. *Applied Sciences (Switzerland)*, *11*(23). <https://doi.org/10.3390/app112311187>
- Van Deursen, A. J. , & Helsper, E. J. (2015). The third-level digital divide: Who benefits most from being online?. *Communication and Information Technologies Annual*, 29–52.
- Van Dijk, J. (2020). *The Digital Divide*. Polity Press.
- Van Dijk, J. A. (2006). Digital divide research, achievements and shortcomings. *Poetics*, *34*(4–5), 221–235.
- Vesnic-Alujevic, L., Nascimento, S., & Pólvara, A. (2020). Societal and ethical impacts of artificial intelligence: Critical notes on European policy frameworks. *Telecommunications Policy*, *44*(6). <https://doi.org/10.1016/j.telpol.2020.101961>
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, xiii-xxiii.
- Wei, K. K., Teo, H. H., Chan, H. C., & Tan, B. C. (2011). Conceptualizing and testing a social cognitive model of the digital divide. *Information Systems Research*, *22*(1), 170–187.
- White, J. M., & Lidskog, R. (2022). Ignorance and the regulation of artificial intelligence. *Journal of Risk Research*, *25*(4), 488–500. <https://doi.org/10.1080/13669877.2021.1957985>
- Wu, Y. (2022). An Overview Analysis of AI Divide: Applications and Prospects of AI Divide in China’s Society. *3rd International Conference on Mental Health, Education and Human Development*, 3–9.



- Yu, L. (2006). Understanding information inequality: Making sense of the literature of the information and digital divides. *Journal of Librarianship and Information Science*, 38(4), 229–252.
- Yu, P. K. (n.d.). *THE ALGORITHMIC DIVIDE AND EQUALITY IN THE AGE OF ARTIFICIAL INTELLIGENCE*. <https://www.nitrd>.
- Yu, P. K. (2020). *CAN ALGORITHMS PROMOTE FAIR USE?* <https://ssrn.com/abstract=3374951>.
- Yu, X., Xu, S., & Ashton, M. (2023). Antecedents and outcomes of artificial intelligence adoption and application in the workplace: the socio-technical system theory perspective. *Information Technology and People*, 36(1), 454–474. <https://doi.org/10.1108/ITP-04-2021-0254>
- Zarouali, B. , Helberger, N. , & De Vreese, C. H. (2021). Investigating algorithmic misconceptions in a media context: Source of a new digital divide?. *Media and Communication*, 9(4), 134–144.
- Zhou, L., Rudin, C., Gombolay, M., Spohrer, J., Zhou, M. , & Paul, S. (2023). From Artificial Intelligence (AI) to Intelligence Augmentation (IA): Design Principles, Potential Risks, and Emerging Issues. *AIS Transactions on Human-Computer Interaction*, 15(1), 111–135.

# The Underlying Practices of Digital Transformation Leadership: Theorising the Practitioner Voice

## Abstract

*Companies don't transform, people transform companies. However, there is still a pressing need to understand 'what' action a Digital Transformation (DT) leader needs to take and 'how' they should enable that action, in order to achieve the best possible outcome in a DT initiative. Therefore, this paper explores the underlying practices associated with DT leadership in the context of pre-digital organisations. We set about understanding these underlying practices in a unique way, using a grounded approach, analysing sixteen key informant's accounts of the 'what' and 'how' of DT leadership. In total we identify six underlying practices (collaborative change, digital influence, collaborative tooling, employee spirit, prioritised platformisation, democratising data) that impact on the outcome of a DT initiative within a pre-digital organisation. We believe that our approach strengthens the relevance of our research outputs for practitioners, where the practitioner voices and their lexicon are central to the theorising and the outputs produced.*

**Keywords:** Digital Transformation, Leadership, Practices, Grounded Approach, Key Informant

## Introduction

Over the past number of decades IT-enabled transformation has significantly influenced the shape of organisational transformation processes. The emergence of key enabling technologies, such as *Mainframe System Architectures* (1960s/1970s), *End User Computing* (PCs 1980s), *Enterprise Architecture* (Client/Server 1990s), and *Service Oriented Architecture* (ERP & Data Warehousing 2000s) laid the foundation for enterprise transitioning, to become a *digitalized organisation* over the past two decades (2010s & 2020s). Therefore, charting the history of these IT-enabled transformation ‘siloes’ provides an appreciation that current Digital Transformation (DT) initiatives are indeed a more comprehensive and integrated recipe using those emergent ingredients of the previous five decades. Furthermore, the key differentiator between Digital Transformation and IT-enabled Transformation is the nature of the organisational change (enterprise-wide and at both strategic & operational levels), and how digital technologies, value propositions, and organisation identity interrelate during the process (c.f. Verhoef et al., 2021; Vial, 2019; Yoo, 2013; Besson & Rowe, 2012).

Digital Transformation (DT) has generated much research and curiosity in recent years from both an academic and practitioner perspective, not least in Information Systems (IS) research. Despite this growing attention around DT, several gaps still exist in our understanding of this complex process (c.f. Vial 2019). As a result, several calls for further research have been made, in areas such as: embedding and sustaining (normalising) a DT (Carroll, 2020), organisational readiness for DT (Nguyen et al., 2021), role of middle management in DT (Nadkarni & Prügl, 2021), leader attributes for a successful DT (Pabst von Ohain, 2019) and prioritising practitioner activities throughout DT phases (Berghaus & Back, 2017). For the purposes of this study, we define DT as the strategic enhancement of an organisation to be customer centric, culturally aware, and innovative, by leveraging emergent and emerging technologies to improve processes, data, and people to disrupt business and operational models, to deliver value propositions to all organisational stakeholders.

Emerging scholarly attention positions DT as a “*leading technology-related phenomenon*” (Wessel et al., 2021 p.102). However, McCarthy et al. (2021) suggest that there is currently a relatively small number of empirical research outputs focusing on DT leadership, based on their analysis of 87 empirical studies (from 93 top ranked ‘information management’ journals and 8 major AIS conferences). The emergence of new digital leadership roles, e.g. the Chief Digital Officer (CDO), is highlighted as being significant (c.f. Haffke et al., 2017; Haffke et al., 2016; Horlacher & Hess, 2016; Singh & Hess, 2017). So while achieving DT success is linked to having certain digital-savvy or *transformational leaders* (c.f. Paavola et al., 2017; Ready et al 2020) in place, less than one-third of organisations have engaged a CDO to support their transformations (c.f. McKinsey, 2018; Wade & Shan, 2020; Barthel et al., 2020; Wade et al., 2017). Notwithstanding this, the emergence of the CDO represents the widespread view of the need to appoint a specialist to take charge of digitally transforming the business (Haffke et al., 2016; Singh & Hess, 2017).

Despite the growing volume of academic research, it is still hard to find comprehensive coverage of the underlying practices of DT leadership (even in the trade press) that are linked to “*what*” action a DT leader needs to take and “*how*” they enable that action, where they are striving for the best possible DT initiative outcome. This takes on a “must know” significance for the IS field when we consider that DT is a complex and multidimensional phenomenon (c.f. Porfirio et al., 2021; Tabrizi et al., 2019), and the DT process “*is not well understood*” within an IS context (Carroll, 2020 p.1). For many organisations, DT begins with trying to identify the ‘what’ and ‘how’; in other words, being able to understand what is required and consequently how to implement those requirements (Ure, 2018). As suggested by McCarthy et al (2021, p.28) “*the leadership required to lead a [DT] programme is perhaps greater than is anticipated, simply because, in many cases, the volume of changes within the business is unprecedented*”. Therefore, the objective of this research is to explore the underlying practices associated with DT leadership, specifically in the context of pre-digital organisations. To fulfil this objective, we pose the following **Research Question:** *What are the underlying practices associated with Digital Transformation (DT) Leadership that impact on the outcome of a DT initiative within a pre-digital organisation?* To answer this research question, we follow a theory building research strategy to develop an understanding of DT leadership from

those practitioners currently “in the DT trenches”. This approach affords us the opportunity to ‘*capture the meaning*’ from those practitioners ‘*living the experience*’ (leading a DT initiative) and ‘*theorise about that experience*’ (Gioia et al., 2012 p.26).

The remainder of this paper is structured as follows. In the next section we present the research approach being followed. This is followed by a presentation of our observations (six underlying practices of DT leadership) gained through coding sixteen key informant interviews (those operating as DT leaders in their respective organisations). The paper concludes with the implications for theory and practice, along with opportunities for further research.

In the next section we present a detailed description of our research approach to building theory.

### **Methodology: Data Gathering & Data Analysis**

For the purposes of this research we follow a theory building research strategy where our ambition is to build theory, and in so doing, we embrace an approach aligned with “*concept development*” as opposed to “*construct elaboration*”, where a concept captures “*qualities that describe or explain a phenomenon of theoretical interest*”, (Gioia et al., 2012 p.16). Therefore, being inspired by features of the Gioia Methodology, which is positioned as a “*systematic inductive approach to concept development*” and assumes that “*the organisational world is socially constructed*”(Gioia et al., 2012 p.17), we aim to conceptualise the practitioner voice and not “*substitute practitioners’ understandings for theory*” (Markus & Rowe, 2021 p.273). As a result, in data collection there is a need to “*give extraordinary voice to informants, who are treated as knowledgeable agents*”; while in data analysis there is a need to maintain “*the integrity of 1<sup>st</sup> order (informant-centric) terms*” during initial data coding, and further “*organise 1<sup>st</sup>-order codes into 2<sup>nd</sup>-order (theory-centric) themes*” (Gioia et al., 2012 p.26). To conclude, “*without research outcomes relevant to practice, the very existence of a research discipline could be questioned because the discipline could well lack impact beyond its own (academic) community*” (Rosemann & Vessey, 2008 p.3).

To answer our research question, we select sixteen key informants based on their organisational perspective (Business or IT) and role (Strategic or Operational). This stratified selection of key informants afforded us the opportunity to “*capture the consonance (or dissonance) between plans [strategic] and their implementation [operational]*” (Day et al., 2009 p.641), while also appreciating the alignment between IT and business perspectives (c.f. Bendig et al., 2022; Yeow et al., 2018) because the “*impact of DT*” on the “*business*” is “*technology-enabled*” (Porfirio et al., 2021 p.616). In fact, (Smith & Watson, 2019 p.98), in using the metaphor of a “*tapestry*” and its “*weavers*” of the “*threads*” explicitly refer to the “*business thread*” and the “*IT thread*” of DT (the *digital tapestry*). The literature also reminds us of the importance of a well-functioning and collaborative strategic partnership between IT and business leadership for the purpose of change adaptations throughout the DT process (c.f. Singh & Hess, 2017; Matt et al., 2015; Hess et al., 2016; Bharadwaj et al., 2013) . In short, in order to design a new digital experience and an improved portfolio of digital offerings to change the way the organization engages with customers, organizations will require the expertise of both IT and business personnel to operate in partnership (*reference withheld for review purposes*).

These key informants are considered DT leaders within their respective organisations (e.g. CEO/CIO/CTO, Business Transformation Director/Manager, Lead Digital Business Analyst) and their voices reflect those of their industry peers. **These key informants were recruited through several means, these were (i) having prior knowledge of, and working relationships with, practitioners currently active in a DT programme, (ii) speakers identified at practitioner conferences and webinars with a history of working in DT programmes, and (iii) LinkedIn connections.** On average these key informants have 15+ years of industry experience in business/IT transformation (within single and multiple DT initiatives in MNC/SME environments across a variety of sectors, e.g. Technology [6], Higher Education [4], Energy [2], Agriculture [2], and Healthcare [1]) **(see Appendix A).** Our approach to key informant selection allows for four types (quadrants) of practitioner voices to be heard (e.g. *Business Strategic, Business Operational, IT Strategic, IT Operational*) as we theorise about the underlying practices of DT leadership.

Interviews are considered the most appropriate data gathering technique for collecting rich and detailed data from industry experts and are a typical data gathering technique with the key informant approach (Barker et al., 2005; Whittaker, 2012). When using semi-structured interviews as part of the key informant technique, it is not uncommon to have a smaller number of interviewees; this can range from 6 interviewees (c.f. Flores & Ekstedt, 2012) to 32 interviewees (c.f. Benova et al., 2019). In this study, we conduct a series of semi-structured interviews (four per practitioner voice type), where each key informant reveals their unique DT leadership experiences. Interviews took place over sixteen months (between November 2018 and February 2020) and ranged in duration from 35 to 75 minutes with an average interview duration of 60 minutes. It is also worth mentioning that the sixteen key informants are affiliated with organisations “born in the pre-digital age” and they are conscious that they are balancing “tensions between the ‘old’ and the ‘new’” when transforming (Oberlander et al., 2021 p.1).

The emphasis of qualitative data analysis is on “sense making” (Bhattacharjee et al., 2012) and coding is one of the techniques widely used in analysing qualitative data in order to build theory (Hund et al., 2021; Buchwald et al., 2014). In this research we follow an inductive open, axial and selective coding approach, where these coding techniques aim to generate concepts from field data (Walsham, 2006). Therefore, *open coding* is a process that aims to identify the concepts or key ideas that are hidden within data that are likely to be related to the phenomenon of interest (Bhattacharjee et al., 2012) and concepts that appear to be similar are grouped together under a higher-order, more abstract concept called a category (Corbin & Strauss, 1990). The second reading of the data is considered during *axial coding* (Dezdar & Sulaiman, 2009), which is performed simultaneously with open coding (Bhattacharjee et al., 2012); (Corbin & Strauss, 1990). During this stage, where the researcher thinks systematically about the data in order to relate them, the categories are refined in order to be linked in the form of relationships (Alhassan et al., 2019). Finally, *selective coding* begins when researchers identify a potential core category (Tan et al., 2015), focusing then on the core categories and related categories that accrued in the axial coding. This involves comparing the core categories with the raw data by telling the story of the core categories that emerge (Corbin & Strauss, 1990). Therefore, “what coding does, above all, is to allow the researcher to communicate and connect

*with the data to facilitate the comprehension of the emerging phenomena and to generate theory grounded in the data”* (Basit, 2003 p.143).

For this research, after preparing each of the sixteen key informant interview transcripts (as the interviews were completed throughout the 16-month data gathering period), the data analysis commenced by reading each transcript sentence-by-sentence and following an inductive open coding approach. During *open coding* we were initially looking for two sides of a key informant’s DTL experience, namely the “what” and the “how”. This simply translates as “what” action they need to take and “how” they enable that action, in their role as a DT leader. These actions are in the context of the key informant striving for the best possible outcome in a DT initiative. The output from our open coding produced 28 categories. After coding the first two interviews, *axial coding* (the second reading of the data) was commenced in an iterative manner (as categories started to emerge) to identify and clarify the relationships (specific examples of ‘causal conditions’, ‘actions/interactions’, and ‘consequences’) between the emerging categories. See Table 2 for a sample of our coding. Therefore, the axial coding procedure resulted in the creation of a relationship matrix (see Table 5) to record the relationships between the emerging categories throughout the coding process. During this iterative process, similarities and differences among the categories emerged thus reducing the number of categories from 28 to 10, while also highlighting relationships between these categories. Therefore, the output from our axial coding produced 10 categories (emerging from 558 key informant excerpts) and 45 relationships between these categories (see Figure 1). Throughout this iterative coding process, the researchers were also looking to identify a core category along with its related categories that emerged from axial coding. This *selective coding* allows us to tell a compelling theorising story around the underlying practices associated with DT leadership in the context of pre-digital organisations.



KI	Key Informant Coded Excerpt	Concept	Category	Relationship
BS	<i>you need to underline what you wish to achieve as part of your strategy and be able to distil that message so people can understand and latch on to it.</i>	execute the digital strategy and vision using digital capabilities	C1: communicating and executing a digital strategy and vision by leveraging executive management support	Democratising Data
ITS	<i>be transparent in how the strategy will be delivered by the underpinning of the digital capabilities available to the organisation</i>			
BO	<i>to concentrate on of what parts of the organisation are suitable for transforming and how transforming can be achieved by using digital capabilities</i>			
ITO	<i>to understand the strategy as to what they want to transform and how they feel transforming helps everybody and grows the company</i>			
BS	<i>need to start with the data strategy as the first stage on the data journey and then examine how it's being captured and analysed, ensuring that it delivers what is expected from it</i>	capture and analyse data for decision making	C8: unlocking the value of data-driven decisions by capturing and analysing high quality data	
ITS	<i>focus on identifying suitable tools such as dashboards, reports, etc, wanting real time data has to be informative and gives them what they want</i>			
BO	<i>once the data is captured, how are we measuring really what's going on, so that we can inform leadership, which they need to support their decision making</i>			
ITO	<i>using digital systems and digital platforms to capture the data, do a deep dive analysis and extract the data properly for usage</i>			
<i>Business Strategic (BS), IT Strategic (ITS), Business Operational (BO), IT Operational (ITO)</i>				

Table 2: Sample Coding for an Underlying Practice (Democratising Data)

Number	Category		Excerpt	Concept	Practitioner Voice Quadrant			
	What	How			IT Strategic	IT Operational	Business Strategic	Business Operational
C1	communicating and executing a digital strategy and vision	by leveraging executive management support	101	20	26	24	28	23
C2	prioritising the customer value proposition	by implementing an integrated digital platform	76	16	21	20	18	17
C3	understanding the journey of organisational change	by embracing digital disruption	62	26	16	14	17	15
C4	inspiring the organisation to change	by adopting an open culture and digital mindset	60	18	16	10	18	16
C5	underpinning the organisational change	by using appropriate digital capabilities	58	15	22	18	10	8
C6	collaborating cross functionally	by adopting a disruptive approach to innovation	55	16	12	16	15	12
C7	redefining the business model	by optimising functionally aligned processes	48	17	14	12	12	10
C8	unlocking the value of data-driven decisions	by capturing and analysing high quality data	44	15	15	8	14	7
C9	realising value creation	by balancing cost reduction and revenue generation	29	10	13	8	5	3
C10	empowering employee experience	by creating a dynamic digital workplace	25	12	5	3	6	11
Total			558	165	160	133	143	122

*Table 3: The DT Leadership Characteristics Frequency across the four Practitioner Voice Quadrant*

Number	Category		Description
	What	How	
C1	communicating and executing a digital strategy and vision	by leveraging executive management support	creating a digital strategy and vision that is transparent to all and communicated from the strategic level to the operational level
C2	prioritising the customer value proposition	by implementing an integrated digital platform	putting the customer at the core of an organisation's digital strategy, so as to ensure greater value to the customer, through the changes that are implemented
C3	understanding the journey of organisational change	by embracing digital disruption	undergoing a digitally enabled organisational change involves having a clear understanding of the 'why', 'what' and 'how' of digitally transforming
C4	inspiring the organisation to change	by adopting an open culture and digital mindset	understanding the culture of the organisation is crucial to promote the necessary cultural shift in the organisation in order to make it 'culturally fit' for transforming digitally
C5	underpinning the organisational change	by using appropriate digital capabilities	creating a well architected digital platform will provide the foundation to deploy digital services across the organisation
C6	collaborating cross functionally	by adopting a disruptive approach to innovation	getting organisations to collaborate, both functionally and cross functionally, involves using collaborative technologies, human resources, and innovative methods
C7	redefining the business model	by optimising functionally aligned processes	enhancement and optimisation of business processes are essential in redesigning an organisation's business model (the remodelling of how the "digital" business operates)
C8	unlocking the value of data-driven decisions	by capturing and analysing high quality data	a clear understanding of the importance of data, why it is an enabler of organisational change, and why it is critical for transformation decision making needs to exist
C9	realising value creation	by balancing cost reduction and revenue generation	making the right investments in digital (technologies and resources), that will enable value creation and provide a meaningful ROI (return on investment)
C10	empowering employee experience	by creating a dynamic digital workplace	encouraging a value proposition for employees needs to be outlined (e.g. monetary, improved working environment, or a greater say in how the organisation operates)

Table 4: DT Leadership Characteristics Descriptions (adapted from: *reference withheld for review purposes*)

Number	Category		to from											
	What	How		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	
C1	communicating and executing a digital strategy and vision	by leveraging executive management support	C1			1					1		1	3
C2	prioritising the customer value proposition	by implementing an integrated digital platform	C2	1		2				1		1		5
C3	understanding the journey of organisational change	by embracing digital disruption	C3							1				1
C4	inspiring the organisation to change	by adopting an open culture and digital mindset	C4	2	1			1	1				2	7
C5	underpinning the organisational change	by using appropriate digital capabilities	C5	1	1	1	1		1	1	1	2	2	11
C6	collaborating cross functionally	by adopting a disruptive approach to innovation	C6	1	1		3	1				1	1	8
C7	redefining the business model	by optimising functionally aligned processes	C7		1							2	1	4
C8	unlocking the value of data-driven decisions	by capturing and analysing high quality data	C8	1	1	1								3
C9	realising value creation	by balancing cost reduction and revenue generation	C9			1								1
C10	empowering employee experience	by creating a dynamic digital workplace	C10			1	1							2
Total				6	5	7	5	2	2	3	2	6	7	

Table 5: Relationships (emerging from axial coding) between the DT Leadership Characteristics (colours show each virtuous relationship)

## **Findings & Discussion: The Underlying Practices of DT Leadership**

In this section we will present our findings. We start by presenting a high-level overview based on some key patterns emerging from our analysis (see Table 3) and also compare our understanding of these patterns against current literature. We then present an overview of the relationships between the ten DTL characteristics (based on our theorizing). The naming of the ten DTL characteristics respects the lexicon of the sixteen key informants involved in this study, notwithstanding the fact that there is an evolution in the description used as part of the data-to-theory process (Gioia et al., 2012).

Examining Table 3, we can identify a number of patterns of interest. These patterns also emerge from our coding of the sixteen key informant interviews. Table 3 presents the ten DT leadership characteristics in descending order of coded excerpts. The distribution of excerpts across the four practitioner voice quadrants (*Business Strategic* (BS), *Business Operational* (BO), *IT Strategic* (ITS), *IT Operational* (ITO)) is also highlighted. Each DT leadership characteristic is presented as a ‘what’ and ‘how’ combination, which emerged as part of the open and axial coding process. Furthermore, a brief description of the ten DT leadership characteristics is presented in Table 4; See (*reference withheld for review purposes*) for a more detailed description of the ten DT leadership characteristics and patterns of interest.

We are conscious that our coding efforts drive the takeaways on this topic of underlying practices associated with DT leadership. Having arrived at ten DT leadership characteristics we have balanced completeness with parsimony, in that what we have found can explain plausibly most of what can happen in the context of leading a DT initiative. In particular, we have arrived at a “*causal structure*” (Akkermans & van Helden, 2002) that could explain the complexity of DT leadership. As a result, in this section, using the analogy of the “*magnifying glass*”, we “*zero in on some parts of the whole image*” to “*find the most interesting and incisive parts to work with and emphasise*” (Akkermans & van Helden, 2002, p.44); (Gioia et al., 2012). Therefore, we specifically present “*deep processes*” in the relationships between the DT leadership characteristics (Gioia et al., 2012, p.17). In the sub-section

below, we state a number of formal propositions as an “*opportunity to speculate*” on where further theorising might lead (Gioia et al., 2012, p.17. Furthermore, taking on board our conclusions presented in this section, we also enfold the DT literature to highlight the novelty in our work and our contribution to the DT leadership conversation. In essence, understanding the relationships between the DT leadership characteristics (as underlying practices) will encourage the exploration of more appropriate DT implementation strategies, along with further improving the sense of organisational readiness and C-Suite appetite.

Our analysis suggests that DT leadership characteristic C5 (*underpinning the organisational change by using appropriate digital capabilities*) emerges as the single most impactful DT leadership characteristic. Reflecting on exiting literature this may come as no surprise, seeing as DT is often defined as *organisational change enabled/triggered by digital technologies* (Nadkarni & Prügl, 2021). Therefore, C5 appears fundamental in shaping the outcome of a DT initiative. Specifically, our analysis reveals a strong enabling relationship between C5 and C9 (*value creation*), and C5 and C10 (*employee experience*). For example, as discussed previously, *using appropriate digital capabilities* (C5) enables both *cost reduction and revenue generation* (C9), while *using appropriate digital capabilities* (C5) also enables the creation of a *dynamic digital workplace* (C10), ultimately, creating an employee value proposition. In fact, our analysis also reveals a strong enabling relationship between the DT leadership characteristic C7 (*redefining the business model by optimising functionally aligned processes*) and C9 (*realising value creation*).

Furthermore, while C3 only enables one DT leadership characteristic (C7 - *redefining the business model*), it is itself enabled by six DT leadership characteristics (the largest number across all characteristics). For example, as discussed previously, *understanding the journey of organisational change by embracing digital disruption* (C3) is enabled by the following: a well communicated *digital strategy and vision* (C1), a focus on *value creation* (C9), delivering a *value proposition* for both the *customer* (C2) and the *employee* (C10), and making more *data-driven decisions* (C8). In the next section we now look at the “*virtuous cycles*” between the DT leadership characteristics.

## The Underlying Practices of DT Leadership

Appreciating that we have produced a *laundry list* (Akkermans & van Helden, 2002), of ten DT leadership characteristics, we now aim to progress the theory on how these DT leadership characteristics affect each other. We reflect on the 45 relationships that exist (based on our theorising) between the ten DT leadership characteristics, and in this section, we focus specifically on six enabling relationships that emerged from our analysis (the underlying practices of DT leadership) as follows: *Collaborative Change*, *Digital Influence*, *Collaborative Tooling*, *Employee Spirit*, *Prioritised Platformisation*, and *Democratising Data*. We also qualify our understanding with current literature. These six relationships are all “*virtuous cycles*” (c.f. Akkermans & van Helden, 2002, p.35), and a virtuous cycle is best explained where two factors are “*seen to reinforce each other*” (Akkermans & van Helden, 2002, p.35) so as one factor goes up, the other factor will also increase.

### Collaborative Change (C6→C4→C6)

The relationship of (C6) collaborating cross functionally by adopting a disruptive approach to innovation and (C4) inspiring the organisation to change by adopting an open culture and digital mindset, illustrates the importance of the interaction and integration between functions in the organisation in adopting new ways of working together to achieve a culture and mindset shift and acceptance by employees and stakeholders so creating an improved and dynamic work environment for them. The enabling relationship between C6 and C4 was the most frequently observed in our analysis of the data. The relationship promotes a virtuous cycle between the DTL characteristics. This relationship is employee-centric and further highlights the criticality of employees to the outcome of a DT initiative. We appreciate that *cross-functional collaboration* is an internal organisational activity that breaks down silos and invites all employees into the activities associated with DT. This virtuous relationship translates into a **principle**, as follows: **(Principle 1)** *innovating around how employees collaborate inspires and enables change through an open and digital mindset which will itself enable cross-functional collaboration.*

Reflecting on existing DT literature, evidence in support of this virtuous cycle exists. For example, DT leaders need to build a digital culture to cultivate a willingness to take risks and to experiment with digital technologies, initially on a small scale,

before scaling such experiments (where successful) to the rest of the organisation ((c.f. Fehér et al., 2017), Dremel et al., 2017)). Furthermore, as stated by (Haffke et al., 2016); *“the level of cross-functional collaboration also affect the need for orchestration of digital change”*. However, reflecting on the stated proposition, connecting employee cross-functional collaboration from adopting a digital culture (promoting digitally enabled change) to a greater desire for cross-functional collaboration, because of the use of digital technologies, has not been examined to date.

#### **Digital Influence (C4→C5→C4)**

The relationship of (C4) inspiring the organisation to change by adopting an open culture and digital mindset and (C5) underpinning the organisational change by using appropriate digital capabilities, illustrates the importance of achieving a culture and mindset shift and acceptance by employees and to do so requires their embracing of digital capabilities which can create an improved and dynamic work environment for them. The enabling relationship between C4 and C5 promotes a virtuous cycle between the DTL characteristics. This virtuous relationship translates into a **principle**, as follows: **(Principle 2)** *fostering an open culture and digital mindset enables organisational change through introducing digital capabilities which will itself enable further openness to change.*

Reflecting on existing DT literature, evidence in support of this virtuous cycle exists. For example, DT leaders must work to ensure that they cultivate a digital mindset within the organisation and build the capability to respond to the disruptions associated with the introduction and use of digital technologies (Haffke et al., 2016); (Hansen et al., 2011). Furthermore, as stated by (Haffke et al., 2016) *“in order for incumbent businesses to take advantage of the opportunities that an increased focus on digitization affords them, they must undertake a digital transformation journey, often altering corporate culture in order to open the organization to new digital opportunities”*. However, reflecting on the stated proposition, connecting digitally enhanced change from an open/digital mindset to a greater desire for change, because of the introduction of digital technologies, has not been examined to date.



### **Collaborative Tooling (C5→C6→C5)**

The relationship of (C5) underpinning the organisational change by using appropriate digital capabilities and (C6) collaborating cross functionally by adopting a disruptive approach to innovation illustrates the process of transformation through digitalisation is clearly dependant on the impetus gained through integration and collaboration of business units throughout the entire enterprise. The enabling relationship between C5 and C6 promotes a virtuous cycle between the DTL characteristics. This virtuous relationship translates into a **principle**, as follows: **(Principle 3)** *embracing digitally enhanced change enables cross-functional collaboration which will itself enable greater use of digital tools and technologies.*

Reflecting on existing DT literature, evidence in support of this virtuous cycle exists. For example, DT leaders need to focus on redesigning the organisational structure in order to promote agility and flexibility, as enablers for creating cross-functional collaboration and alignment between organisational functions; (c.f. (Svahn et al., 2017); (Li et al., 2018)), (Maedche, 2016). Furthermore, as stated by (Singh & Hess, 2017), “exploiting ideas from both internal and external sources, for instance in the form of crowdsourcing and cross-company collaboration, using digital technologies”. However, reflecting on the stated proposition, connecting cross-functional collaboration from digitally enhanced change to a greater use of digital technologies, has not been examined to date.

### **Employee Spirit (C4→C10→C4)**

The relationship of (C4) inspiring the organisation to change by adopting an open culture and digital mindset and (C10) empowering employee experience by creating a dynamic digital workplace illustrates the need for a culture of openness and transparency to be fostered to create a conducive work environment that will get the best out of its employees. The enabling relationship between C4 and C10 promotes a virtuous cycle between the DTL characteristics. This virtuous relationship translates into a **principle**, as follows: **(Principle 4)** *encouraging employees to believe, thereby inspiring the organisation to change enables a digitally enhanced employee-centric workplace which will itself enable an ongoing culture and mindset shift.*

Reflecting on existing DT literature, evidence in support of this virtuous cycle exists. For example, DT leaders need to appreciate the significant role that employees play in DT. In particular, the flexibility and ambidexterity of employees coupled with their differing perspectives (IT and business) have a significant impact on the outcome of a DT initiative, especially where employees contribute to the DT decision making processes (c.f. (Yeow et al., 2018); (Dremel et al., 2017); (Hess et al., 2016)). However, coordinating the skills mix of employees, for both existing and new employees, is a key requirement for a positive DT outcome (c.f. (Hess et al., 2016); (Colbert et al., 2016); (Watson, 2017)). Furthermore, as stated by (Dery et al., 2017) “management prioritizes the activities that focus on the development and continuous improvement of employee experience in the organization... to develop workplace effectiveness”. These leaders encouraged experimentation with “new technologies and new approaches to work”. However, reflecting on the stated proposition, connecting an enhanced employee-centric workplace from inspiring change to a greater desire for ongoing digitally enhanced change, has not been examined to date.

#### **Prioritised Platformisation (C2→C7→C2)**

The relationship of (C2) prioritising the customer value proposition by implementing an integrated digital platform and (C7) redefining the business model by optimising functionally aligned processes, illustrates how customer value and experience supported by an effective digital architecture energises transformation of the business model through the digitalisation of business processes. The enabling relationship between C7 and C2 promotes a virtuous cycle between the DTL characteristics. This relationship is customer-centric (externally focused) and ensures that the DT initiative maintains a clear strategic focus and avoids drift, thereby, also avoiding waste. This virtuous relationship translates into a **principle**, as follows: **(Principle 5)** *prioritising a digital customer-focused platform enables the operational efficiencies of a digital business model which will itself enable the delivery of a customer-value inspired digital platform.*

Reflecting on existing DT literature, evidence in support of this virtuous cycle exists. For example, DT leaders need to design a digital platform that reflects the digital strategy (c.f. (Burgelman & Grove, 2007); (Vial, 2019)) and also prioritise the welfare of customers. By doing so, leadership will be able to identify new interaction and

engagement opportunities for customer value creation through digitisation, realised through gathering and analysing customer data (c.f. (Haffke et al., 2016); (Tumbas et al., 2015); (Setia et al., 2013)). Furthermore, as stated by (Singh & Hess, 2017) “a company undergoing a digital transformation uses new digital technologies such as social media, mobile access, analytics or embedded devices to enable major business improvements like enhancing customer experience, streamlining operations or creating new business models”. However, reflecting on the stated proposition, connecting the operational efficiencies of a digital business model, inspired by a customer-focused digital platform, to the actual delivery of a customer-centric digital platform, has not been examined to date.

#### **Democratising Data (C8→C1→C8)**

The relationship of (C1) communicating and executing a digital strategy and vision by leveraging executive management support and (C8) unlocking the value of data-driven decisions by capturing and analysing high quality data illustrate the closeness of the relationship with a digital strategy and vision and the use of data to create value propositions. The enabling relationship between C8 and C1 promotes a virtuous cycle between the DTL characteristics. This virtuous relationship translates into a **principle**, as follows: **(Principle 6)** *embracing data-driven decisions enables and steers the digital strategic vision which will itself enable data-driven decisions and the need for high quality data*. This relationship reflects the powerful sentiment of one of the key informants in this study, who state “*without data, you’re blind*”! Therefore, while data can be considered “*the soul of a DT*”, unfortunately, its value as a business asset is still unappreciated within the cultural fabric of many organisations.

Reflecting on existing DT literature, evidence in support of this virtuous cycle exists. For example, DT leaders will never understand the “*true value of a customer*” if data is maintained in “*silos*” in an “*undisciplined*” fashion (Fisher, 2009). Furthermore, as stated by (Singh & Hess, 2017, p.3) leadership should have a strong appreciation for the role of data in DT, and leaders, such as the Chief Data Officer, “*instead of treating data merely as a by-product of running the business, they devise strategies for exploiting the business’s data*”. However, reflecting on the stated proposition, connecting the existence of a data-driven digital strategy, to an increased appetite for

high quality data to support data-driven DT initiative decisions, has not been examined to date.

## **Conclusions and Implications**

In this paper we set ourselves the challenge of conducting research that is both rigorous and relevant. Throughout our inductive approach, we have maintained an “*analytical discipline*” in order to produce “*credible interpretations of data*” and conclusions that are both “*plausible and defensible*” (Gioia et al., 2012, p.15). Our efforts at concept development (underlying practices associated with DT leadership) and our making sense of the organisational world, that IS practitioners live in, affords others with the opportunity to continue discovering and developing concepts/categories similar to those we present in this paper. We enable such further research by providing sufficient transparency into our theorising efforts as we progressed from data-to-theory. In the next section we focus on the implications (of our theorising) for theory and practice.

## **Implications for Theory and Practice**

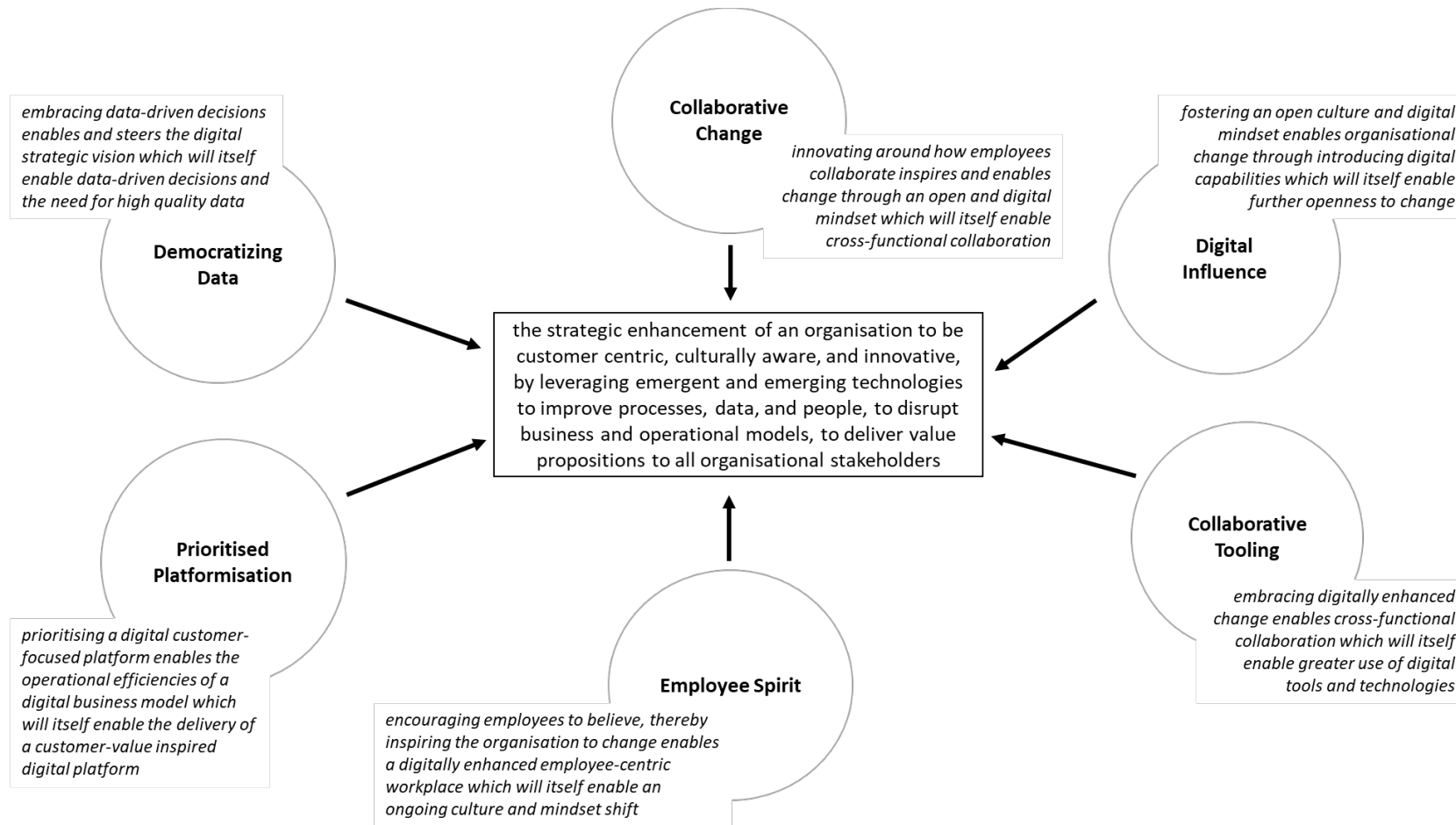
IT leadership has undergone much change over the past five decades (1970s to 2020s) especially with regards to its role in influencing organisational change through IT enablement. Historically, the role of IT was very much something to leverage (1970s & 1980s), to reduce cost, be siloed in business functions and provide limited volume of applications and services. The responsibility lay with IT and IS managers to drive investment and manage IT resources and to provide the necessary technology, applications, and services to support and underpin organisational change (Schein, 1996; Benjamin & Levinson, 1993; Porter & Millar, 1985). However, the mandate of IT changed to that of expansion (1990s & 2000s), whereby from originally supporting organisations, they were now enabling them, and with that the role of IT leadership continued to grow in influence. Organisations looked to IT Managers and CIOs, to lead change by championing the enablement of business strategies, drive further customer engagement, lower operational costs, foster workplace improvements for employees, and grow internal cross-functional collaboration (Bresciani et al., 2021; Baiyere et al., 2020; Cross et al., 1997).

So, what has changed for IT leadership with the movement to digitalisation (2010s & 2020s), with organisations now focused on digital transformation (DT) initiatives. The result was the IT mandate further changing from that of an enabler to that of transformer, taking on a greater emphasis and responsibility for organisational change (Wessel et al., 2021). DT leadership became responsible for leading the changes those in IT transformational positions were traditionally tasked with (c.f. Paavola et al., 2017; Ready et al 2020). This change saw new roles such as the chief digital officer (CDO) and chief data officer (CDaO) taking key positions in leading out those DT initiatives (Haffke et al., 2016; Singh & Hess, 2017). The priority for DT leadership is now to focus on convergence, collaboration, and continuity by ensuring an aligned digital strategy (Chaniyas et al., 2019; Hess et al., 2016), improved value creation (Bharadwaj et al., 2013; Matt et al., 2015), customer enhancement (Sia et al., 2016; Tumbas et al., 2015), employee experience (Liere-Netheler et al., 2018, Remane et al., 2017), and have a positive impact on redesigning the business model, increasing IT capabilities and developing an openminded culture (Singh & Hess, 2017; Horlacher, 2016).

It is reported that *importance* is the most critical dimension of relevance for IS practitioners. Similar to (Rosemann & Vessey, 2008 p.3) we view *importance* as research that “*meets the needs of practice by addressing a real-world problem in a timely manner [currently significant], and in such a way that it can act as the starting point for providing an eventual solution*”. While DT is a current hot topic and a top concern for many practitioners (both business and IT), the ability to lead a DT initiative, and be successful in doing so, is an area of IS research not yet well established. Therefore, the work presented in this paper is an effort at addressing this current shortfall. Using the key informant technique to capture the sixteen practitioner voices, allowed the DT stories to be interrogated, the outcome of which leads to the emergence of the underlying practices associated with **DT leadership** within the context of pre-digital organisations.

We argue that this research offers new insights at the individual level of analysis. Specifically, our work unveils the underlying practices of DT leadership within the context of pre-digital organisations. We show that DT leaders (the *who* across the four quadrants) focus on different practices (the *what* and the *how*) when leading a DT

initiative. To note, as per (Rosemann & Vessey, 2008 p.3), *accessibility* is understood as “*the research is understandable, readable, and focuses on results*” and *applicability* is understood to be “*whether it provides guidance and/or direction, and whether it provides concrete recommendations*” that are easy to apply in practice. Our six underlying practices of DT leadership offers practical guidance to organisations undertaking DT (see Figure 1). Such organisations can ask probing questions around their own *who*, *what* and *how* combinations. Therefore, we offer a practical toolkit to make sense of leading a DT initiative and guide a DT process toward a desirable outcome. For example, organisations can establish groups of DT leaders to cultivate each of the six virtuous cycle relationships (that define the underlying practices) between the DT leadership characteristics. For example, *Employee Spirit* highlights the importance of *Business Operational* practitioner voices to ensure that the DT process keeps the employee front and centre in all aspects of the journey.



**Figure 1: The Underlying Practices (with principles) of DT Leadership**

## References

- Akkermans, H. and van Helden, K., 2002. Vicious and virtuous cycles in ERP implementation: a case study of interrelations between critical success factors. *European journal of information systems*, 11(1), pp.35-46.
- Alhassan I, Sammon D & Daly M (2019), "Critical success factors for data governance": a telecommunications case study, *Journal of Decision Systems*, 28:1, 41-61, DOI:10.1080/12460125.2019.1633226
- Barker, K. K., Bosco, C., & Oandasan, I. F. (2005). Factors in implementing interprofessional education and collaborative practice initiatives: Findings from key informant interviews. *Journal of Interprofessional Care*, 19(sup1), 166-176. <https://doi.org/10.1080/13561820500082974>
- Barthel, P., Stark, N., & Hess, T. (2020). Exploring New areas for Project portfolio Management-Evolving Practices for Digital Transformation Projects. ECIS, [https://aisel.aisnet.org/ecis2020\\_rip/19\\_p.1-11](https://aisel.aisnet.org/ecis2020_rip/19_p.1-11)
- Basit, T. (2003). Manual or electronic? The role of coding in qualitative data analysis. *Educational research*, 45(2), 143-154. <https://doi.org/10.1080/0013188032000133548>
- Bayiere A., Salmela, H. & Tapanainen, T. 2020. Digital transformation and the new logics of business process management. *European Journal of Information Systems*, 29, 238-259.
- Bendig, D., Wagner, R., Jung, C., & Nüesch, S. (2022). When and why technology leadership enters the C-suite: An antecedents perspective on CIO presence. *The Journal of Strategic Information Systems*, 31(1), 101705. <https://doi.org/10.1016/j.jsis.2022.101705>
- Benjamin, R.I. and Levinson, E., 1993. A framework for managing IT-enabled change. *Sloan Management Review*, 34(4), pp.23-33
- Benova, L., Moller, A., & Moran, A. (2019). Qualitative data for: " The landscape of validation of global maternal and newborn health indicators through key informant interviews". <https://datacompass.lshtm.ac.uk/id/eprint/1403/>
- Berghaus, S., & Back, A. (2017). Disentangling the fuzzy front end of digital transformation: Activities and approaches. International Conference on Information Systems (ICIS), Seoul. [http://aisel.aisnet.org/icis2017\\_p.1-17](http://aisel.aisnet.org/icis2017_p.1-17)
- Bharadwaj, A., El Sawy, O. A., Pavlou, P. A., & Venkatraman, N. (2013). Digital business strategy: toward a next generation of insights. *MIS quarterly*, 37(No 2), 471- 482. <https://www.jstor.org/stable/43825919>
- Bhattacharjee, A., Limayem, M., & Cheung, C. M. (2012). User switching of information technology: A theoretical synthesis and empirical test. *Information & Management*, 49(7-8), 327-333. <https://doi.org/10.1016/j.im.2012.06.002>
- Bresciani, S., Huarng, K. H., Malhotra, A., & Ferraris, A. (2021). Digital transformation as a springboard for product, process and business model innovation. *Journal of Business Research*, 128, 204–210.
- Buchwald, A., Urbach, N., & Ahlemann, F. (2014). Business value through controlled IT: Toward an integrated model of IT governance success and its impact. *Journal of Information Technology*, 29(2), 128-147. <https://doi.org/10.1057/jit.2014.3>
- Burgelman, R.A. and Grove, A.S., 2007. Let chaos reign, then rein in chaos—repeatedly: Managing strategic dynamics for corporate longevity. *Strategic management journal*, 28(10), pp.965-979.



- Carroll, N. (2020). Theorizing on the normalization of digital transformations. Proceedings of the 28th European Conference on Information Systems (ECIS) p.1-17, [https://aisel.aisnet.org/ecis2020\\_rp/75](https://aisel.aisnet.org/ecis2020_rp/75)
- Chanias, S. (2017). Mastering digital transformation: The path of a financial services provider towards a digital transformation strategy. In Proceedings of the 25th European Conference on Information Systems, ECIS 2017.
- Colbert, A., Yee, N. and George, G., 2016. The digital workforce and the workplace of the future. *Academy of management journal*, 59(3), pp.731-739.
- Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology*, 13(1), p.3-21.
- Cross, J. "IT Outsourcing: British Petroleum," *Harvard Business Review*(73:3), May-June 1995,pp. 94-102.
- Day, J. M., Junglas, I., & Silva, L. (2009). Information flow impediments in disaster relief supply chains. *Journal of the Association for Information Systems*, p.638-659 10(8), DOI: 10.17705/1jais.00205, <https://aisel.aisnet.org/jais/vol10/iss8/>
- Dery, K., Sebastian, I.M., Meulen, N.v.d., (2017). The digital workplace is key to digital innovation. *MIS Quart. Executive*; 16 (2), 135–152.
- Dezdar, S. and Sulaiman, A. (2009), "Successful enterprise resource planning implementation: taxonomy of critical factors", *Industrial Management & Data Systems*, Vol. 109 No. 8, pp. 1037 1052 <https://doi.org/10.1108/02635570910991283>
- Dremel, C; Wulf, J; Herterich, M .; Waizmann, J.C; Brenner, W (2017); *MIS Quarterly Executive* . Jun2017, Vol. 16 Issue 2, p81-100.
- Fehér, P., Szabó, Z. and Varga, K., 2017. Analysing digital transformation among Hungarian organizations; 30<sup>th</sup> BLED eConference
- Fisher, B., Turner, R.K. and Morling, P., 2009. Defining and classifying ecosystem services for decision making. *Ecological economics*, 68(3), pp.643-653.
- Flores, W. R., & Ekstedt, M. (2012). A model for investigating organizational impact on information security behavior. *WISP 2012 Proceedings*. <https://aisel.aisnet.org/wisp2012/12>
- Gioia, Corley, K., & Hamilton, A. (2012). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational research methods*, 16(1), 15-31. <https://doi.org/10.1177/1094428112452151>
- Haffke, I., Kalgovas, B., & Benlian, A. (2017). Options for Transforming the IT Function Using Bimodal IT. *MIS Quarterly Executive*, 16(2). DOI:10.24251/HICSS.2017.660
- Haffke, I., Kalgovas, B. J., & Benlian, A. (2016). The Role of the CIO and the CDO in an Organization's Digital Transformation. International Conference on Information Systems ICIS2016, p.1-20 <https://aisel.aisnet.org/icis2016/ISStrategy/Presentations/3>
- Hansen, A.M., Kraemmergaard, P., Mathiassen, L., (2011). Rapid adaptation in digital transformation: a participatory process for engaging IS and business leaders. *MIS Quart. Exec.* 10 (4), 175–185.
- Hess, T., Matt, C., Benlian, A., & Wiesböck, F. (2016). Options for formulating a digital transformation strategy. *MIS Quarterly Executive*: Vol. 15 : Iss. 2 , Article 6. p.123 -139 <https://aisel.aisnet.org/misqe/vol15/iss2/6>
- Horlacher, A., & Hess, T. (2016). What does a Chief Digital Officer do? Managerial tasks and roles of a new C-level position in the context of digital

- transformation. Hawaii International Conference on System Sciences (HICSS), DOI: [10.1109/HICSS.2016.634](https://doi.org/10.1109/HICSS.2016.634)
- Hund, A., Wagner, H.-T., Beimborn, D., & Weitzel, T. (2021). Digital innovation: Review and novel perspective. *The Journal of Strategic Information Systems*, 30(4), 101695. <https://doi.org/10.1016/j.jsis.2021.101695> (<https://www.sciencedirect.com/science/article/pii/S0963868721000421>)
- Li, L., Su, F., Zhang, W. and Mao, J.Y., 2018. Digital transformation by SME entrepreneurs: A capability perspective. *Information Systems Journal*, 28(6), pp.1129-1157
- Liere-Netheler, K., Packmohr, S., & Vogelsang, K. (2018). Drivers of digital transformation in manufacturing.
- Maedche, Alexander (2016).,"Interview with Michael Nilles on “what makes leaders successful in the age of the digital transformation?”." *Business & Information Systems Engineering* 58, no. 4 (2016): 287-289. DOI: [10.1007/s12599-016-0437-1](https://doi.org/10.1007/s12599-016-0437-1)
- Markus, M. L., & Rowe, F. (2021). Guest Editorial: Theories of Digital Transformation: A Progress Report. *Journal of the Association for Information Systems*, 22(2), 1 <https://aisel.aisnet.org/jais/vol22/iss2/11> p.273 - 280
- Matt, C., Hess, T., & Benlian, A. (2015). Digital transformation strategies. *Business & information systems engineering*, 57(5), 339-343. DOI [10.1007/s12599-015-0401-5](https://doi.org/10.1007/s12599-015-0401-5)
- Miranda, Shaila; Berente, Nicholas; Seidel, Stefan; Safadi, Hani; and Burton-Jones, Andrew. 2022. "Editor's Comments: Computationally Intensive Theory Construction: A Primer for Authors and Reviewers," *MIS Quarterly*, (46: 2) pp.iii-xviii.
- McCarthy, P., Sammon, D., & Alhassan, I. (2022). "‘Doing’ digital transformation: theorising the practitioner voice", *Journal of Decision Systems*, 31(sup1), pp.341-361, DOI: [10.1080/12460125.2022.2074650](https://doi.org/10.1080/12460125.2022.2074650)
- McCarthy, P., Sammon, D., & Alhassan, I. (2021). Digital Transformation Leadership Characteristics: a Literature Analysis. *Journal of Decision Systems*, p.1-31. <https://doi.org/10.1080/12460125.2021.1908934>
- McKinsey. (2018). Insights Report Winning in Digital Ecosystems January 2018. *MIS Quarterly*, January 2018, p.411–436.
- Nadkarni, S., & Prügl, R. (2021). Digital transformation: a review, synthesis and opportunities for future research. *Management Review Quarterly*, 71(2), p.233-341. DOI: [10.1007/s11301-020-00185-7](https://doi.org/10.1007/s11301-020-00185-7)
- Nguyen, D. K., Broekhuizen, T., Dong, J. Q., & Verhoef, P. (2021). Are You Ready for Digital Transformation? Digital Readiness Configurations and Business Value across Industries. *International Conference on Information Systems ICIS 2021 Proceedings*. 7. [https://aisel.aisnet.org/icis2021/dig\\_innov/dig\\_innov/7](https://aisel.aisnet.org/icis2021/dig_innov/dig_innov/7) p.1-9
- Oberländer, A. M., Röglinger, M., & Rosemann, M. (2021). Digital opportunities for incumbents—a resource-centric perspective. *The Journal of Strategic Information Systems*, 30(3), 101670. <https://doi.org/10.1016/j.jsis.2022.101705>
- Pabst von Ohain, B. (2019). Leader Attributes for Successful Digital Transformation. *International Conference on Information Systems ICIS 2019 Proceedings*. 5., p.1-17 [https://aisel.aisnet.org/icis2019/practice\\_is\\_research/practice\\_is\\_research/5](https://aisel.aisnet.org/icis2019/practice_is_research/practice_is_research/5)

- Paavola, R., Hallikainen, P., & Elbanna, A. (2017). Role of middle managers in modular digital transformation: the case of Servu. (ECIS), Guimarães, Portugal, June 5-10, 2017 (pp. 887-903). ISBN 978-989-20-7655-3 Research Papers.  
[https://aisel.aisnet.org/ecis2017\\_rp/58](https://aisel.aisnet.org/ecis2017_rp/58)
- Porfirio, J. A., Carrilho, T., Felício, J. A., & Jardim, J. (2021). Leadership characteristics and digital transformation. *Journal of Business Research*, 124, 610-619. <https://doi.org/10.1016/j.jbusres.2020.10.058>
- Porter, M. E., and Millar, V. E. "How Information Gives You Competitive Advantage," *Harvard Business Review* (63:4), 1985, pp. 149-160
- Ready, D., Cohen, C., Kiron, D., and Pring, B. (2020) "The New Leadership Playbook for the Digital Age," *MIT Sloan Management Review*.  
<https://sloanreview.mit.edu/leadership2020>
- Remane, G., Hanelt, A., Nickerson, R.C. and Kolbe, L.M., 2017. Discovering digital business models in traditional industries. *Journal of Business Strategy*, 38(2), pp.41-51.
- Rosemann, M., & Vessey, I. (2008). Toward improving the relevance of information systems research to practice: the role of applicability checks. *MIS quarterly*, p.1-22. <https://doi.org/10.2307/25148826>
- Schein, E.H., 1996. Culture: The missing concept in organization studies. *Administrative science quarterly*, pp.229-240.
- Setia, P., Venkatesh, V., Joglekar, S., (2013). Leveraging digital technologies: how information quality leads to localized capabilities and customer service performance. *MIS Quarterly*. 37 (2), 565–590.
- Sia, S.K., Soh, C., Weill, P., (2016). How DBS Bank pursued a digital business strategy. *MIS Quarterly Executive*. 15 (2), 105–121.
- Singh, A., & Hess, T. (2017). How chief digital officers promote the digital transformation of their companies. *MIS Quarterly Executive*, 16(1), 1-17.  
<https://aisel.aisnet.org/misqe/vol16/iss1/5>
- Smith, H. A., & Watson, R. T. (2019). Digital transformation at carestream health. *MIS Quarterly Executive*, 18(1), 8. <https://aisel.aisnet.org/misqe/vol18/iss1/8>
- Svahn, F., Mathiassen, L., and Lindgren, R. (2017). "Embracing Digital Innovation in Incumbent Firms: How Volvo Cars Managed Competing Concerns," *MIS Quarterly* (41:1), pp.239-253.
- Tabrizi, B., Lam, E., Girard, K., & Irvin, V. (2019). Digital transformation is not about technology. *Harvard business review*, 13(March), p.1-6.  
<https://hbr.org/2019/03/digital-transformation-is-not-about-technology>
- Tan, B., Pan, S. L., Lu, X., & Huang, L. (2015). The role of IS capabilities in the development of multi-sided platforms: The digital ecosystem strategy of Alibaba. com. *Journal of the Association for Information systems*, 16(4), 2.
- Tumbas, S., Berente, N., and vom Brocke, J. (2017). "Three Types of Chief Digital Officers and the Reasons Organizations Adopt the Role," *MIS Quarterly Executive* (16:2), pp. 121– 134.
- Ure, I. (2018). Digital Transformation: The What, Why and How For Transformational Leaders <https://thefutureshapers.com/digital-transformation-transformational-leaders/>
- Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*, 28(2), Pages 118-144, ISSN 0963-8687, <https://doi.org/10.1016/j.jsis.2019.01.003>

- Wade, M., Noronha, A., Macaulay, J., & Barbier, J. (2017). Orchestrating digital business transformation. *Global Center for Digital Business Transformation, IMD and Cisco*. <https://www.imd.org/research-knowledge/reports/digital-orchestra/>
- Wade, M., & Shan, J. (2020). Covid-19 Has Accelerated Digital Transformation, but May Have Made it Harder Not Easier. *MIS Quarterly Executive*, 19(3). p.213 - 220. <https://aisel.aisnet.org/misqe/vol19/iss3/7>
- Walsham, G. (2006). Doing interpretive research. *European Journal of Information Systems*, 15(3), 320-330. <https://doi.org/10.1057/palgrave.ejis.3000589>
- Wessel, L., Baiyere, A., Ologeanu-Taddei, R., Cha, J., & Blegind-Jensen, T. (2021). Unpacking the difference between digital transformation and IT-enabled organizational transformation. *Journal of the Association for Information Systems*, 22(1), 102-129. <https://doi.org/10.17705/1jais.00655>
- Whittaker, R. (2012). Issues in mHealth: findings from key informant interviews. *Journal of medical Internet research*, 14(5), e129. [doi:10.2196/jmir.1989](https://doi.org/10.2196/jmir.1989)
- Yeow, A., Soh, C., & Hansen, R. (2018). Aligning with new digital strategy: A dynamic capabilities approach. *The Journal of Strategic Information Systems*, 27(1), 43-58. [10.1016/j.jsis.2017.09.001](https://doi.org/10.1016/j.jsis.2017.09.001)
- Yoo, Y. (2013). The Tables Have Turned: How Can the Information Systems Field Contribute to Technology and Innovation Management Research? *Journal of the Association for Information Systems*, 14(5), 227–236.

## Appendix A: Key Informant Overview

Key Informant Position	KI Quadrant	DT Initiative	DT Classification	Experience (years)	Sector	Industry	Org Type	Org Size (employees) (all figures are less than)
IT Services Director	ITS	S	L	20-25	PU	H.Ed.	SME	2.5K
Chief Information Officer	ITS	M	L	15-20	PR	Agri	SME	.2K
Global Director of Digital Services Platform	ITS	M	G	25-30	PR	Energy	MNC	10K
Senior Software Development Manager	ITS	S	L	20-25	PR	S/ware	MNC	2K
Chief Executive Officer & VP	BS	M	G	25-30	PR	Energy	MNC	10K
Director of Academic Affairs & Digital Services	BS	S	L	20-25	PU	H.Ed.	SME	2.5K
Chief Information Officer & VP	BS	M	G	25-30	PR	Tech	MNC	50K
Senior Global Business Transformation Director	BS	M	G	20-25	PR	Tech	MNC	50K
Senior Digital Solutions Engineering Manager	ITO	S	G	15-20	PR	Tech	MNC	15K
IT Manager	ITO	S	L	15-20	PU	H.Ed.	SME	1.8K
Chief Technology Officer	ITO	S	L	15-20	PR	Agri	SME	.15K
Lead Digital MIS Analyst	ITO	S	L	10-15	PU	H.Ed.	SME	2.5K
Business Transformation Officer	BO	M	G	15-20	PR	Tech	MNC	50K
Director of Operations & Global Support Services	BO	M	G	20-25	PR	Tech	MNC	15K
Business Transformation Director	BO	M	G	20-25	PR	H/care	MNC	80K
Business Transformation Manager	BO	M	G	15-20	PR	Tech	MNC	15K
Business Strategic (BS), IT Strategic (ITS), Business Operational (BO), IT Operational (ITO) DT initiative (S-Single; M-Multiple)   DT Classification (G-Global; L-Local)   Sector (PU-Public; PR-Private)								

# Generative AI in the Workplace: A Dive into Employees' Acceptance in China

*Research In progress*

## **Abstract (around 150 words)**

*This paper explores factors influencing the adoption of generative AI (GenAI) among employees in China endeavouring to alleviate the excessive overtime work culture. Built on UTAUT and the institutional theory, a research model examines adoption determinants at the individual (performance expectancy, effort expectancy, personal innovativeness and trust), the organisational (social influence, facilitating conditions, organisational innovation climate), and the national (AI policy landscape) levels. A survey methodology will collect data from Chinese employees across industries. The paper argues GenAI can augment human capabilities and automate repetitive tasks, enabling employees to focus on creative, strategic activities and restore work-life balance. By revealing multi-level drivers of GenAI adoption, theoretical contributions include extending technology acceptance research to emerging GenAI while providing practical implications for organizations undergoing AI transformation. As one of the first examinations of employee GenAI adoption in the Chinese context, this study provides timely insights into harnessing AI's benefits for workers' well-being.*

**Keywords:** Generative Artificial Intelligence, Technology Adoption, Overtime, Quality of life, Digital Work

## **1.0 Introduction**

### **1.1 Overtime in China**

A salient feature of China's work culture in recent years has been the emergence and prevalence of the '996' working regime – a demanding routine that involves working from 9 a.m. to 9 p.m., six days a week (Li et al., 2021). According to the latest data from the National Bureau of Statistics of China (NBSC, 2023), the average workweek for employees in Chinese enterprises is 48.7 hours which exceeds the legally mandated minimum of 40 hours (NPCC, 2009). The extended work hours, often surpassing legally mandated limits, have been linked to increased stress, burnout, and a host of mental health issues (Afonso, et al., 2017). This toxic work culture has raised concerns regarding employee quality of life, work-life balance, overall job satisfaction and organisational commitment (Fagan et al., 2012).

### **1.2 Gen AI As a Potential Solution**

The dawn of the digital age has ushered in a plethora of technological advancements, with Artificial Intelligence (AI) standing at the forefront of this revolution. AI's potential to redefine job roles, influence organizational dynamics, reshape industries, and boost global economy is undeniable (McKinsey & Company, 2023). Particularly, generative AI, which is capable of creating novel content and artifacts such as text, code, images, video, and code. Even more, GenAI can automate tasks, assist in decision-making and augment human capabilities (García-Peñalvo and Vázquez-Ingelmo, 2023). It's helping employees lift the dirty, dull, dangerous and difficult tasks from their shoulders so then they can simply focus on the very essence of our work, the vision, the idea, and the purpose. By training on large data sets, users can reduce the burden of repetitive tasks and complex computations, enabling them to focus on more creative and more strategic activities which is the very essence of what work is really about and to have more time focusing parts their skills that are more human-centric, like creativity, problem solving, empathy, and leadership.

With the release of systems like DALL-E for generating images and ChatGPT for natural language text, generative AI is gaining rapid traction because it lowers the threshold of using AI, which means that users don't need to know much professional knowledge of AI, but only need simple training or even no training to use GenAI. Therefore, GenAI is considered as a solution to the overtime issue in China. Organizations are seeking to explore and adopt applications by generative AI across areas like content creation, software development,

customer service, and data analysis (IBM, 2022). However, the implementation of GenAI into the workplace does not occur in a vacuum. Many factors play an essential role in shaping the adoption of GenAI. Therefore, given this backdrop, this research endeavours to address the following research question:

What are the primary factors influencing the adoption of generative AI among employees in the workplace in China?

## **2.0 Literature Review**

In this part, we seek to summarize the theoretical models in AI adoption studies in different fields as a solid theoretical foundation of our research on GenAI adoption among employees in the workplace.

### **2.1 AI Adoption**

Prior research on AI adoption in the workplace has largely focused on organizational-level factors, finding drivers like top management support and strategic vision (Ransbotham et al., 2017). However, challenges around data, skills, and ethics have slowed adoption (Duan et al., 2019). Studies indicate mixed employee sentiments towards AI automation (World Economic Forum, 2023). Risks include deskilling and heightened monitoring (Christin, 2020; Kellogg et al., 2020). Still, there is limited understanding of facilitators and barriers influencing individual employees' adoption of AI, especially emerging generative AI. Research is predominantly set in Western contexts, with minimal examination of how cultural differences shape employee AI perceptions and use. National-level policies on AI governance could also affect adoption but remain understudied (Gasser and Almeida, 2017). In summary, current literature provides inadequate insight into drivers of employee-level generative AI adoption and the effects of cultural context. This developmental paper aims to address these gaps through an empirical study of employees' adoption of GenAI in the workplace.

### **2.2 AI Adoption Models**

Prior research on AI adoption employs a range of theoretical models to understand and predict the factors that drive the adoption process. The Technology Acceptance Model (TAM) has been extensively used to study acceptance of AI technologies. TAM predicts adoption intentions from perceived usefulness and perceived ease of use (Davis, 1989). Many studies have augmented TAM with additional factors like trust, risk and social influence (Huang et al., 2019;



Seo and Lee, 2021). TAM has been applied across contexts including healthcare (Ye et al., 2019), education (Kim et al., 2020) and customer service (Gao and Huang, 2019). The Unified Theory of Acceptance and Use of Technology (UTAUT) is another prominent model applied in AI acceptance research. UTAUT considers performance expectancy, effort expectancy, social influence and facilitating conditions as key determinants of adoption intentions and use behavior (Venkatesh et al., 2003). UTAUT has shown higher explanatory power than TAM and has been utilized to study AI adoption decisions in areas like organizations (Andrews et al., 2021), healthcare (Fan et al., 2020) and consumer products (Gansser and Reich, 2021).

The Theory of Planned Behavior (TPB) has also been used, albeit less extensively, in some AI acceptance studies. TPB highlights attitudes, subjective norms and perceived control as drivers of adoption intentions (Ajzen, 1980). For instance, Chai et al. (2021) applied TPB to assess primary school students' intentions to learn AI. The newer AI Device Use Acceptance Model (AIDUA) was designed specifically to explain consumer adoption of AI technologies (Gursoy et al., 2019). AIDUA incorporates technology-specific factors like anthropomorphism, social influence, and hedonic motivation. Initial applications in hospitality have been promising, but more research is needed to validate AIDUA across contexts relative to established models like TAM and UTAUT. Since these model cannot achieve the aim of our research, so we develop a theoretical framework based on UTAUT and the institutional theory.

### 3.0 Research Model and Hypotheses

In this part, we seek to propose a theoretical framework, including the AI adoption factors and hypotheses.

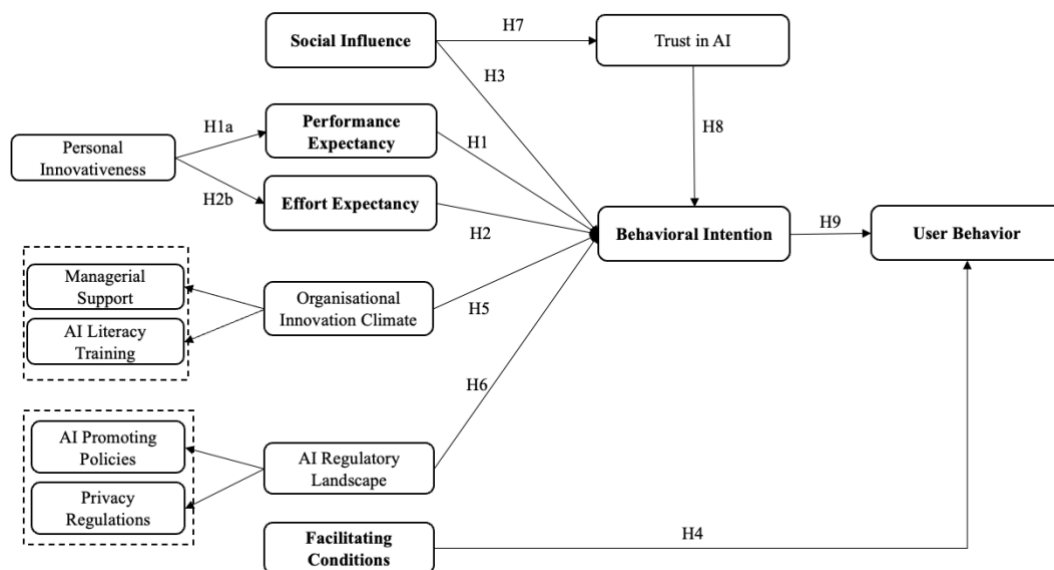


Figure 1. Proposed Theoretical Framework

### 3.1 Theoretical Foundation and Hypothesis Development

This research applies an extended Unified Theory of Acceptance and Use of Technology (UTAUT) model (Venkatesh et al., 2003) and combine the institutional theory (Meyer and Rowan, 1977) to examine factors influencing employees' adoption of generative AI in the workplace in China from three levels, the individual level, the organisational level and the national level. The base UTAUT model proposes that performance expectancy, effort expectancy, social influence, and facilitating conditions determine behavioral intention and use behavior (Venkatesh et al., 2003). Building on this, for individual level factors, we add Personal Innovativeness as an antecedent. In addition, we incorporate additional constructs relevant to AI adoption, which are Organisational Innovation Climate at the organisational level and AI Regulatory Landscape at the national level.

Several hypotheses can be developed based on the proposed model:

H1	Performance expectancy has a positive influence on employees' behavioural intention to use GenAI in the workplace in China.
H1a:	Personal innovativeness has a positive influence on Performance expectancy for GenAI.
H2	Effort expectancy positively influences employees' behavioural intention to use GenAI in the workplace in China.
H2b	Personal innovativeness has a positive influence on effort expectancy for GenAI.
H3	Social influence has a positive influence on employees' behavioural intention to use GenAI in the workplace in China.
H4	Facilitating conditions positively influence employees' use behaviour of GenAI in the workplace in China.
H5	Organizational innovation climate has a positive influence on employees' behavioural intention to GenAI in the workplace in China.
H6	AI regulatory landscape positively influences employees' behavioural intention to use GenAI in the workplace in China.
H7	Social influence has a positive influence on employees' trust to GenAI.
H8	Trust in AI positively influences employees' behavioural intention to use GenAI.
H9	Behavioural intention positively influences employees' use behaviour for GenAI in the workplace in China.

This theoretical model incorporates factors identified in the literature as relevant to employee AI adoption, adapted to the context of generative AI. Testing the hypothesized relationships will provide greater understanding of what drives employee usage of this emerging technology.

### 4.0 Research Method and Data Collection

This research is to carry out a quantitative survey of employees within Chinese organizations. The questionnaires would be self-administered and constructed using established and validated

scales and the questionnaires would be designed in bilingual format, encompassing both English and Chinese languages. The eligible individuals would encompass all individuals who are currently employed within mainland China. Participants could be possibly enlisted through social media platforms such as Weibo, Zhihu and Xiaohongshu.

## **5.0 Conclusion**

By examining the factors that influence employees' adoption of GenAI in the Chinese context, the proposed research could make contributions to provide new theoretical perspectives to the technology adoption literature in emerging economies. Additionally, it offers valuable insights for organizations that are currently undergoing AI transformation, particularly regarding the influence of national AI policy and organizational innovation climate.

## References

- Afonso, P., Fonseca, M. and Pires, J. (2017) Impact of working hours on sleep and mental health. *Occupational Medicine*, 67(5), pp.377-382.
- Ajzen, I. and Fishbein, M. (1980) *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Andrews, J.E., Ward, H. and Yoon, J. (2021) UTAUT as a model for understanding intention to adopt AI and related technologies among librarians. *The Journal of Academic Librarianship*, 47(6), p.102437.
- Chai, C. S., Lin, P. Y., Jong, M. S. Y., Dai, Y., Chiu, T. K., and Qin, J. (2021) Perceptions of and behavioral intentions towards learning artificial intelligence in primary school students. *Educational Technology & Society*, 24(3), 89-101.
- Christin, A. (2022) *Metrics at work: Journalism and the contested meaning of algorithms*. Princeton: Princeton University Press.
- Davis, F. D. (1989) 'Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology', *MIS Quarterly*, 13(3), pp. 319-340.
- Duan, Y., Edwards, J. S., and Dwivedi, Y. K. (2019) Artificial intelligence for decision making in the era of Big Data—evolution, challenges and research agenda. *International journal of information management*, 48, 63-71.
- Fagan, C., Lyonette, C., Smith, M. and Saldaña-Tejeda, A. (2012) *The influence of working time arrangements on work-life integration or 'balance': a review of the international evidence* (No. 32). Geneva: ILO.
- Gansser, O.A. and Reich, C.S. (2021) A new acceptance model for artificial intelligence with extensions to UTAUT2: An empirical study in three segments of application. *Technol. Soc.* 65, 101535 <https://doi.org/10.1016/j.techsoc.2021.101535>.
- Gao, B. and Huang, L. (2019) Understanding interactive user behavior in smart media content service: An integration of TAM and smart service belief factors. *Heliyon*, 5(12).
- García-Peñalvo, F. and Vázquez-Ingelmo, A. (2023). What do we mean by GenAI? A systematic mapping of the evolution, trends, and techniques involved in Generative AI.
- Gasser, U. and Almeida, V.A. (2017) A layered model for AI governance. *IEEE Internet Computing*, 21(6), pp.58-62.

- Gursoy, D., Chi, O.H., Lu, L. and Nunkoo, R. (2019) Consumers acceptance of artificially intelligent (AI) device use in service delivery. *International Journal of Information Management*, 49, pp.157-169.
- Huang, Y.K., Hsieh, C.H., Li, W., Chang, C. and Fan, W.S., 2019, December. Preliminary Study of Factors Affecting the Spread and Resistance of Consumers' Use of AI Customer Service. In *Proceedings of the 2019 2nd Artificial Intelligence and Cloud Computing Conference* (pp. 132-138).
- IBM (2022) *IBM Global AI Adoption Index 2022*. IBM Corporation.
- Kellogg, K. C., Valentine, M. A., & Christin, A. (2020) Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, 14(1), 366-410.
- Kim, J., Merrill, K., Xu, K. and Sellnow, D.D., 2020. My teacher is a machine: Understanding students' perceptions of AI teaching assistants in online education. *International Journal of Human-Computer Interaction*, 36(20), pp.1902-1911.
- Li, D., Wang, Z., and Yang, C. (2021) A Controversial Working System in China: The 996 Working Hour System. *2021 International Conference on Arts, Law and Social Sciences (ALSS 2021)*. DOI: 10.23977/ALSS2021030
- McKinsey & Company (2023) *What's the future of generative AI? an early view in 15 charts*, McKinsey & Company. Available at: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/whats-the-future-of-generative-ai-an-early-view-in-15-charts> (Accessed: 30 October 2023).
- Meyer, J. W. and Rowan, B. (1977) Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340-363.
- NBSC (2023) *The economy started well in the first quarter.*, NBSC. Available at: [http://www.stats.gov.cn/xxgk/sjfb/zxfb2020/202304/t20230418\\_1938732.html](http://www.stats.gov.cn/xxgk/sjfb/zxfb2020/202304/t20230418_1938732.html) (Accessed: 10 June 2023).
- NPCC (2009) *Labour Law of the People's Republic of China, Social Law*. Available at: [http://www.npc.gov.cn/zgrdw/englishnpc/Law/2007-12/12/content\\_1383754.htm](http://www.npc.gov.cn/zgrdw/englishnpc/Law/2007-12/12/content_1383754.htm) (Accessed: 2023).
- Ransbotham, S., Kiron, D., Gerbert, P., & Reeves, M. (2017) Reshaping business with artificial intelligence: Closing the gap between ambition and action. *MIT Sloan Management Review*, 59(1).
- Seo, K.H. and Lee, J.H., 2021. The emergence of service robots at restaurants: Integrating trust, perceived risk, and satisfaction. *Sustainability*, 13(8), p.4431.

- Venkatesh, V., Morris, M. G., Davis, G. B. and Davis, F. D. (2003) 'User Acceptance of Information Technology: Toward a Unified View', *MIS Quarterly*, 27(3), pp. 425-478.
- Wenjuan, F., Liu, J., Shuwan, Z. and Pardalos, P.M., 2020. Investigating the impacting factors for the healthcare professionals to adopt artificial intelligence-based medical diagnosis support system (AIMDSS). *Annals of Operations Research*, 294(1-2), pp.567-592.
- World Economic Forum (2023) *The Future of Jobs - World Economic Forum*. Available at: [https://www3.weforum.org/docs/WEF\\_Future\\_of\\_Jobs\\_2023.pdf](https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf) (Accessed: 30 October 2023).
- Ye, T., Xue, J., He, M., Gu, J., Lin, H., Xu, B. and Cheng, Y., 2019. Psychosocial factors affecting artificial intelligence adoption in health care in China: Cross-sectional study. *Journal of medical Internet research*, 21(10), p.e14316.

# MOVING ACTION LEARNING SETS ONLINE: REFLECTING ON PRIVACY, INTERSECTIONALITY AND GROUP FAILURE

Eleni Tzouramani, University of the West of Scotland, UK [eleni.tzouramani@uws.ac.uk](mailto:eleni.tzouramani@uws.ac.uk)

## Abstract

*This paper provides an account of my experience and insights gained while navigating the transition of Action Learning Sets (ALS), initially conducted face-to-face, into an online format during the COVID 19 lockdowns. Employing autoethnography as the methodological approach, this account allows for a critical reflection of my professional practice and the lessons I learnt from this process in relation to the wider social landscape that influenced these experiences during this particular period. Reflecting on 6 Action Learning Sets that moved online, I explore the online space as the context in which professional, student and family identities intersect. Within this context, issues of privacy, trust, intersectionality and group failure surface, calling for the need to redesign ALS for online contexts in ways that provide inclusive, safe and effective learning.*

## 1.0 Introduction and conceptual framework

The change we thought might take years to materialize happened overnight when all university classes, workshops and meetings moved online in 2019. There is no scope here to discuss the specifics of why and how it happened or go into long discussions about COVID 19, as the focus of this paper lies in an autoethnographic exploration of conducting Action Learning Sets online. Evaluating what worked, what didn't work and what perhaps could have worked better, this study aims to shed light on the challenges and potential of collaborative problem solving online.

According to Pedler (2008) Action Learning is an approach to individual and organisational development. Working in small groups known as 'sets', people tackle important organisational issues or problems and learn from their attempts to change things. Developed by physicist Reg Revans in the 1950's as a way to educate managers, Action Learning challenged traditional, prescriptive methods of learning and repositioned learning and development into a social context. Action Learning occurs mainly in Action Learning Sets (ALS) providing a space for productive dialogue, problem solving and testing hypotheses (Kehdr et al. 2022). Revans (1998) designed this methodology to integrate: Research into what is unclear, Learning about what is unknown and Action to resolve a problem. Involving complex interpersonal

relationships and processes (Yeadon-Lee 2013) ALSs can be sensitive to changes and power dynamics. Vince (2004) acknowledges the political dimension within ALS which can potentially be revealed, discussed and transformed.

Aiming to start a discussion on redesigning ALS for online spaces, in this autoethnographic account, I connect my experience of interactions and processes in online ALS to uncover wider issues of privacy, intersectionality and the challenges inherent in group learning. This is important in the context of current educational practices and the paradigm shift towards co-creation in learning (Bovill 2020, Elg et al. 2012) which emphasizes active participation in learning with the potential of creating inclusive and innovative learning environments. The theoretical framework guiding this paper follows process relational perspectives exploring mainly the ways in which intersecting identities (Crenshaw 1991) are constituted within online interactions and learning experiences.

From this perspective, the online space is approached as a context within which different identities intersect in complex ways, challenging the ways in which traditionally collaborative learning modalities work. Shotter's (2011) 'witness thinking' provides a valuable framework for dialogic interaction, exploring how learning unfolds within an action learning set. Proximity is not necessarily a prerequisite for embodying learning or for our 'expressive responsive understanding' (Shotter 2005). Instead, it is the quality of presence, of listening and being aware of the 'spontaneous responsiveness of our living bodies' (Shotter 2004) that shapes the dialogic relationship. The conditions within which this can happen can offer insights for designing or rather redesigning Action Learning Sets for online contexts in ways that maintain privacy and allow for safe enactment and exploration of identity intersections.

I have intentionally omitted Vygotsky in my analysis to concentrate on the dynamics and complexity of interactions instead of the process of constructing meaning. To explore ALS resilience during change, I draw upon Haslam and Reicher's (Haslam and Reicher 2006, 2007 and Reicher and Haslam 2006) ideas on social identity and group failure.

## 2.0 Writing Autoethnography

Reflection is often used in qualitative research in the form of reflexivity (Alvesson and Skoldberg 2000, Taylor and White 2000) or reflective practice (Schon 1983). Critical reflection



as a methodological approach, is part of ethnographic research, where the researcher becomes deeply immersed in the research, engaging in observations while reflecting on both the research process and personal experiences. This process demands a deliberate and systematic introspection by the researcher, constituting a critical aspect of qualitative research. Critical reflection is particularly prevalent within phenomenological and hermeneutic approaches, where the researcher's subjectivity and experience play a pivotal role in understanding their positionality.

Autoethnography (Ellis and Bochner 2000) as a qualitative research method, involves reflective writing that delves into the researcher's lived experience to uncover and address social and cultural challenges. While sharing some commonalities with other reflective methods like autobiography or storytelling, what sets autoethnography apart is its utilization of personal experiences to critically evaluate broader cultural contexts. Sparkes (2013) suggests it is a mode of being and a mode of knowing, involving cultural interpretation and connecting self and others. What the autoethnographic researcher needs to pay attention to is to not indulge in self-ruminations (Sparkes 2002) but to mostly use this method to live an 'examined life' so as to increase caring for others and contribution to the common good (Holman Jones, Adams and Ellis 2013). In this way, autoethnography serves as a tool for self-transformation and for reshaping both our thinking and research practices.

Autoethnography involves navigating multiple layers of consciousness and facilitates the connection of personal experiences to broader cultural contexts. The process involves going back and forth several times between introspection and the exploration of sociocultural aspects of personal experience until they meld into an intertwined whole. In my exploration, I found it to be a process that requires vulnerability, where I often came to examine, question and challenge aspects of my identity, role, memory and my own writing. While this paper draws upon a) my notes for six ALS b) my personal journal/ diary and c) my recollections of that period of time, it is important to clarify that the perspective presented in this work does not represent the viewpoints of the ALS participants but rather my individual perspectives and interpretations of the context of the COVID 19 lockdowns and the transition of face-to-face ALS to an online format.

From a practical perspective, in my process, I encountered challenges with my diaries, which, while extensive, presented issues due to unclear handwriting, a mix of Greek and English text

and the use of nicknames for individuals whose identities I can no longer remember. Thankfully, my notes from the ALS were more organised and legible, though they primarily consisted of factual information and lacked evaluative comments.

Critiques of autoethnography often revolve around the potential for self-indulgence in the process and around questions regarding its validity, as this methodology is considered to be too subjective to ensure rigorous research outcomes. To address these concerns, I adopt a continuous questioning approach drawing on Winkler's (2018) emphasis on regularly evaluating to what extent does my story enable me and the readers to understand culture as well as on Gorichanaz' (2021) suggestion to keep asking: 'How does this move us forward?' as a valuable tool to improve the methodological rigour of this research.

### 3.0 Action Learning Sets as learning spaces.

During the transition to online, the ALS setting changed. Working at the university provided a professional space designed specifically to facilitate group processes whereas for the online ALS, we were sitting at sofas at home, in bedrooms or dining tables. This also allowed for individual 'spacing out' looking at others' backgrounds, colours and to add more distraction, family members were moving about at the back of screens, often turning to look at the screen. Due to this being a lockdown, the whole family was at home, there was not enough professional space for everyone in the household and the allocation of space seemed to prioritise work over studies and from this managerialist way of thinking the person with the greater financial contribution would be allocated more resources and not the student. It was different for everyone, but I wondered if it was always the husband/father who is allocated 'the office' the most suitable space at home for work or even the person with the most important or highest paid job. We were all from different backgrounds, genders and ethnicities which calls for further research into the sociocultural dimensions that influence the allocation of space and by extension the experience of learning.

### 4.0 Online ALSs as contexts for identity experimentation

Within this context, participants were bringing their professional, student and personal/family selves into the set, holding interesting polarities between intersecting identities. The boundaries that once kept our professional, academic and personal identities fragmented started dissolving, inviting a confluence of identities within and between us, further increasing the complexity of

the learning space. Here the focus of the discussion shifts towards an exploration of identity intersections that lie in the context of interaction as shaped by the online ALS environment as well as on the complexity of identities.

Participants were also enacting their relational identities, introducing their significant others, children, spouses and pets to the set. The introduction of elements from our personal lives into the academic setting intensified identity work.

## 5.0 Online ALS as spaces of self-regulation and control

Within the online setting there can be a 'Panopticon' (Foucault 1977) effect where individuals do not know when they are being observed so they are in a state of constant self-regulating during the action learning set. This invisible surveillance was felt as being in a constant state of self-monitoring demanding a continuous performance of attentiveness. This creates the effect of allowing the mind, as the part that is not being observed, to rebel or to seek the freedom to wander.

What mostly concerned me was the issue of privacy, also mentioned by several facilitators and students. When family members were seen to be walking back and forth at the background, it was not possible for facilitators to safeguard a safe space, resulting in discussions not being as open as they could be. This intrusion, however unintentional, restricted my ability as facilitator to create a safe space for participants and reduced the depth and openness of our discussions. I sensed clear hesitancy by ALS participants to share as openly, possibly being self-conscious about being overheard by their and others' family or just to avoid family scrutiny.

The change left us disconcerted, in a fragmented state, thinking the set can continue from where we were left working in person together. From this I learnt that a new norming stage should have been actioned where the ALS would discuss how their ground rules would work in the online setting and what new ground rules and values should be set. This would have allowed the set to start anew and develop in the same way any new group develops.

To me, the compromise of the privacy and safety of the ALS space undermines its purpose and reduces its effectiveness as a method for learning. This calls for rethinking the evolving online ALS modalities to secure their privacy, safety and authenticity.

## 6.0 Online ALS as contexts for enacting group dynamics

Apart from the issue of privacy mentioned above, the online environment tends to allow for disengagement. This does not necessarily reflect the individual or the team. Haslam and Reicher (2006) suggest that, although the context as well as individual characteristics are important for the development of group, it is group failure that can lead to chaos. This resonates with my experience as I observed ALS groups dissociating, dismantling or malfunctioning when moving online. The question then shifts towards the underlying causes of disengagement. Was it the change in setting that unsettled the groups, or were there other, more intricate dynamics at play? Why did one of the ALS maintain its structure and effectiveness? These questions have taught me there are multiple ways of navigating the complexities of group interaction. In this section I will be exploring the question How do we analyse the complex interrelation of resilience factors within individual lives and across shared experiences?

## 7.0 Conclusions

Through this autoethnographic journey, the challenges encountered do not suggest that ALS cannot take place online but, rather they emphasise the need for innovation, calling for reimagining and redesigning ALS for the particularities of the digital space, ensuring privacy, inclusivity and safety in learning. The scope of autoethnography is not to generalize but to reveal nuanced, personal experiences that can inform broader understandings of intersectionality, online ALS interactions and innovation in learning practices. In its entirety, this working paper will seek to explore the connections between deep personal experiences and wider sociocultural challenges. Further, systematic research is needed to help provide insights on designing Action Learning Sets for online contexts which can facilitate inclusive, effective and safe learning.

## References

- Alvesson, M., and Skoldberg, K. (2000) *Reflexive Methodology: New Vistas for Qualitative Research*. London: Sage.
- Bovill, C. (2020) Co-creation in learning and teaching: The case for a whole-class approach in higher education. *Higher education*, 79(6):1023-1037.
- Brockbank, A. and McGill, I. (2002) *Reflective Learning in Practice*. Aldershot: Gower
- Crenshaw, K. (1991) Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review*, 43(6): 1241–1299.
- Elg, M., Engström, J., Witell, L. and Poksinska, B. (2012) Co-creation and learning in health-care service development. *Journal of Service Management*, 23(3): 328-343.
- Ellis, C. and Bochner, A.P. (2000) Autoethnography, personal narrative, reflexivity: researcher as subject. In *Handbook of Qualitative Research*, edited by N.K. Denzin and Y.S. Lincoln (2nd ed). Thousand Oaks, CA: Sage:733–768.
- Foucault, M. (1977) *The Discipline and Punish: The birth of the prison*. Translation by Μπέτσελος Τάσος. Πλέθρον Φιλοσοφία (2011).
- Fourie, I. (Ed.) (2021) *Autoethnography for Librarians and Information Scientists*. London: Routledge.
- Gorichanaz, T. (2021) “How does this move us forward?”: A question of rigour in autoethnography. In Fourie, I. (Ed.) *Autoethnography for librarians and information scientists*, pp: 79–91. London: Routledge.
- Haslam, S. A. and Reicher, S.D. (2006) Stressing the group: social identity and the unfolding dynamics of responses to stress. *Journal of applied psychology*, 91(5): 1037-1052.
- Haslam, S. A. and Reicher, S. D. (2007) Beyond the banality of evil: Three dynamics of an interactionist social psychology of tyranny. *Personality and Social Psychology Bulletin*, 33: 615-622.
- Holman Jones, S., Adams, T.E. and Ellis, C. (Eds.) (2013) *Handbook of Autoethnography*. Walnut Creek: Left Coast Press.
- Khedr, A.M., El-Henawy, W. and El Bassuony, J. (2022) From Business to Education: Action Learning in EFL Teacher Preparation Programs. *Port Said Journal of Educational Research*, 1(1): 113-131.
- McGill, I. and Brockbank, A. (2004) *The Action Learning Handbook*. London: Routledge.
- Nakamura, L. (2013) *Cybertypes: Race, ethnicity, and identity on the Internet*. London: Routledge.
- Pedler, M. (2008) *Action Learning for Managers*. Aldershot: Gower.
- Reicher, S. D. and Haslam, S. A. (2006) Rethinking the psychology of tyranny: The BBC Prison Study. *British Journal of Social Psychology*, 45: 1–40.
- Revans, R.W. (1980) *Action Learning: New Techniques for Action Learning*. London: Blond and Briggs.

- Revans, R.W. (1998) *The ABC of Action Learning: empowering managers to act and to learn from action* (3rd edition), London: Lemos & Crane.
- Rogers, C. (1957) The necessary and sufficient condition for therapeutic personality change. *Journal of Consultative Psychology*, 21: 95-103.
- Rogers, C. (1979) *Carl Rogers on Personal Power*. London: Constable.
- Schon, D.A. (1983) *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.
- Shotter, J. (2004) Responsive expression in living bodies: The power of invisible ‘real presences’ within our everyday lives together. *Cultural Studies*, 18(2,3): 443-460.
- Shotter, J. (2005) ‘Inside the moment of managing’: Wittgenstein and the everyday dynamics of our expressive-responsive activities. *Organization studies*, 26(1): 113-135.
- Shotter, J. (2011) *Getting it: Witness Thinking and the Dialogical in Practice*. Hampton Press Communication Series Social Construction in Practice.
- Sparkes, A.C. (2002) Autoethnography: self-indulgence or something more? In *Ethnographically Speaking: Autoethnography, Literature, and Aesthetics*, edited by A.P. Bochner & C. Ellis. Walnut Creek, CA: AltaMira Press: 209–232.
- Sparkes, A.C. 2013. Introduction: autoethnography as a mode of knowing and a way of being. In *Handbook of Autoethnography*, edited by S. Holman Jones, T.E Adams & C.S Ellis. Walnut Creek, CA: Left Coast Press: 511–517.
- Taylor, C and White, S. (2001) Knowledge, Truth and Reflexivity: The Problem of Judgement in Social Work. *Journal of Social Work*. 1(1): 37-59.
- Vince, R. (2004) Action Learning and Organizational Learning: Power, politics and emotion in organizations. *Action Learning: Research & Practice*, 1(1): 63-78.
- Watson, T. J. (2011) Ethnography, reality, and truth: The vital need for studies of ‘how things work’ in organizations and management. *Journal of Management Studies*, 48: 202–217.
- Winkler, I. (2018) Doing autoethnography: facing challenges, taking choices, accepting responsibilities. *Qualitative Inquiry*, 24(4):236–247.
- Yeadon-Lee, A. (2013) Action learning: understanding interpersonal relationships within learning sets. *Journal of Management Development*, 32(9): 984-994.

# Leveraging Generative AI in Information Systems Development

*Research In progress*

## **Abstract**

*The philosophy of Agile development has been hugely influential in the software development and other fields, at the same time, Design Thinking has been widely embraced by many leading companies, promising better user-centred innovation, and design. There are still pragmatic difficulties in making both operational. In this paper, we suggest the use of modern generative AI systems, underpinned by Foundation Models, to help designers operationalise the Agile philosophy, and its manifestation in the Agile process models, and put Design Thinking into practice more cost-effectively and thoroughly by leveraging foundation models, tuned with the target substantive domain, to help in ideating and selecting the most optimum experiments, and in prioritising the selection and roll-out of features, embodied in user stories in the various hierarchical levels.*

**Keywords:** Design Thinking, Agile Development, Generative AI, Foundation Models, Methodology in Systems Development

## **1.0 Introduction**

This paper describes a novel approach to leveraging the huge power of the AI tool known as Foundation Models in operationalising the two popular approaches to systems development and the strongly related digital innovation, namely Agile development, and design thinking. The paper begins by

## **2.0 The Role of Methodology in Information Systems is Still Relevant**

It has been acknowledged for years that information systems developed activities need to follow a methodology (Avison and Fitzgerald, 2006) to maximise the chances of information systems. However, more recently, the spotlight has moved away from methodology per se to modelling paradigms and representations (for example, object-oriented modelling, epitomised by modelling notation such as UML), as well as a host of new “agile” approaches to systems development. We prefer to characterise the latter Agile approaches as process models, rather than afford them the more

substantial “methodology” label. To us the role of methodology is still relevant, we demonstrate this via a quick exposition of our preferred meaning of the term.

### **1.1 Methodology vs Life Cycles**

The first notable thing about the “methodology” term is the “ology” suffix. It means a subject of study or a branch of knowledge, which we take to point at the sense of *knowing why*. Knowing why you do something in a certain way, means that you have a philosophy, or vision, as to why this way of doing things is good or beneficial. Here’s a definition that we advocate:

#### Methodology

There are two features that we believe must exist in a methodology:

Firstly, a methodology offers an integration of a number of tools with a number of techniques and methods for the application of those tools. The tools are largely concerned with modelling (or representing) certain aspects that are seen to be critical or important to the endeavour; the type of system being considered. A methodology *may or may not* prescribe a sequence of activities that suggests a particular life-cycle model. See "Cookbook vs. Toolbox Methodologies" below.

Secondly, and this is central to our view, a methodology should be issue-based. That is, the methodology should be devised to address one or more issues, in systems development, such as ease of maintenance, correctness of code or acceptance by the users, etc. An important corollary of the above is that a methodology should be able to provide guidance on where it can be applied by matching the issues and quality goals (see next section) which the methodology supports, with those associated with the type of system or application to be developed.

Issue: Although the intuitive meaning of this term is usually sufficient, we give slightly more formal definition because of the importance of the concept of a methodology being issue-based. An issue is an acknowledged problem or difficulty that is related to the development process from the technical, business, or human points of view.

Note that issues include system quality goals, such as adaptability, efficiency, or robustness, etc. It could also include issues surrounding the system, such a complex



human activity system or social context that can highly impact the resultant system (or *informational artefact*).

### Life-Cycle Model (LCM)

A life cycle or process model represents a high-level generic plan for the system development process. Its focus is largely on how the development activities are ordered in time and managed. The (in)famous Waterfall model, is but a top-level plan of ordered activities, so it is the life-cycle model. It does not proffer any modelling tools for the informational or technical content of the system. Accordingly approaches that belong under the Agile development paradigm are essentially life cycle (or process) models, but not proper methodologies. We can classify development life cycles (alternatively: process models) into three main categories:

- **Sequential process models:** these are development process models that executes the development phases, or activities, in a linear sequence with iteration and rework being limited to only adjacent phases, as in when obvious errors are discovered in requirements, so those are fixed before moving to the next (design) phase.
- **Non-sequential process models:** these are models that divide the systems development up into increments, deliveries or
- **Interleaved process models:** this is process models that mix the above two styles, as in the case of a waterfall process, that is front ended with a prototyping phase (as in experimental or throwaway prototyping). Another example is the DevOp model (for Development-Operations) that moves from development to operations and back in a way that maximises communications between the development and operations team. The Spiral model could also be considered as an interleaved process model, as it is capable of accommodating other process models, including Prototyping and the Waterfall (Boehm, 1988).

One of the approaches that gained much interest recently in the areas of innovation and new product development is design thinking, which has proven applicability in several domains, including information systems design. As such, design thinking offers a philosophy that can benefit a range of information systems types that require innovation and user-centric design.

### **Design Thinking -DT**

The Design Thinking process is a process that is centered around close user knowledge and is basically defined by the stages of: (i) empathize, (ii) define, (iii) ideate, (iv) prototype, and (v) test, (Interaction, 2023). During the empathize phase, the focus is on gaining knowledge of the users' needs and challenges through field research methods like interviews and/or observations. Following empathizing, is the

*define* phase, where a problem is identified and more sharply framed. In the *ideation* phase, the team and users come up with possible solutions to solve the problem identified in the previous stage. One or more of the *prototypes* is selected to be developed and then *tested*, with the testing and evaluation results being fed into any of the earlier phases on the DT process, and to modify the prototype(s) being tested or to select another. A core issue in the define phase is to frame the problem properly using all the data available, possibly through ethnographic style research. One important decision in the prototype selection is: which of the possible generated ideas or experiments would yield the highest learning. An exploratory study was conducted by (Souza et al. 2021) to investigate how Design Thinking benefits the development of a data-driven requirements elicitation tool. Their results show that DT can be incorporated along with development to help provide a better understanding of the problem and solutions that are more user-centered. The research done by (Canedo et al., 2020) observed that DT accords with the improvement of requirements elicitation process and that, through prototyping, aids to identifying errors in requirements understanding prior to implementation. Research by (Micheli et al., 2019) found that Design Thinking has attracted significant interest from practitioners and academics, as it offers a novel approach to innovation and problem-solving. Requirements Engineering can be seen as an iterative process that is performed all the way through the software development process instead of a single phase at the beginning of the project, which makes this activity in tune with a design thinking attitude. Also, the research by (Corral et al., 2018) presented a study that maps the two recognizable methodologies in software engineering, Agile and Design Thinking; they assessed methods based on artefacts and final products delivered by students in their learning path. Agile development was shown to be more rigorous in strict software engineering aspects, while Design Thinking led to more outside-the-box products. Our proposal is developing a problem domain model will help the designers make better decisions about the requirements, in a user-focused way, and in prioritizing them, as well support the generation of ideas as well as in assessing and prioritizing them. A domain model has been proposed before to be developed using qualitative data analysis (Galal-Edeen, 2005). However, conventional qualitative data analysis is time consuming and is sensitive to the researchers carrying out the data collection and analysis.

## Foundation Models

We are exploring the utility of the concept of a Foundation Model in resolving some of the issues pointed at above. Foundation models are one of the novel artificial intelligence paradigms, developed via training artificial neural network on massive amounts of broad data of different types (text, images, graphs, audio, etc.) to learn the latent semantics and structure of a corpus of data. They can support the generation of new text and other output such classification, summarization, story generation and question answering, based on what they have “learnt” from the data, this is self-supervised learning that doesn’t require intervention in the early stages. These models can be specialized (fine-tuned) to serve narrower domains of knowledge and related practice (Bommasani et al., 2021).

Using a suitable foundation model, the question (or more appropriately: *prompt*) could be posed as to what are the core problems or pain points that the user community of a certain domain or artifact. It seems plausible that such models can aid the designers greatly in sharpening up their problem framing effort by putting the distillation of a very large amount of data at their fingertips. The output can also provoke the imagination of the designers during the ideation by proposing new combinations of ideas distilled from the data.

To give an example, when Google’s Bard (a Generative AI system) was prompted as follows:

Prompt: **suggest requirements for a mobile app that can help patient compliance with treatment programs**

The response was: Here are some suggested requirements for a mobile app that can help patient compliance with treatment programs:

- User-friendly interface: The app should be easy to use and navigate, even for patients with limited technical experience.
- Personalized reminders: The app should allow patients to set personalized reminders for taking medications, completing exercises, or attending appointments.
- Educational resources: The app should provide patients with educational resources about their condition and treatment program. This could include information about the benefits of

<p>compliance, how to manage side effects, and how to live a healthy lifestyle.</p> <ul style="list-style-type: none"> <li>• Tracking and progress reporting: The app should allow patients to track their progress and generate reports that they can share with their healthcare providers. This could help patients to stay motivated and accountable.</li> <li>• Social support: The app could also include features to help patients connect with other patients who have similar conditions or treatment programs. This could provide emotional support and motivation.</li> </ul>
<p>Additional features that could be considered include:</p> <ul style="list-style-type: none"> <li>• Integration with electronic health records (EHRs): This would allow the app to access and share patient data with healthcare providers, which could improve communication and coordination of care.</li> <li>• Medication refills: The app could allow patients to request medication refills directly from their pharmacy.</li> <li>• Symptom tracking: The app could allow patients to track their symptoms and share this information with their healthcare providers. This could help providers to identify and address potential problems early.</li> <li>• Gamification: The app could use gamification elements to make it more engaging and motivating for patients. This could include features such as badges, leaderboards, and rewards.</li> </ul>

**Table 1. An example of a prompt and its response to a Generative AI system to help with requirements ideating for a mobile application to aid patients comply with a treatment regimen.**

### **Agile Systems Development**

Agile development is a highly popular process model for software and systems development, as well as many other endeavors such as general project management, although it has been proposed based on a strong push from the latter. Agile development embodies a specific philosophy, practices, and values to address rapidly changing or incomplete user requirements. Its philosophical stance favors working software (or systems) over comprehensive documentation, individuals and interactions over processes and tools, collaboration with the customer over contract negotiation, and responding to change over following a plan. Typically, projects that purport to follow agile divide the available requirements to small chunks, usually referred to “user stories”. The gathered requirements, in form of user stories, are prioritised and a number of those is selected for implementation in the next iteration (termed “sprints”). Usually, the prioritisation of user stories is done according to one of a number of approaches such as MoSCoW, on Business Value, etc., they all have issues and none is totally satisfactory (see Popli et al., 2014). The method proposed in Popli et al. (2014) is to combine various methods in the literature, but it lacks a way of



The salient concepts in the diagram (representing a preliminary domain architecture, that can be layered on top of a suitably chosen domain model), can guide the prioritisation of user stories by establishing the degree of cohesion of the user stories with the most salient domain concepts<sup>1</sup>. Also, by juxtaposing user stories against a qualitative model of the domain, user stories are more grounded and “given voice” in a way that facilitates choice -in each sprint or iteration- thus making its justification more visible.

## Conclusions

This paper outlined the relevance of Design Thinking and Agile Development to information systems development, and proposed how modern generative AI systems may support ideation, idea and requirements evaluation and prioritisation. This is still research in progress, and we plan to explore how suitable Foundation Models may be selected and tuned to serve in the ways that we propose in this paper; and whether the modality of such support can be altered to aid the operationalisation of design thinking and agile development principles in complex information systems development settings.

## References

- Boehm, B. W. (1988). “A Spiral Model for Software Development and Enhancement.” *Computer*(May 1988): 61-72.
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., ... & Liang, P. (2021). On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258.
- Brown (2008) ‘Design Thinking.’, *Harvard business review.*, 86(6), pp. 84–92.  
Available at: <https://doi.org/info:doi/>.
- Canedo, Edna Dias, Ana Carolina dos Santos Pergentino, Angelica Toffano Seidel Calazans, Frederico Viana Almeida, Pedro Henrique Teixeira Costa, and Fernanda Lima. 2020. “Design Thinking Use in Agile Software Projects: Software Developers’ Perception.” In *ICEIS 2020 - Proceedings of the 22nd International Conference on Enterprise Information Systems*, 2:217–24. SciTePress. <https://doi.org/10.5220/0009387502170224>.

---

<sup>1</sup> This is akin to the concept of urban grid integration measure, as expounded by the Space Syntax method for studying the morphological nature of urban and architectural spaces (Hillier, 1996). But space limitation prevents us from expanding on such concepts here.

- Deng, B. & Jia, K. 2023, *Universal Domain Adaptation from Foundation Models*, Cornell University Library, arXiv.org, Ithaca.
- Galal-Edeen, G. H., (2005) Information Systems Requirements Engineering: An Interpretive Approach, *The Egyptian Informatics Journal* Vol. 6 No.2. Pp.154-174.
- Global Business Review (2022) “How smarter AI will change creativity: Foundation models”.
- Hillier, B., *Space is the machine - A configurational theory of architecture*. 1996, Cambridge: Cambridge University Press. 463.
- Interaction Design (2023) <https://www.interaction-design.org/literature/topics/design-thinking>
- Micheli, Pietro, Sarah J.S. Wilner, Sabeen Hussain Bhatti, Matteo Mura, and Michael B. Beverland. (2019) “Doing Design Thinking: Conceptual Review, Synthesis, and Research Agenda.” *Journal of Product Innovation Management* 36 (2): 124–48. <https://doi.org/10.1111/jpim.12466>.
- Popli, R., Chauhan, N., and Sharma, H. (2014) "Prioritising user stories in agile environment," 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), Ghaziabad, India, pp. 515-519, doi: 10.1109/ICICT.2014.6781336.
- Probert, S.K. (2002) *Relevance and Rigour, Theory and Practice in Systems*. In: Ragsdell, G., West, D., Wilby, J. (eds) *Systems Theory and Practice in the Knowledge Age*. Springer, Boston, MA. [https://doi.org/10.1007/978-1-4615-0601-0\\_4](https://doi.org/10.1007/978-1-4615-0601-0_4)
- Souza Filho, José de, Walter Nakamura, Lígia Teixeira, Rógenis da Silva, Bruno Gadelha, and Tayana Conte. 2021. “Towards a Data-Driven Requirements Elicitation Tool through the Lens of Design Thinking.” In: 283–90. Scitepress. <https://doi.org/10.5220/0010443402830290>.

# Virtual Representations as Boundary Objects: A Case of Building Information Modelling

*Research In Progress*

## **Abstract**

*An increasing body of research on cross-boundary collaboration is investigating how digital artefacts (e.g., 3D representations, data visualisation) when viewed as boundary objects, influence group dynamics and collaborative performance. Central to this conversation is the inter-team and interpersonal communications mediated by such digital artefacts. We investigate large-scale construction projects based on Building Information Modelling (BIM), where multiple stakeholders collaborate across a variety of organisational and disciplinary boundaries. BIM is a process-based virtual representation shared by project team members in their collaboration. The promise of its efficacy as a boundary object lies within its capability of affording 'model-oriented' communications. To understand this process, we draw on the literature of semiotics to analyse how virtual representations are used to signify objects, meanings, and processes during communication practices. Our findings suggest that virtual representations bear multiple semiotic natures that allow them to become effective communicative boundary objects, affording diverse ways to convey meaning among project teams. This theoretical alignment demonstrates a robust explanatory mechanism when applied to the context of multi-team cross-boundary collaboration.*

**Keywords:** virtual representation, boundary objects, mediated communication, semiotics, cross-boundary collaboration, Building Information Modelling, digital artefacts

## **1. Introduction**

Virtual representation uses digital artifacts to present visual information, approximating physical objects. The focus has recently shifted toward creating digital twins and metaverses, especially in the construction industry. Building Information Models (BIM) are instrumental in crafting these digital environments. Narrowly defined, BIM is a modelling tool that consolidates various building details (geomatic and nongeomatic) into an integrated digital model. Broadly, its advanced visualization and analytical features transform the collaborative and functional roles of participating organizations and disciplines within the collaboration process (Volk et al., 2014). The digital world, as rendered by the digital model, enables comprehensive imagination and strategic arrangement of the physical world (Deng et al., 2021), which co-creates the digital twin process. Model-based virtual representation aids in both sense-making and future planning (Boland et al., 2007; Comi & Whyte, 2018). Employing digital models for representation is complex, requiring process-based integration and the



establishment of collective sense-making and expectations (Miettinen & Paavola, 2018). Digital models frequently act as boundary objects in practice, promoting information exchange, negotiation, and knowledge generation within project teams. Current research investigates digital models as boundary objects in communication and collaboration to enhance cross-boundary collaboration (Neff, 2010). However, few studies have pinpointed the role of model-based virtual representation in collaborative efforts. Recent studies have highlighted the functional affordances that digital artifacts, as boundary objects, provide in communication (Leonardi, 2019).

Rapid technological changes in the Architecture, Engineering, and Construction (AEC) industry pose challenges for project management (Travaglini et al., 2014). Digital collaboration may offer solutions, yet the disruptive nature of new digital technologies can create complications. Contrary to expectations that digital models enhance visualisation, they may intensify rather than mitigate the complexities of communication. For instance, increased visualisation could diminish interpretative flexibility among team members from diverse backgrounds (Neff et al., 2010). In this environment, the digital model, as a virtual representation, might function distinctively as a boundary object within project-based collaboration (Succar, 2009; Arayici et al., 2011). The semiotic significance of the BIM model is vital for collaboration within construction project teams, as members from varied knowledge domains collaborate across their knowledge boundaries. BIM technology's visualisation alters communication methods among these groups, with the efficiency and effectiveness of shared meaning conveyed by the visualised building information being key to collaborative performance (Neff, 2010). Digital models in the AEC industry aim to enhance knowledge management across project boundaries, with communication being affected differently in various scenarios involving these models. Recent socio-technical studies underscore the critical role of communication dynamics among project teams. This paper delves into the role of virtual representation of BIM models as boundary objects within project-based collaboration. It addresses the research question: How do virtual representations facilitate various communication methods to support cross-boundary collaboration? The study applies semiotics to conceptualise BIM, identifying different types of virtual representations of BIM models and their implications as boundary objects. A case study of BIM-enabled collaboration is conducted to examine how BIM models function as boundary objects between teams and organisations.

## **2. Literature Review**

### **2.1 Virtual representation and boundary object**

Few studies have sought to understand how virtual representation functions practically as boundary objects to convey shared understanding. Elucidating this can enhance our comprehension of the use of boundary objects in practice. Bailey et al. (2012, p. 1485) note, "Virtuality occurs when digital representations stand for, or in some cases completely substitute for, physical objects, processes, or people they represent." Virtual representation is the digital embodiment of objects, processes, or systems that aids in understanding, manipulation, and communication (Carlile, 2002). Serving as a type of boundary object (Star and Griesemer, 1989), it mediates among diverse groups, facilitating meaningful interaction. These representations range from simple 2D diagrams to intricate multi-dimensional models and act as common reference points for discussions and decision-making in collaborative environments (Orlikowski, 2007; Levina and Vaast, 2005). The role of virtual models in collaborative contexts is recognized for enhancing communication efficiency among different groups, especially as a boundary object that bridges the sensemaking of these groups.

The concept of a boundary object was proposed by Star and Griesemer (1989) and refers to an object which can bridge interaction among different groups with its plastic structure to localise each group's needs and its flexibility to build a common identity across groups. In other words, the boundary object serves as a translation or interpretation tool for communication and understanding, facilitating interaction among individuals, or group members, originating from various social worlds, who have different viewpoints and interests (Uppström & Lönn, 2017). Boundary objects serve as repositories that contain or link to multiple interpretations and meanings across diverse communities of practice (Bechky, 2003). Carlile (2004) proposes that boundary objects can be utilised through three processes to manage these knowledge boundaries: 1) transferring, which facilitates knowledge sharing across different knowledge boundaries by establishing a shared language for people in different communities to present their specific knowledge at the syntactic knowledge boundary; 2) translating, which allows actors from different communities to have a common understanding of the knowledge at the semantic knowledge boundary; and 3) transforming, which empowers actors to transform their interests and develop from their previous knowledge to the shared needs at the pragmatic boundary depends on the novelty of the shared knowledge and the knowledge background of the entities on either side of the boundary.

Boundary objects manifest a fruitful perspective in exploring the nature of digital technology-enabled communication and collaboration (Papadonikolaki et al., 2019; Leonardi et al., 2019). Recent research indicates that a boundary object is not stable but exhibits emerging dynamics in practice. For example, digital artefacts transforming into boundary objects show that the boundary object is developed in ongoing practice (Levina & Vaast, 2005). The object is perceived to afford the expectation of goals that become the boundary objects in practice across boundaries (Leonardi et al., 2019). 3D model-based virtual representation is recognised as boundary objects across team members for simulating the building information at different stages of the project and contributing to project teams with both roles of symbolic and semantic objects (Meyer et al., 2018). BIM is not merely a 3D model but a comprehensive virtual representation that consists of a wide array of information relevant to different stakeholders, from architects and engineers to contractors and clients. It enables these disparate groups to collaborate effectively by providing a shared but malleable space where information can be added, modified, or viewed according to the specific needs of each stakeholder (Sacks et al., 2010). Therefore, the boundary object role of BIM models in facilitating multidisciplinary and multi-organisational collaboration manifests in its nature as a virtual representation (Succar, 2009; Arayici et al., 2011).

## **2.2 Virtual Representations as Signs in Communication**

Virtual representations mediate communication through their capabilities of signifying. This understanding aligns with the concept of signs in traditional semiotics, where semioticians argue that through various mechanisms of signification, signs play a quintessential role in communicative activities (Eco, 1976). Noted by Frege (1892/1980), a sign can signify a definitive referent when it points to a physical entity, or it can also signify an abstract meaning when it is interpreted subjectively. Peirce (1974) further classifies signs in their distinct ways of signifying, namely as indices, icons and symbols. An overview of these concepts and their implication in communication is as follows.

Indices are signs that have a direct physical relationship with the objects they represent. This signification of the existential link between the sign and the referent could facilitate an understanding of their sequential and causal co-presence. In communication, they indicate a processual logic that helps demonstrate operational consequences to the receiver. (e.g., simulation of conflicting components in a construction model exhibits a direct indication of

operational infeasibility.) Icons are based on resemblance or similarity to the objects they represent. Such resemblance is often physical or visual, making icons self-explanatory to a broad audience irrespective of cultural or linguistic differences (Peirce, 1974). This nature of icons indicates an immediate recognisability in communication. (e.g., a blueprint representing a future building visually presents the layout and structure of the intended construction). Symbols operate on convention, agreement, or learned associations rather than resemblance or causal relationships. They emphasise the necessity of a social context such as a cultural or community of practice for their interpretation (Saussure, 1916, Lave & Wenger, 1991). With this nature, they often typify the representation of abstract ideas or meanings. (e.g., a financial report may represent value creation within business conventions via learned notations such as numbers and terms.). This theoretical perspective provides us with an instrumental lens, through which the communicative affordances of virtual representation are analysed.

### **3. Research settings and methods**

Our research setting is a large-scale construction project in China, which employs BIM (Building Information Modelling) as a core technology. Contextually, BIM technology is strongly backed by the Chinese government and AEC institutions, which makes its adoption increasingly prevalent in construction projects at large urban centres like Beijing and Chengdu. Guided by literature, we take advantage of a revelatory single-case study (Yin, 2009). The case for this research was selected due to its revelatory nature that addresses our unique research objectives. A key criterion was the participation of organisations with three to five years of BIM-enabled project experience. This criterion ensured the project was overseen by an experienced body, epitomising a benchmark BIM initiative. Preliminary data was collected through the lead researcher's disciplinary connections with the management of construction projects and key roles in AEC institutions. This network streamlined the process of pinpointing potential organisational stakeholders and moulding an embedded case study design. Projects were chosen based on contracts stipulating BIM technology utilisation and comprehensive BIM process deliverables. The embedded design accentuates organisations pivotal across the diverse phases of the BIM-enabled project. The fieldwork primarily took place in Chengdu, for engaging with the general contractor organisation, and in Beijing, for the design and sub-construction organisations. Further data collection was conducted via digital media from the proprietor organisation.

The case study aimed to unravel the virtual representation of BIM models in communication among project team stakeholders. The primary data source was semi-structured interviews, comprising 27 participants from diverse construction project roles (Table 1). Other data emerged from observations, focusing on real-time collaboration in meetings, especially BIM-centric activities, and project documents, including drawings, standards, and pictures of collaborative activities. The results from our analysis of the different semiotic nature of BIM models are presented in the daily communication among construction projects. To reflect these different types of virtual representations, we present our preliminary findings based on reflective thematic analysis from Braun and Clack (2020) and we show the findings in three vignettes based on different boundaries appearing within the construction project-based collaboration.

Table1: The involved participants, their organisations and their roles within construction project teams

<b>Organisation</b>	<b>Participants and role in the construction project team</b>
The client organisation	Project manager
	BIM consultant
	Design consultant
	Onsite manager
	Estimator manager
Construction organisation	Project manager
	BIM supporter
	Security manager
	Onsite manager
	Technical engineer
	Quality manager
	Construction technician
Sub-construction organisation	Project manager
	MEP engineer
	BIM technician
	BIM managers
	Technical engineer

	Onsite manager
	Construction technician
	Installation engineer
	Estimator
Designing organisation	Architect designer
	Engineering designer
	BIM managers
	BIM designer
	BIM executive
	Project Assistant

#### 4. Preliminary findings and discussion

The analysis guides us further to consider the communicative process of virtual representation used in project collaboration. We provide a brief demonstration in the following vignettes of our findings.

##### **Vignette 1: BIM model as an icon – crossing the disciplinary boundary**

During the project, the 3D model is frequently used to ‘visualise’ a technical problem. The following quote illustrates how sharing the BIM model facilitates communication between technical roles from different disciplines, aiding in problem-solving.

*The general practice is that when I encounter a problem and need to consult with others for a solution, I show them [the client] the [BIM] model. Showing the model gives them an ‘intuitive view’. However, they primarily rely on their own 2D drawings. After I provide the model, it's to help them decide on a solution, especially when a certain aspect is not feasible. The [BIM] model gives them an immediate visual impression that confirms the problem. What they see is considered more intuitive, although it doesn't differ much from reviewing drawings. When I communicate with them this way, it's not the case that we ask them to review the model; they only look at it for a general understanding. (BIM technician, general construction team)*

Between the general construction team and the client, the boundary lies within their differing expertise and knowledge. The BIM model in this instance serves as an iconic representation, by providing an immediate and intuitive view of the problem at hand and affords demonstrative and explanatory communication. Through the realisation of such affordances, expert technicians can bridge knowledge boundaries between disciplines. This aligned with the recent assertions that virtual representation can effectively be a boundary object to facilitate process and product-oriented communication (Wlazlak et al., 2019). However, it extends the current understanding of virtual representation by demonstrating its iconic roles in transferring different disciplinary knowledge.

### **Vignette 2: BIM technology as a symbol – crossing the organisational boundary**

Technicians working in the general construction team often mention their deliberate reference to the BIM model in their communication with clients. This quote demonstrates that this practice can sometimes become excessive in extreme cases.

*“They [the clients] don't care [about reporting with the BIM model], but I have to. It's part of my job. I need to show it [BIM model] to him just as he has the right to know. For example, when he [the client representative] asks for a report, I can provide it. When they want to know the status of my current progress, I could show him, right? Even if he doesn't ask, I would still show him. Since they've invested so much in BIM, I must make its value apparent. It reassures the client about our methods, showing them we're scientific and can make technology serve the actual project. Our ultimate goal is value creation, regardless of the means.” (BIM technician, general construction team)*

In this instance, BIM is referred to as an advanced technology, which in its appearance embodies the scientific, methodical approach to construction. By symbolising a shared goal of value creation embedded in the multi-organisational project, BIM thus can afford negotiatory interactions, in constructors' attempt to align with client stakeholders across organisational boundaries. This echoes the recent findings that digital technology contributes value-cocreation through facilitating inter-organizational collaboration across knowledge boundaries semantically and syntactically (Shi et al., 2023). Furthermore, it enriches the understanding of pragmatic interests across different organisational perspectives.

### **Vignettes 3: BIM simulation as index – crossing the operational boundary**

When different components (e.g., architectural design, structural design, piping design) are integrated into the BIM model, a simulation is run to visualise conflicts and problems. Upon sharing the simulated model, technicians from different design teams propose operational solutions by making change suggestions to the model. The following is a demonstration of this process.

*“In the coordination and communication work we do, they might propose a solution [to the simulated problem] today, asking you to adjust the BIM model accordingly. Tomorrow they may change yesterday's solution. How they notify you about these changes is usually through a phone call first to discuss the changes. Once they propose an idea, we have to adjust it in the model. However, the model is basic, not detailed. It*

*doesn't mean that every time they propose something, we need to modify the [BIM] model significantly. It's part of our daily coordination. If a plan is firmly decided, then we'll make the necessary changes. If it's just minor local adjustments, our model generally remains unchanged. If they propose something on-site, they will send an official document. After making the changes, they will send a document specifying design alterations to be made on-site. Sometimes, after they have an idea and deem it suitable, they'll send an official document detailing the changes. We, as the construction unit, will then proceed according to their changes, provided they are reasonable.” (project manager, sub-construction team)*

Construction projects are carried out by divided labour among various compositing teams. These teams are formed based on their operational function, implying boundaries when orchestration attempts fail. When a BIM model is created for simulation (as inputs from different teams are compiled together), it exposes the causal relationship between the future-projecting model and an infeasible operational process. In this sense, it indexes a processual incompatibility in operating teams. Later, the proposals and feedback from different teams again indicate the indexical nature of BIM simulation in reflecting a causal relation between change in model and change in operation. BIM's affordances here become apparent in facilitating coordinated action and orchestrating communication across operational boundaries.



## **5. Conclusion**

Our synthesis of the literature on virtual representation and signification extends traditional semiotics conceptualisation into the area of communication practices. We argue that virtual representations bear multiple semiotic natures which suggest their versatility in conveying meanings which, upon recognition, afford goal-oriented communicants' different ways of communication. This theoretical alignment demonstrates a robust explanatory mechanism when applied to the context of multi-team cross-boundary communication and collaboration. We also contribute to the understanding of boundary objects by specifying their communicative affordances among collaborators. Our future research will build on preliminary analysis and findings, delving into data analysis to discern how iconic, symbolic, and indexical representations of BIM models affect perceptions of virtual BIM representations in cross-boundary collaboration. The subsequent analysis aims to enhance our comprehension of digital artifacts as dynamic boundary objects and the capabilities of model-based virtual representations. By delineating the varied roles of the BIM model across communication scenarios, this study clarifies how virtual representation imparts distinct semiotic meanings in object-oriented collaboration and communication, thereby informing the ongoing discourse on the function of boundary objects in practice.

## References

- Arayici, Y., Coates, P., Koskela, L., Kagioglou, M., Usher, C., & O'Reilly, K. (2011). Technology adoption in the BIM implementation for lean architectural practice. *Automation in Construction*, 20(2), 189–195. <https://doi.org/10.1016/j.autcon.2010.09.016>
- Bailey, D. E., Leonardi, P. M., & Barley, S. R. (2012). The Lure of the Virtual. *Organization Science*, 23(5), 1485–1504. <https://doi.org/10.1287/orsc.1110.0703>
- Bechky, B. A. (2003). Sharing Meaning Across Occupational Communities: The Transformation of Understanding on a Production Floor. *Organization Science*, 14(3), 312–330. <https://doi.org/10.1287/orsc.14.3.312.15162>
- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE.
- Boland, R. J., Lyytinen, K., & Yoo, Y. (2007). Wakes of Innovation in Project Networks: The Case of Digital 3-D Representations in Architecture, Engineering, and Construction. *Organization Science*, 18(4), 631–647. <https://doi.org/10.1287/orsc.1070.0304>
- Carlile, P. R. (2002). A Pragmatic View of Knowledge and Boundaries: Boundary Objects in New Product Development. *Organization Science*, 13(4), 442–455. <https://doi.org/10.1287/orsc.13.4.442.2953>
- Carlile, P. R. (2004). Transferring, Translating, and Transforming: An Integrative Framework for Managing Knowledge Across Boundaries. *Organization Science*, 15(5), 555–568. <https://doi.org/10.1287/orsc.1040.0094>
- Comi, A., & Whyte, J. (2018). Future Making and Visual Artefacts: An Ethnographic Study of a Design Project. *Organization Studies*, 39(8), 1055–1083. <https://doi.org/10.1177/0170840617717094>
- Deng, M., Menassa, C. C., & Kamat, V. R. (2021). From BIM to digital twins: A systematic review of the evolution of intelligent building representations in the AEC-FM industry. *Journal of Information Technology in Construction*, 26, 58–83. Scopus. <https://doi.org/10.36680/J.ITCON.2021.005>
- Eco, U. (1976). *A Theory of Semiotics*. Indiana University Press.
- Miettinen, R., & Paavola, S. (2018). Reconceptualizing object construction: The dynamics of Building Information Modelling in construction design. *Information Systems Journal*, 28(3), 516–531. <https://doi.org/10.1111/isj.12125>

- Frege, G. (1892/1980). On sense and reference. P. Geach, M. Black (Eds.). *Translations from the Philosophical Writings of Gottlob Frege*. Blackwell, Oxford, UK, 25–50.
- Faraj, S., & Azad, B. (2012). The Materiality of Technology: An Affordance Perspective. In *Materiality and Organizing: Social Interaction in a Technological World* (pp. 237–258). <https://doi.org/10.1093/acprof:oso/9780199664054.003.0012>
- Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.
- Levina, N., & Vaast, E. (2005). The Emergence of Boundary Spanning Competence in Practice: Implications for Implementation and Use of Information Systems. *MIS Quarterly*, 29(2), 335–363. JSTOR. <https://doi.org/10.2307/25148682>
- Leonardi, P. M., Bailey, D. E., & Pierce, C. S. (2019). The Coevolution of Objects and Boundaries over Time: Materiality, Affordances, and Boundary Salience. *Information Systems Research*, 30(2), 665–686. <https://doi.org/10.1287/isre.2018.0822>
- Meyer, R. E., Jancsary, D., Höllerer, M. A., & Boxenbaum, E. (2018). The Role of Verbal and Visual Text in the Process of Institutionalization. *Academy of Management Review*, 43(3), 392–418. <https://doi.org/10.5465/amr.2014.0301>
- Neff, G., Fiore-Silfvast, B., & Dossick, C. S. (2010). A Case Study of the Failure of Digital Communication to Cross Knowledge Boundaries in Virtual Construction. *Information, Communication & Society*, 13(4), 556–573. <https://doi.org/10.1080/13691181003645970>
- Orlikowski, W. J. (2007). Sociomaterial Practices: Exploring Technology at Work. *Organization Studies*, 28(9), 1435–1448. <https://doi.org/10.1177/0170840607081138>
- Papadonikolaki, E., van Oel, C., & Kagioglou, M. (2019). Organising and Managing Boundaries: A structural view of collaboration with Building Information Modelling (BIM). *International Journal of Project Management*, 37(3), 378–394. <https://doi.org/10.1016/j.ijproman.2019.01.010>
- Peirce, C. S. (1974). *Collected Papers of Charles Sanders Peirce*. Harvard University Press.
- Sacks, R., Eastman, C., Lee, G., & Teicholz, P. (2018). *BIM Handbook: A Guide to Building Information Modeling for Owners, Designers, Engineers, Contractors, and Facility Managers*. John Wiley & Sons.
- Saussure, F. de. (1916). *Course in General Linguistics*. Open Court Publishing.

- Star, S. L., & Griesemer, J. R. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19(3), 387–420. JSTOR.
- Shi, Y., Cui, T., & Kurnia, S. (2023). Value co-creation for digital innovation: An interorganizational boundary-spanning perspective. *Information & Management*, 60(5), 103817. <https://doi.org/10.1016/j.im.2023.103817>
- Succar, B. (2009). Building information modelling framework: A research and delivery foundation for industry stakeholders. *Automation in Construction*, 18(3), 357–375. <https://doi.org/10.1016/j.autcon.2008.10.003>
- Travaglini, A., Radujković, M., & Mancini, M. (2014). Building Information Modelling (BIM) and Project Management: A Stakeholders Perspective. *Organization, Technology & Management in Construction : An International Journal*, 6(2), 1001–1008. <https://doi.org/10.5592/otmcj.2014.2.8>
- Uppström, E., & Lönn, C.-M. (2017). Explaining value co-creation and co-destruction in e-government using boundary object theory. *Government Information Quarterly*, 34(3), 406–420. <https://doi.org/10.1016/j.giq.2017.08.001>
- Wlazlak, P., Eriksson, Y., Johansson, G., & Ahlin, P. (2019). Visual representations for communication in geographically distributed new product development projects. *Journal of Engineering Design*, 30(8–9), 385–403. <https://doi.org/10.1080/09544828.2019.1661362>
- Yin, R. K. (2009). *Case study research: Design and methods* (Fourth Edition). Los Angeles : SAGE. <https://trove.nla.gov.au/work/11329910>

### **Additional note**

We would like to note that while this manuscript is submitted as a work-in-progress, we intend to complete the paper ahead of the UKAIS 2024 conference. We have included more text than a typical work-in-progress submission in an attempt to provide more details. A shortened draft can be provided upon request.

# Becoming Analytical Champions: A Simple Hermeneutics-inspired Learning Analytics Model for Modules

*Completed Research*

## Abstract

The objective of this paper is to present a Learning Analytics (LA) tool designed for use within modules. The LA tool, from this point referred to as SHLAM<sup>2</sup> (Simple Hermeneutics-inspired Learning Analytics Model for Modules), enables learners to participate in a series of weekly learner self-assessments throughout the duration of a module. This paper reports on the preliminary use of SHLAM<sup>2</sup> as part of a 5-credit research methods module (for doctoral students). The 5-credit module was completed by three doctoral students over a 12-week period (from January to March 2023) in an Irish Higher Education Institution (HEI). The module leaders (co-authors of this paper) are also the co-designers of the SHLAM<sup>2</sup> LA tool. Therefore, the insights presented in this paper are based on the observations of the module leaders as opposed to the learners registered on the module. Our SHLAM<sup>2</sup> insights reveal that 'Seeking Clarity' (through reading and re-reading papers) is the strongest pattern of learning, while 'Building Confidence' (through evaluating an evolving design artefact) and 'Finding Voice' (through peer-led discussion and in-class presentations) are consistently lower throughout the delivery of the 12-week module.

**Keywords:** Hermeneutics, Learning Analytics, Evidence-Based Nudges, Curriculum Design, Feedback.

## 1.0 Introduction

The theory and practice of Learning Analytics (LA) is gaining popularity since the term was first mentioned in 2012 (Leitner et al., 2019). Notwithstanding the fact that LA is still “in its infancy” (Leitner et al., 2017, p.5), it is argued that irrespective of the chosen definition, LA should always “provide actionable insights” to be deemed a success (Leitner et al., 2019, p.2). Despite the growing attention on this interdisciplinary field, there are still significant challenges to implementing effective LA in HEIs (Leitner et al., 2019). For example, one such significant challenge is the integration of the theoretical understanding of learning with the practical development of the LA tools (Kaveri, et al., 2023). As a result, existing LA implementation efforts are sometimes criticised for not focusing on the process of learning and *micro* (learner-centric) patterns but instead targeting larger datasets to find *macro* (institution-centric) patterns (c.f. Guzmán-Valenzuela et al., 2021; Leitner et al., 2017). In fact, this challenge highlights a notable misalignment between the capabilities of the LA tool and the actual needs of the learner (Kaveri, et al., 2023). Furthermore, the “development and implementation of LA” in Higher Education Institutions (HEIs) is “often ad-hoc” and lacks

“replication” by others, given the “lack of standardised design knowledge to guide LA development” (Nguyen *et al.*, 2021, pp.541-542).

In total, Leitner *et al.* (2019) identify six categories of challenges to the implementation of LA (see Table 1). These challenges span across the socio-technical components that define our appreciation of Information Systems, namely: *people, process, technology, and data.*

Challenge	Key Requirement	Summary	Socio-Technical Component
1: Ineffective Leadership	Capability Building	Leadership often lacks the capability to implement LA effectively across institutional environments.	People & Technology
2: Disengaged Stakeholders	Shared Understanding	Understanding of LA varies widely among stakeholders, limiting institutional acceptance.	People & Process
3: Technical Conversation	Pedagogical Practice	LA tool design often overlooks pedagogical approaches, focusing more on technical aspects.	Technology & Process
4: Insufficient Training	Stakeholder Competence	Insufficient training limits understanding of LA's benefits for all stakeholders.	Process & People
5: Limited ROI	Proof of 'Value-in-Use'	Decisions on budget allocation for LA are challenged by difficulties in empirically evaluating its impacts.	Data & People
6: Inappropriate Policy	Clear Guidance	Institutions often lack specific policies and codes of practice for LA.	Process & People

**Table 1. Challenge Categories of LA Implementation (adapted from Leitner *et al.*, 2019).**

Ultimately these challenges can impact on academic staff buy-in: (i) to embrace the collection of learner-centric data and (ii) to use the data to drive change in their teaching practices. Therefore, the LA opportunity for HEIs lies in the fact that academic staff need to see the ‘value-in-use’ when they invest their time into an LA initiative. However, meaningful outputs may not be possible where a HEI adopts a technology-led, pedagogy-agnostic “one size fits all” approach to LA, failing to recognize the diversity of pedagogical approaches that define the learning and teaching space within each HEI (c.f. Kaveri, *et al.*, 2023; Leitner *et al.*, 2019). As a result, our approach responds to these LA implementation challenges by promoting a simple bottom-up approach, enabling academic staff to embed LA (our SHLAM<sup>2</sup> tool) into their modules. Therefore, our approach is in direct contrast to most LA initiatives that adhere to a more top-down roll-out. However, these LA initiatives often fail to gain sufficient traction locally (amongst the academic staff), where data-driven changes in teaching practices become a reality.

Therefore, in this paper we are responding to “*calls to consider how LA should be applied to support learning and teaching activities in higher education*” (Nguyen *et al.*, 2021, p.542). As a result, we leverage the outcome of our scholarly approach to learning and teaching (which started over 10 years ago) and was recently awarded the (*name withheld for the review process*) *President’s Award for Excellence in Teaching (2022)*. Leveraging our experience in educational co-creation and successfully building and implementing multiple pedagogical artefacts, we have designed SHLAM<sup>2</sup>, and believe it has the potential to transform the ‘*what*’, ‘*why*’ and ‘*how*’ of learning-based feedback, where the feedback provided is not reliant on the grading of learners’ submitted assessment(s), but more on the *interim struggles* of learners as they learn over a period of time (e.g. 12-week period of a 5-credit module).

The remainder of this paper is structured as follows. The background to the SHLAM<sup>2</sup> approach is first presented and this is then followed by a description of the SHLAM<sup>2</sup> approach in use. The insights from this instantiation of SHLAM<sup>2</sup> are then presented and patterns of interest discussed. Finally, the paper concludes with key learnings and implications for IS educators.

## 2.0 Background to the Story (*the retrospective*)

In an effort to appreciate if the way we are teaching a 5-credit<sup>1</sup> research methods module (Action Design Research) is effective, we decided to analyse the learning experiences of recent graduates of the module. As a result, in June 2022, we conducted and analysed a conversational account workshop with four graduates (doctoral students) of the module. An in-depth insight into the conversational account is available in (*reference withheld for review purposes*). On completing the first-round analysis of this conversational account, we embarked on a second-round of analysis using the seven concepts underpinning Hermeneutics (c.f. Myers, 2004, 2009). See **Figure 1** for a brief description of these seven concepts. This decision was inspired toward the end of the first-round analysis, where the graduates’ learning stories were most often focused on the time invested in engagement with literature (to make sense of the ADR methodology) and the challenge in effectively sharing

---

<sup>1</sup> 5-credit ECTS represents a notional student workload of 125 hours: [www.study.eu/article/what-is-the-ects-european-credit-transfer-and-accumulation-system](http://www.study.eu/article/what-is-the-ects-european-credit-transfer-and-accumulation-system)

their interpretations (captured in their respective design artefacts – *checklists to evaluate the execution of ADR in empirical studies*). Furthermore, an in-depth insight into the hermeneutics-inspired analysis is available in (*reference withheld for review purposes*).

## hermeneutic concept

<p><b>engagement</b> meaning emerges from the engagement of the reader and the text</p>	<p>hermeneutic circle movement of understanding a text (from the whole to the part and back to the whole)</p>
<p><b>appropriation</b> to understand the meaning of a text, we need to make it our own</p>	
<p><b>autonomization</b> text has an autonomous, 'objective' existence independent of the author</p>	
<p><b>distanciation</b> a text takes on a life of its own, and text is the medium through which we understand ourselves</p>	
<p><b>prejudice</b> prior knowledge plays an important part in understanding a text</p>	
<p><b>historicity</b> historically informed present informs our interpretation of a text</p>	

**Figure 1.** The Seven Hermeneutics Concepts (adapted from: *reference withheld for review purposes*)

As part of our second-round analysis, (using the seven hermeneutic concepts of *historicity*, *the hermeneutic circle*, *prejudice*, *autonomization*, *distanciation*, *appropriation*, and *engagement*, on the confessional accounts of the four doctoral students), we also visualised our understanding of the *hermeneutics inspired learning-by-doing pedagogical approach* as a learning flow (see Figure 2). Specifically, we appreciate that four hermeneutics concepts (*engagement*, *appropriation*, *prejudice*, and *the hermeneutic circle*) are directly linked to the doctoral students' personal efforts at meaning making (their personal *inputs* to the learning-by-doing approach). Furthermore, the other three hermeneutics concepts (*historicity*, *distanciation*, and *autonomization*) are linked to the findings/results of the doctoral students' personal efforts (the *outputs* of the learning-by-doing approach and their willingness to share these outputs).

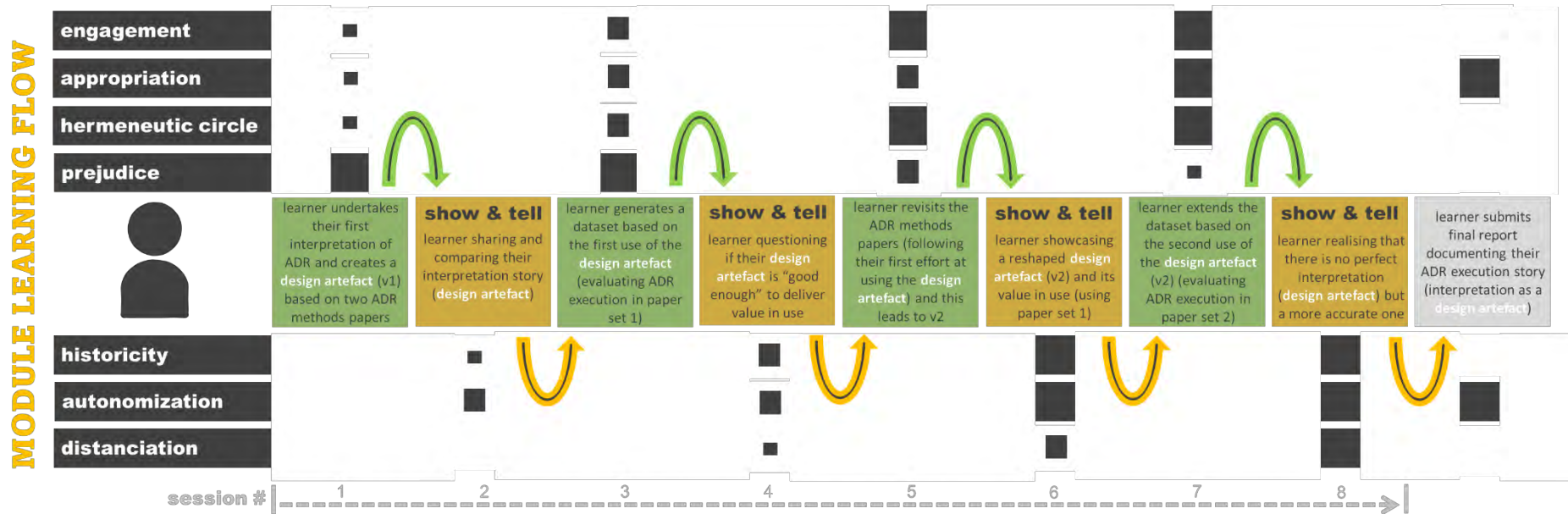


As presented in (*reference withheld for review purposes*), the learning flow visual (see Figure 2) highlights that throughout the 8-sessions the doctoral student's *prejudice* shapes their interpretations less and less, whereas their ever-evolving *historicity* starts to have a bigger influence and part to play in their learning-by-doing. However, *appropriation*, *engagement* and *the hermeneutic circle* move somewhat in tandem over the 8-session period. In particular, the doctoral student's appreciation of the importance of movement between "texts" (e.g. the prescribed literature being reviewed, and the commentary of their peers during the "show & tell" sessions) increases, and their new and improved interpretations are perceived as being more accurate solutions to the problem (*how well documented is the execution of ADR in empirical studies?*). Finally, over time the *distanciation* between the doctoral student and their interpretations (influencing their design artefact) increases, and this affords the doctoral student (learner) the opportunity to critically evaluate their own interpretations as if they were the work of someone else. Furthermore, following the first "show & tell" session the sense of *autonomization* increases, and on submission of the final assignment, the level of *autonomization* between the doctoral student and their interpretations (design artefact) is at its greatest.

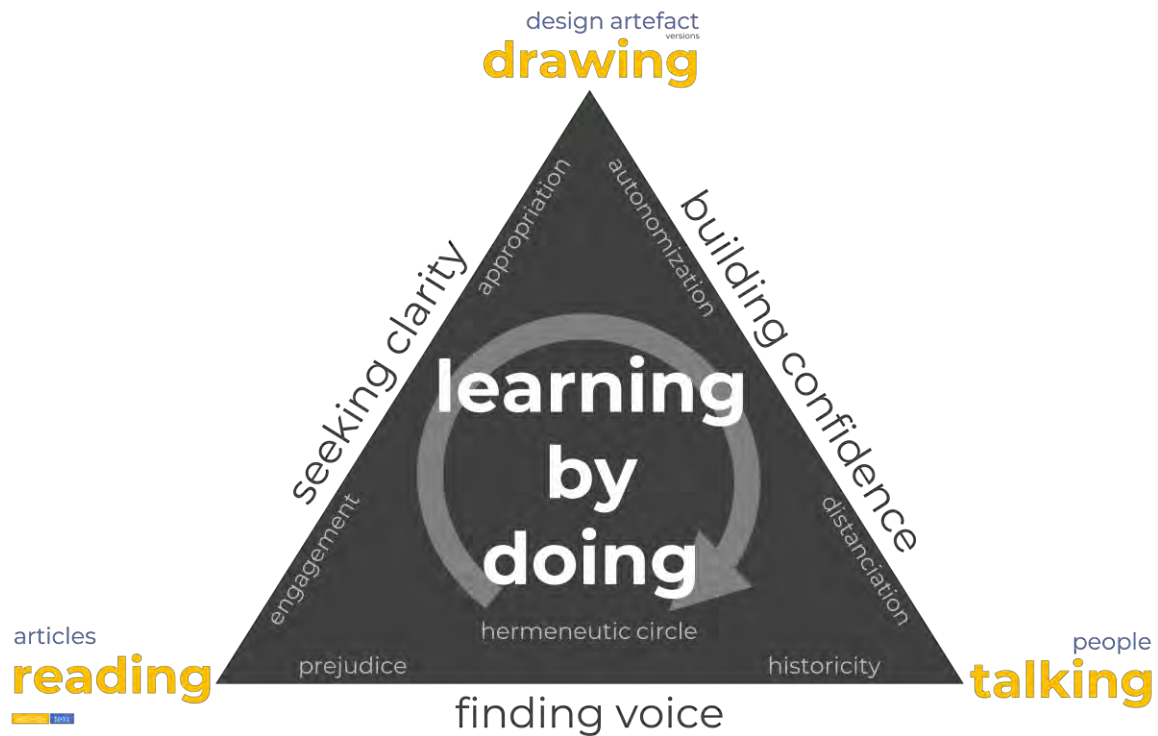
Finally, and most importantly, three patterns emerged from our hermeneutics-inspired analysis. These three patterns are: *Seeking Clarity*, *Finding Voice*, and *Building Confidence* (see Figure 3). Each of these patterns reflects the interrelationships between specific hermeneutic concepts, as follows:

- **Seeking Clarity:** (engagement, appropriation, the hermeneutic circle)
- **Finding Voice:** (prejudice, historicity)
- **Building Confidence:** (distanciation, autonomization)

Furthermore, Figure 3 highlights that reading, drawing, and talking play crucial roles in the learning process, as part of our 'learning-by-doing' pedagogical approach, contributing to different aspects of cognitive development and knowledge acquisition. This espouses a complementary learning approach through incorporating reading (research articles), drawing (the design artefact), and talking (in-class presentations about the design artefact) to encourage social learning and reinforce understanding. Therefore, incorporating such multiple modalities accommodates diverse learning styles, and combining these activities helps to create a well-rounded and engaging learning experience.



**Figure 2.** The ADR Module Learning Flow (a hermeneutics inspired learning-by-doing pedagogical approach) (source: *reference withheld for review purposes*)



**Figure 3.** A Module Curriculum Design Principle

‘Seeking Clarity’ highlights the interplay between *engagement*, *appropriation*, and *the hermeneutic circle*. For example, for *the hermeneutic circle* to work, the learner needs to be actively reading, writing (reflecting on their reading), and interpreting (their writings and the original texts from where their written interpretations come). ‘Building Confidence’ highlights the impact of an environment that forces *distanciation* and *autonomization* between the learner and their design artefact (the shape of which reflects their “current” ADR understanding) over time. For example, in our module design, the four “show & tell” sessions ensure that the learner shares their interpretation, harvests peer feedback, and evaluates the shape of their design artefact continuously. Finally, these four “show & tell” sessions can also be somewhat disarming, as the learner must present their design artefact at each session. Therefore, this “show & tell” discourse is building the shared language of the module group and growing the confidence of each learner to share and challenge their respective perspectives (*historicity* and *prejudice*), thereby ‘Finding Voice’.

### 3.0 The Emergence of SHLAM<sup>2</sup> (the prospective)

Having spent some time theorising around the four learners’ experiences (retrospectively) we turned our attention to exploring the possibility of having such learner-based insights

available to the lecturer (prospectively) during the delivery of the module. It was decided that one way of making this a reality (in a low fidelity way) was to ask questions of the learner (on a weekly basis) around the activities undertaken in that specific week in the name of ‘Seeking Clarity’, ‘Finding Voice’, and ‘Building Confidence’. A full list of the questions designed for this purpose (weekly self-reflection) is provided in Appendix A.

Therefore, the preliminary use of the LA tool (SHLAM<sup>2</sup> - Simple Hermeneutics-inspired Learning Analytics Model for Modules) was designed into the rhythm of the module delivery for the next run of the 5-credit research methods module (Action Design Research). This module run was completed by three doctoral students over a 12-week period (from January to March 2023). Despite this small number of learners, the module leaders were very conscious of the value in being able to track the learning activity patterns, associated with a ‘learning-by-doing’ pedagogical approach (see Figure 4), playing out over the duration of the module. Therefore, an MS Forms version of these questions was circulated to each learner through the Canvas VLE at a similar time each week (Wednesday afternoon) from week two to week eleven of the module. See Table 2 for the questions that map to the three patterns (highlighted in grey in Appendix A).

Hermeneutic Concept	Emergent Pattern	Learning Activity	SHLAM <sup>2</sup> Learning Flow Question
Engagement	Seeking Clarity	Reading	Question 1
Appropriation			Question 2, Question 4*, Question 8, Question 10*, Question 14*, Question 15*
Hermeneutic Circle			Question 7
Prejudice	Finding Voice	Talking	Question 14*, Question 15*
Historicity			Question 3, Question 9
Autonomization	Building Confidence	Drawing	Question 13, Question 14*, Question 15*
Distanciation			Question 4*, Question 10*
* Question provides an insight into more than one hermeneutic concept			

**Table 2.** SHLAM<sup>2</sup> Learning Flow Questions.

In essence, the SHLAM<sup>2</sup> approach delivers learner-generated, learning-specific, near real-time, module-based insights. These analytical insights are visualised week-by-week to ‘see’

and ‘celebrate’ the learner efforts (*interim struggles*), along the flow of learning. Therefore, by analysing the SHLAM<sup>2</sup> questions each week affords the lecturer with the opportunity to appreciate how the learners (as a collective) are finding their way in learning about ADR (in this case). These analytical insights are easily accessible and provide near real-time learning patterns within a module. Hermeneutics, (the process of ‘meaning making’) underpins this SHLAM<sup>2</sup> approach and enables us to unpack the complexities of learning through the three fundamental learning patterns: (i) *seeking clarity on new concepts/content*, (ii) *finding voice through verbalising interpretations*, and (iii) *building confidence by translating interpretations into visual artefacts*.

## 4.0 The SHLAM<sup>2</sup> Insights

In this section we present our insights through unpacking our observations around the learning patterns as part of the ‘learning-by-doing’ pedagogical approach. In fact, we view learning-by-doing as “*the process whereby people make sense of their experiences, especially those experiences in which they actively engage in making things and exploring the world*” (Bruce and Bloch, 2012, p.1821).

### 4.1 Unpacking the Learning Patterns

The SHLAM<sup>2</sup> approach makes the lecturer aware of the *interim struggles* of learners as they learn and enables the lecturer to take action based on these insights. Specifically, the SHLAM<sup>2</sup> approach has two affordances for the module lecturer, as follows: (i) *evidence-based nudges*, and (ii) *week-on-week comparative analysis*. The *evidence-based nudges* create an opportunity for lecturers to prospectively provide weekly direction to the learners that positively impacts on the learner’s journey, based on the lecturer’s interpretation of the learner’s self-assessment of their learning activities. Furthermore, the *week-on-week comparative analysis* creates an opportunity for lecturers to see how the learning flow is materialising throughout the delivery of the module.

Using SHLAM<sup>2</sup> to capture learner-generated data around these three patterns affords both learners and lecturers the opportunity to see and appreciate: (i) a learner’s personal efforts at meaning making, and (ii) the outputs of a learners’ personal efforts, and (iii) their willingness to share these outputs. Therefore, through a series of *evidence-based nudges* and *week-on-week comparative analysis* we are elevating the position of feedback throughout the flow of

learning, at the module level. As an example, **Figure 5** presents a visual of the SHLAM<sup>2</sup> insights across the three patterns (showcasing the week-on-week comparative analysis) while **Table 3** presents **an example of an** evidence-based nudge, **drafted by the lead author** (during iteration 1).

Pattern	Observable Insight	Evidence-Based Nudge
Seeking Clarity	<b>STRONG</b> The percentage of learners reading a 'new' paper, as well as re-reading a previously read paper, finding something new, and making changes to their design artefact (capturing their interpretation).	celebrate those learners continuing to read 'new' papers and encourage more learners to re-read papers they have read previously, as they may produce new or improved interpretations.
Finding Voice	<b>MODERATE</b> The percentage of learners discussing the paper(s) they read/re-read with a peer within and outside of the timetabled sessions.	encourage all learners to discuss their interpretations with peers (outside of the "show & tell" sessions) as it will produce new or improved interpretations and avoid an overreliance on past experiences.
Building Confidence	<b>WEAK</b> 33% of learners shared their design artefact (capturing their interpretation), took onboard a peer comment, but did not make changes to their design artefact.	celebrate those learners taking the risk (at such an early stage) to share their design artefact (for feedback) and encourage more learners to also evaluate their interpretations in this way.

**Table 3.** Evidence Based Nudges (generated from learners' responses - session #2).

For example, as we can see (**Figure 5**) for session two ("show & tell" #1), 'Seeking Clarify' was strong, 'Finding Voice' was moderate, and 'Building Confidence' was weak. Therefore, the *evidence-based nudges* (see **Table 3**) for 'Seeking Clarify' encouraged learners to re-read more, for 'Finding Voice' encouraged learners to discuss interpretations with peers outside in-class sessions, and for 'Building Confidence' encouraged learners to gain more feedback through sharing their design artefacts (and their interpretations). **It is worth noting that the visual presented in Figure 5 should be viewed more as portraying the 'art of the possible' from adopting such a simple LA tool within modules.**

To provide an example of the operational value of the '*week-on-week comparative analysis*' affordance of the SHLAM<sup>2</sup> approach, based on **Figure 5**, we can see that 'Finding Voice' and 'Building Confidence' are continuously moving between a rating of moderate and weak. Therefore, here is a sample narrative (drafted by the lead author) to evaluate the learning flow movement (for a *week-on-week comparative analysis* of week two and week three of the module).

- Overall, the learners have moved slightly this week, when compared to last week, based on the cumulative score of the three patterns ('Seeking Clarity', 'Finding Voice', 'Building Confidence'). Specifically, the move from 6 to 5 is accounted for in the lowered self-assessment around 'Building Confidence'. Therefore, it is very important to encourage learners to share their design artefact (outside of the "show & tell" sessions) as often as they can (so that it can be evaluated by others). It is important to remember that such evaluations are feedback opportunities on their interpretations of Action Design Research (ADR). These feedback opportunities will further enhance their learning experience and highlight what it means to 'be' an ADR researcher as opposed to simply 'do' ADR. Furthermore, 'Finding Voice' needs to be monitored as it remains weak over the past two weeks. Without 'Finding Voice' and 'Building Confidence' the learner will 'never' know if their interpretations of ADR make sense.

Outside of the learnings from the weekly use of the *evidence-based nudges* and the *week-on-week comparative analysis*, the somewhat more strategic value (to-date) from implementing the SHLAM<sup>2</sup> approach is best captured in a hermeneutics-inspired vignette, namely: *the engagement assumption*.

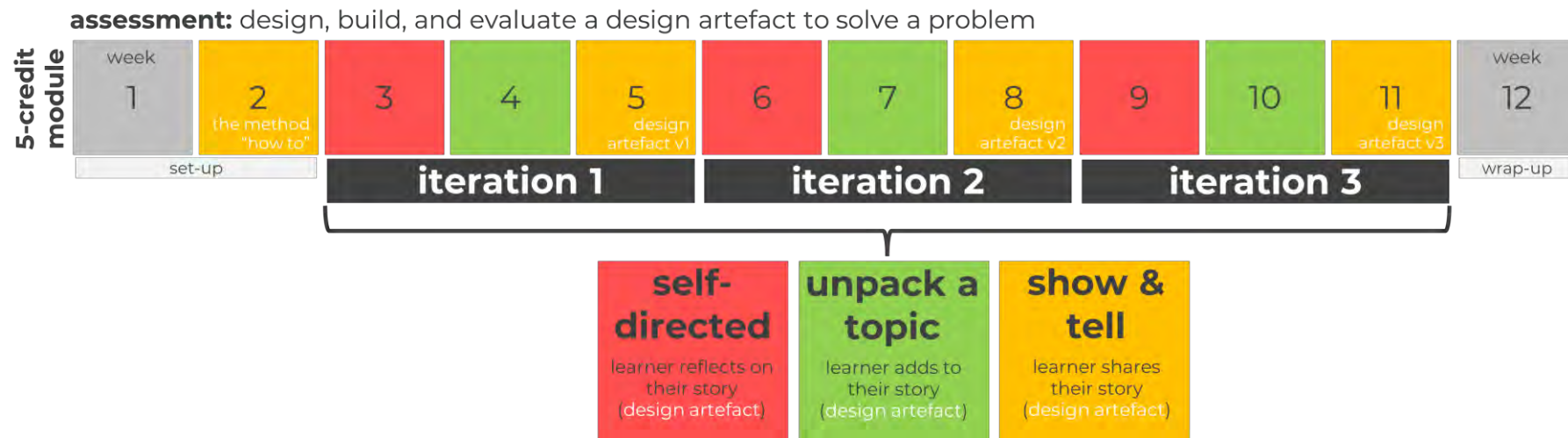
- **The Engagement Assumption:** In session four ("show & tell" #2) one learner presented their evaluation artefact to great applause. The independent evaluator (a seasoned ADR scholar) commented that the learner was displaying all the qualities of an ADR researcher (a sense of *being* and not just *doing*). However, one week prior to this "show & tell" session, the lead author (module lecturer) has a sense that this learner was not engaging with the module material (based on viewing their content engagement analytics on the VLE - Canvas). At that point in time, it crossed the mind of the module lecturer that the learner was disengaged from the module and the "learning-by-doing" pedagogical approach. Notwithstanding this assumption, it all made sense during the "show & tell" session (when the learner presented the evolution of their design artefact across two iterations). In iteration one, version 1 of the design artefact was designed using a small set of prescribed ADR methodology papers. Thereafter, in iteration two, version 2 of the design artefact emerged, where another set of prescribed ADR empirical papers was used to evaluate version 1, and the learnings from which led to the emergence of version 2 (of the design artefact). Therefore, the learner's story during the "show & tell" session presented the *'facts of the case'*, and what seemed like a lack of learner engagement was in fact a considered iterative approach to the design, build, and evaluation of the learner's emerging ADR design artefact.

Therefore, in this *engagement assumption* vignette the use of the SHLAM<sup>2</sup> approach highlights the value of "show & tell" sessions (as opportunities to 'see' and 'celebrate' the learner's progress and as challenges to module lecturer assumptions around learner engagement – originally formed by using content engagement analytics on Canvas). Furthermore, this hermeneutics-inspired vignette also highlights the uniqueness of learning styles and the link between the hermeneutic concepts of *engagement* and *appropriation*. Therefore, not everything we design into our curriculum (as module lecturers) is consumed according to our plan. However, ensuring that we can 'see' and 'celebrate' such learner engagement and appropriation patterns is the critical strategic value.

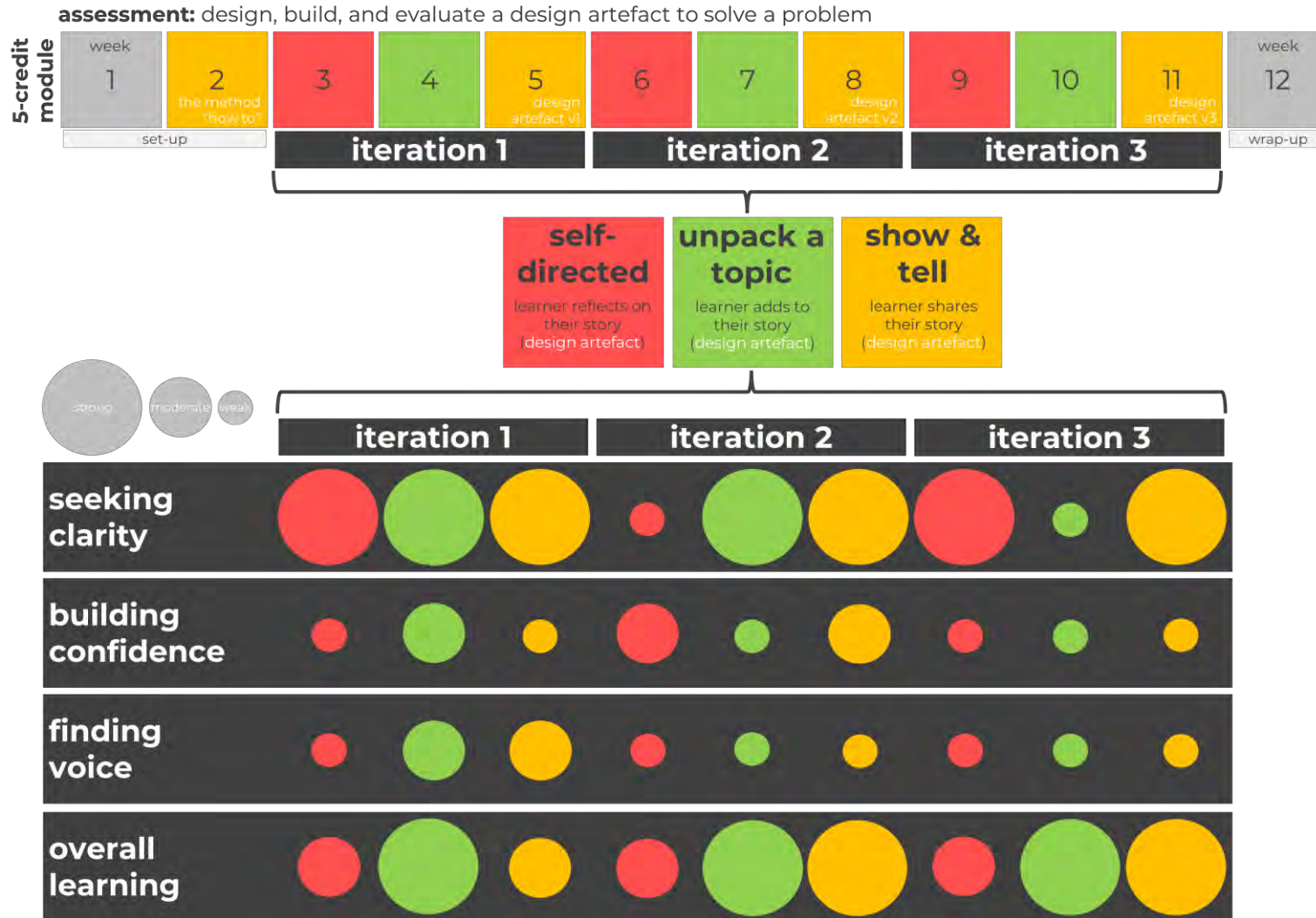
Overall, our observations suggest that learner confidence is growing over the duration of the module, specifically in the four “show & tell” sessions, and especially as the learners’ respective design artefacts evolve. This makes sense, given the fact that when the learner starts session one of the module, they are assumed to be a neophyte researcher (e.g. in Action Design Research). Furthermore, their design artefact does not yet exist. However, when the learner completes the module, they are assumed to be confident in their ability to execute a research study and their design artefact has matured (e.g. a checklist to evaluate how well the execution of ADR is documented in IS literature). As visualised in Figure 4, this evolution is made possible through the actual existence of a design artefact (tangible output) and evaluation of its efficacy in use, along with presentations of each learner’s emerging story during the “show & tell” sessions. Therefore, the module leaders believe that the learner is not only learning about the topic but is also actively evaluating and reflecting on their understanding through the lens of their design artefact. As a result, the design artefact serves as a tangible representation of their evolving knowledge and provides a platform for self-assessment and improvement.

Throughout the 12-week module ‘Seeking Clarity’ is the strongest pattern. However, based on our observations, the time dedicated to reading and re-reading papers is always a challenge for the learners. Furthermore, ‘Building Confidence’ and ‘Finding Voice’ are the weakest patterns of learning, with the former being marginally stronger, given the existence and evolution of the design artefact throughout the iterations. Therefore, aligned with evidence-based nudges, the module leaders often encouraged the learners to keep progressing each week and to be more confident about the shape of their respective design artefacts. Interestingly, the module leaders could see this confidence during the “show & tell” sessions, where presenting during in-class sessions and overcoming the challenge of sharing their design artefact was contributing positively to the learning experience.





**Figure 4. The hermeneutics inspired 'learning-by-doing' pedagogical approach.**



**Figure 5. SHLAM<sup>2</sup> insights showcasing the week-on-week comparative analysis (across the three patterns) and overall learning.**

## 5.0 Key Learnings

The *hermeneutics inspired learning-by-doing pedagogical approach* is an effective way to expose neophyte researchers to a new methodology (ADR in this case). Furthermore, hermeneutics is a very useful lens for analysing learning in the context of a research methodology (ADR in this case). For example, hermeneutics examines learner *engagement* with new texts, as well as re-engagement with previously used texts. This highlights *the hermeneutic circle* in action, where the learner goes back to check/re-evaluate their initial understanding, moving between texts and sections of texts, to appreciate differing contexts, in an effort to generate new and/or improve existing interpretations.

Such engagement contributes to increased *appropriation* and decreases an overreliance on *prejudice*. Over time this engagement also leads to the development of learner *historicity* and their ability to converse with their peers, about the things they read, improves. Furthermore, hermeneutics also highlights the openness of the learner to feedback, along with the importance of *historicity* (ability to converse with peers) and its impact on learner *prejudice*. In fact, seeing as the learners work together on the ADR module (over the 12-weeks) they develop a sense of community and growing awareness of their ‘new’ language being spoken and practices being appropriated. It is fair to say, they now have history and can talk to each other.

Therefore, the relationship between learner *appropriation* and *the hermeneutic circle* leads to improved understanding over time. This is often visible in the reshaping of the design artefact, following a reshaping of the learner’s interpretation. Furthermore, hermeneutics establishes if the learner is open to embracing *distanciation* and creating a sense of *autonomization* by releasing their interpretations (their design artefact) into the wild. We view both the *autonomization* and *distanciation* concepts as great indicators of the increasing confidence of a learner as they move through the 12-week module.

When we design curriculum, we should do so in such a way that we are ensuring (as much as is possible) that our design (learning flow) will enable a learner “*to seek clarity, to find voice, to build confidence*” (see Figure 2). This ‘design principle for

curriculum’ does in fact capture the essence of the theorising process that we observed from our use of the SHLAM<sup>2</sup> approach. Furthermore, following our theorising, as: (i) the designers of the ADR module, and (ii) the researchers of this topic, we conclude that this ‘design principle for curriculum’ also reflects the progress on the learner’s “*interim struggle*” toward “*the qualities of generality, accuracy, and simplicity*” in theoretical explanation (c.f. Weick, 1995, pp.389-390). In the context of this work, *simplicity* translates as ‘simple interpretations of the world’, *generality* translates as ‘the more that can be explained the better’, and *accuracy* translates as ‘closeness to a true or accepted value’. Therefore, some key learnings that have an impact on how we design curriculum are as follows:

- There can be a ‘loss of learning momentum’ during self-directed weeks.
- There is a continuous struggle for learners in reading/re-reading papers (the time and effort). So less prescribed reading is more valuable.
- Learners can often fail to take every and any opportunity to evaluate their design artefact. Therefore, “show & tell” sessions are invaluable to promote the sharing of the design artefact (and interpretations) for feedback.
- The value of a ‘learning-by-doing’ approach and being feedback-centric on weekly efforts (accelerating weekly improvements) also carries a cost in the form of the lecturer effort involved in using the SHLAM<sup>2</sup> insights to shape the module sessions prospectively.

## 6.0 Conclusions

Are we analytically impaired in our role as Higher Education Institution (HEI) educators? In many ways our curriculum designs and pedagogical approaches can challenge our abilities to celebrate the efforts (*interim struggles*) of learners as they learn. Therefore, to be analytical champions, we need to improve our module level storytelling capabilities. To conclude, what are the implications of this work (the SHLAM<sup>2</sup> approach)? In essence, SHLAM<sup>2</sup> is an innovative hermeneutical data-driven approach to assessing learning that provides a unique dataset currently unavailable across many, if not all, HEIs. Furthermore, the insights generated will enhance:

- Student success by transforming the quality of student learning.
- Storytelling for accreditation purposes.
- Evidence-based decision-making through providing meaningful insights as feedback.

In fact, these SHLAM<sup>2</sup> insights (*evidence-based nudges* and *week-on-week comparative analysis*) can become the *de facto* standard for all module lecturers. Therefore, we believe we have the possibility to impact (positively and prospectively) on the *Assurance of Learning* and *Student Engagement* concepts within the HEI context. For example, for *Assurance of Learning*, SHLAM<sup>2</sup> provides greater granular

data to enhance the AOL story and showcases learning prospectively at the module level. Furthermore, for *Student Engagement*, SHLAM<sup>2</sup> grows a community of action-oriented students (as enquirers) and lecturers as reflective practitioners. As we continue with this exploratory work, the following impacts are expected from the SHLAM<sup>2</sup> approach:

- The impacts on the learner are:
  - An appreciation of the importance of certain activities/behaviours during the flow of learning (based on receiving *evidence-based nudges*).
  - An ability to see their progress (e.g. week-on-week) over the duration of a module, independent of their performance in assessments only.
  - An awareness to perform a self-reflection on their performance and associate various activities/behaviours with that performance.
- The impacts on the lecturer are:
  - An ability to provide *evidence-based nudges* to learners in an effort to raise their awareness of the importance of certain activities/behaviours during their learning.
  - An opportunity to pick-up on near to real-time positives/negatives in the module feedback and alter their approach (where necessary) in order to prospectively enhance the student experience while learning.
  - An appreciation of the need to continue to develop as an educational professional (in the scholarship of teaching and learning) and design/deliver impactful content that engages learners.

It is reported that *importance* is the most critical dimension of relevance for IS practitioners, and similar to (Rosemann and Vessey, 2008 p.3) we view *importance* as research that “*meets the needs of practice by addressing a real-world problem in a timely manner [currently significant], and in such a way that it can act as the starting point for providing an eventual solution*”. Therefore, the work presented in this paper is an effort at addressing current shortfalls in curriculum design. For example, can our research inform HEI policy around learner-based feedback throughout the delivery of a module? Is there a gap between the policy discourse and the actual support needed by lecturers to provide meaningful feedback, and learners to receive meaningful feedback on learning (on a weekly basis throughout the delivery of a module)? It is hoped that the practicality of the work presented in this paper will help IS educators to avoid the *hidden traps* (c.f. Hammond, et al., 1998) in their decision making (e.g. *status quo trap, sunk-cost trap, overconfidence trap, etc.*) while promoting a “*focal awareness versus a subsidiary awareness*” with regard to designing learning experiences (i) aligned with a *hermeneutics inspired learning-by-doing pedagogical approach*, and (ii) embracing the continuous use of the SHLAM<sup>2</sup> approach.

To conclude, we appreciate that hermeneutics “helps a researcher to interpret the text such that it makes sense [and] helps the researcher produce a story that is believable” (Myers, 2009, pp.183-184). However, we are conscious that by putting the lived experiences (extracted from the SHLAM<sup>2</sup> weekly dataset) of the three doctoral students (learners) into narrative form “the resulting stories do not duplicate the experience.... the experience is filtered... events in a story are resorted and given order, typically one in which a sequence is created” (Weick, 1995, p.128). Notwithstanding this, our motivation for such an approach is practice-inspired, as we know that “practitioners can relate to stories”, and this is an effective way of “making our research more relevant to practice” (Myers, 2009, p.218). Finally, we conclude our work by encouraging the use of the SHLAM<sup>2</sup> approach which is designed to encourage module lecturers to be reflective (prospectively) on how learning is unfolding from a learner’s perspective. SHLAM<sup>2</sup> advocates for ‘non-assessment based’ feedback which is a departure from the norm and requires a mindset shift on both the side of the lecturer and the learner. The SHLAM<sup>2</sup> approach is built by academics for academics in order to make learning visible through feedback insights!

## Acknowledgement

*Acknowledgement withheld for review purposes.*

## References

- Bruce, B.C. and Bloch, N. (2012). Learning by Doing. In: Seel, N.M. (eds) Encyclopedia of the Sciences of Learning. Springer, Boston, MA. [https://doi.org/10.1007/978-1-4419-1428-6\\_544](https://doi.org/10.1007/978-1-4419-1428-6_544)
- Guzmán-Valenzuela, C., Gómez-González, C., Rojas-Murphy Tagle, A., and Lorca-Vyhmeister, A. (2021) “Learning analytics in higher education: a preponderance of analytics but very little learning?” International Journal of Educational Technology in Higher Education 18, pp.1-19.
- Hammond, J. S., Keeney, R. L. and Raiffa, H. (1998) “The hidden traps in decision making”, Harvard Business Review, September-October, <https://hbr.org/1998/09/the-hidden-traps-in-decision-making-2>
- Kaveri, A., Silvola, A., and Muukkonen, H. (2023). Supporting Student Agency with a Student-Facing Learning Analytics Dashboard: Perceptions of an Interdisciplinary Development Team. *Journal of Learning Analytics*, 10(2), pp.85-99. <https://doi.org/10.18608/jla.2023.7729>
- Leitner, P., Ebner, M., Ebner, M. (2019). Learning Analytics Challenges to Overcome in Higher Education Institutions. In: Ifenthaler, D., Mah, DK., Yau, J.YK.

(eds) Utilizing Learning Analytics to Support Study Success. Springer, Cham.  
[https://doi.org/10.1007/978-3-319-64792-0\\_6](https://doi.org/10.1007/978-3-319-64792-0_6)

- Leitner, P., Khalil, M. and Ebner, M. (2017) “Learning analytics in higher education—a literature review”. Learning analytics: Fundamentals, applications, and trends: A view of the current state of the art to enhance E-learning, pp.1-23.
- Myers, M. D. (2004) Hermeneutics in Information Systems Research, In Social Theory and Philosophy for Information Systems. (eds., Mingers, J. and Willcocks, L.) Wiley, UK, pp.103-128.
- Myers, M. D. (2009) Qualitative research in business & management, Sage Publications, London.
- Nguyen, A., Tuunanen, T., Gardner, L., and Sheridan, D. (2021) “Design principles for learning analytics information systems in higher education”, European Journal of Information Systems, 30(5), pp.541-568, DOI: 10.1080/0960085X.2020.1816144
- Rosemann, M. and Vessey, I. (2008) “Toward improving the relevance of information systems research to practice: the role of applicability checks”, MIS Quarterly, pp.1-22. <https://doi.org/10.2307/25148826>
- Sein, M., Henfridsson, O., Purao, S., Rossi, M. and Lindgren, R. (2011) “Action Design Research”, MIS Quarterly, 35(1), pp.37-56. <https://doi.org/10.2307/23043488>
- Weick, K. E. (1995) Sensemaking in Organizations, Sage Publications, California.

## Appendix A

### SHLAM<sup>2</sup> Learning Flow Questions

#### Reading Papers

1. Did you read a 'new' paper that you have not read previously?
2. Did you find something new in the paper?
3. Did you discuss the paper with a peer (outside of the timetabled sessions)?
4. Did you make any changes to your design artefact (because you read a 'new' paper)?
5. How much time are you investing in reading 'new' papers?
6. Rate your reading of papers at this stage.

#### Re-Reading Papers

7. Did you go back to re-read a paper that you have read previously?
8. Did you find something new in the paper?
9. Did you discuss the paper with a peer (outside of the timetabled sessions)?
10. Did you make any changes to your design artefact (because you re-read a paper)?
11. How much time are you investing in re-reading papers?
12. Rate your re-reading of papers at this stage.

#### Design Artefact Sharing

13. Did you share your design artefact with a peer (outside of the timetabled sessions)?
14. Did you take on board a comment made by a peer (specifically on your interpretation, captured in your design artefact)?
15. Did you make any changes to your design artefact (because you shared the artefact)?
16. How much time are you investing in sharing your artefact?
17. Rate your artefact sharing at this stage.

#### Overall Assessment of Learning

18. How important is your background (prior knowledge/experience) in shaping your design artefact at this stage?
19. How important is reading/re-reading papers in shaping your design artefact at this stage?
20. How important is sharing your design artefact (for feedback) in shaping your design artefact at this stage?
21. Please provide a short update on your learning at this stage in the module (a **positive** and a **negative** aspect).
22. Rate your learning **overall** on the module to date.



# A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach (TEAeM)<sup>1</sup>

Laurence Brooks<sup>1</sup>, Sara Canizzarro<sup>2</sup>, Nitika Bhalla<sup>2</sup>, Kathleen Richardson<sup>2</sup>

<sup>1</sup> Information School, University of Sheffield, Sheffield, UK. <sup>2</sup> Centre for Computing and Social Responsibility (CCSR), De Montfort University, Leicester, UK.

## Abstract

*This paper presents an approach which aims to build on earlier approaches to the ethical study of emerging technologies. The ethics of emerging technologies is a growing field, as the range and potential impact of technologies is ever expanding in our increasingly technological world. The ethics of emerging technologies and their study is first reviewed by examining a number of key ethical frameworks (including ATE, eIA, Future Studies and ATE+). However, these approaches lack some elements, which the paper looks to strengthen by integrating policy and empirical elements. The resulting approach is the TechEthos Anticipatory ethics Matrix (TEAeM) framework. This framework provides a methodological foundation for the study of ethical issues, especially for emerging technologies. In this way the TEAeM framework contributes to the field of emerging technologies ethics analysis.*

**Keywords:** ethics, emerging technologies, ethics issues, frameworks, method, TEAeM

## 1.0 Introduction

Technologies in the 21<sup>st</sup> century are emerging all the time. Most of this aims to enhance the world and society that we live in and hopes to benefit people in their everyday lives, in some way. However, there is also the possibility that these technologies can, intentionally or otherwise, have a negative effect on society. If we drill down into this, it might be that it benefits one part of society while it disadvantages another part, and so we need to be aware of the wider implications of each of these emerging technologies. For example, facial recognition enables a plethora of technologies to function more easily, such as unlocking a smartphone or laptop, but might have biases built in causing it to have trouble recognising darker-skinned females ([Buolamwini & Gebru, 2018](#)). Other examples of potential negative social impacts of emerging technologies include; an increase in social inequality, exacerbation of existing power asymmetries, creation of bias and job displacement.

---

<sup>1</sup> Acknowledgement: The TechEthos project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No.101006249.

The central problem for the ethics of emerging technologies is that we humans cannot predict the future, and therefore do not know which ethical issues will play out once the technology is fully developed and entrenched in society. As the emerging technology is still evolving, many questions can arise about its nature, its future use, and its social consequences. However, if an ethical framework is to be useful in an area of emerging technology.

This paper, therefore, argues for the development of a method to analyse the ethical issues involved that emerge as the technologies themselves are emerging, thus anticipating potential inequalities and concerns, rather than just reacting to them. Further, to suggest the possibility of building a greater ethical sensitivity into the people playing a role in shaping the development of the technologies themselves, or at least avoid the worst of the potential fallout.

TechEthos ([www.techethos.eu](http://www.techethos.eu)) is an EU-funded project that deals with the ethics of the new and emerging technologies anticipated to have high socio-economic impact. The project involves ten scientific partners and six science engagement organisations and ran from January 2021 to December of 2023. TechEthos aimed to facilitate “ethics by design”, namely, to bring ethical and societal values into the design and development of new and emerging technologies from the beginning of the process. Technologies covered are “climate engineering”, “digital extended reality” and “neuro-technologies”. The project aims to produce operational ethics guidelines for these technologies for users such as researchers, research ethics committees and policy makers. To reconcile the needs of research and innovation and the concerns of society, the project explores the awareness, acceptance and aspirations of academia, industry and the general public alike and reflects them in the guidelines.

In presenting a method to approach the ethics of emerging technologies, this paper distils some guidance for a method for analysing ethical issues in emerging technologies, from existing frameworks such as Anticipatory Technology Ethics (ATE), Ethical Impact Assessment (eIA), Future ethics and the nascent Anticipatory Technology Ethics plus (ATE+).

## 2.0 Existing Ethical Frameworks

Starting with the review of existing ethical frameworks and following the example set by Brey ([2012a](#)) this paper first defines emerging technologies, then examines key existing ethical approaches.

### 2.0.1 Existing Approaches to Ethics of Emerging Technologies?

Technologies are developing and expanding all the time, and as they say, it is an ‘ever expanding field’. However, there does need to be some way to define what we mean by the term, “emerging technologies”, at least to create a shared understanding of the boundaries around them ([Haessler et al., 2022](#)). One of the most quoted approaches to this discusses five key attributes that appear to help identify a technology as emerging, as derived from a review of relevant literature by ([Rotolo et al., 2015](#)):

- a) radical novelty,
- b) relatively fast growth,
- c) coherence (persisting over time),
- d) prominent impact (on the socio-economic domain), and
- e) uncertainty and ambiguity (as we don’t really know what the future holds and therefore what the impact of a technology will bring).

In the TechEthos horizon scan task (carried out as the first part of the project), we use the term new and emerging technologies to identify any type of technology that performs a new function or improves some function significantly better than other commonly used technology, which is expected to be developed and deployed in the next 5 to 10 years (adapted from the Organization for Economic Co-operation and Development, ([OECD, 2017](#))).

Emerging technologies are also generally seen as having higher risk (in a general sense of risk arising from the uncertainties associated with the unknown), from both the uncertainties around their impact and also the potential issues with timely ‘fixing’ of any unwanted consequences which might arise from these impacts ([Munoko et al., 2020](#)). This paper, along with others ([Stahl et al., 2017](#); [Wright, 2011](#)), looks to find a way to identify and analyse the ethical issues that arise from these risky emerging

technologies (in the sense of identifying what the potential ethical risks might be early on).

## **2.1 Key ethical frameworks**

The paper first briefly reviews three approaches to ethical analysis that exist in the literature and have previously been applied in a range of technology contexts. While this is not a comprehensive account, it aims to identify the key criteria in each approach. The three approaches selected for review are: Anticipatory Technology Ethics (ATE), Ethical Technology Assessment (eTA) and Future Studies.

### **2.1.1 Anticipatory Technology Ethics (ATE)**

This approach focuses on emerging technologies from the perspective of trying to identify what is both good and bad about them. However, as these technologies are being developed, it is one thing to say what ethical issues are known, or can be reliably expected, but then there are also the ethical issues that will emerge over time as a consequence of use. Brey ([2012a](#)) reviews four approaches to technology assessment focused on ethics, namely ethical Technology Assessment (eTA) ([Palm & Hansson, 2006](#)), ethical Impact Assessment (eIA) ([Wright, 2011](#)), techno-ethical scenarios ([Boenink et al., 2010; Stahl, 2011](#)), ETICA approach ([Stahl, 2011](#)). additional insights on the other ethical frameworks can be found in TechEthos deliverable D5.1 (Bhalla *et al.*, 2023). Based on his analysis of these four approaches, Brey proposes a fifth approach, ATE, which he says has “the potential to meet all the criteria that a sound approach to ethical analysis of emerging technologies should have” ([Brey, 2012a](#)).

ATE has three levels of ethical analysis: technology, artifact and application level (Figure 1). It then defines what it calls ‘objects of ethical analysis’ for each of these levels, as properties or processes that might lead to ethical issues.

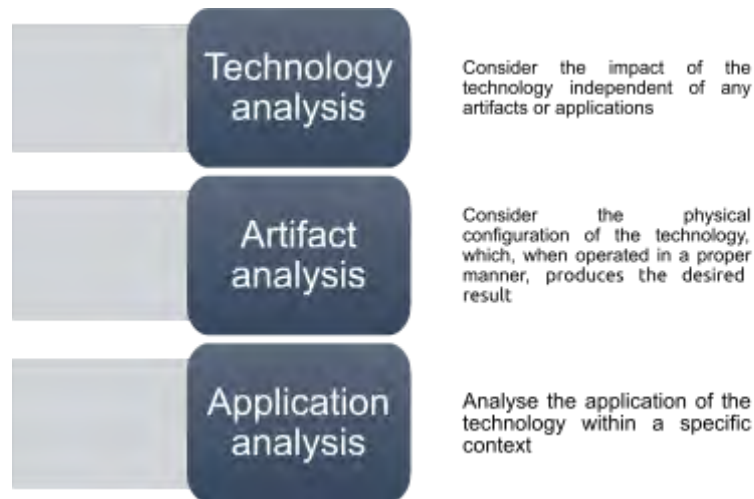


Figure 1: Original ATE levels ([Brey, 2012a](#))

One of the issues for the early stages of ATE is how to identify the appropriate ethical values to be mapped with the specific technology. Brey ([2012b](#)) proposes an ethics checklist (see Table 1), which encompasses a range of ethical values and principles, based on ones that have been seen in earlier ethical approaches and commonly found within society (and also acknowledges that variations in culture and/or specific types of technology, might need a more specific list). The four categories of ethical principles are: Harms and risks, Rights, Justice (distributive) and Well-being and the common good.

**Table 1: The anticipatory technology ethics checklist (Brey, 2012b)**

<ul style="list-style-type: none"> <li>● Harms and risks <ul style="list-style-type: none"> <li>○ Health and bodily harm</li> <li>○ Pain and suffering</li> <li>○ Psychological harm</li> <li>○ Harm to human capabilities</li> <li>○ Environmental harm</li> <li>○ Harms to society</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>● Rights <ul style="list-style-type: none"> <li>○ Freedom</li> <li>○ Freedom of movement</li> <li>○ Freedom of speech and expression</li> <li>○ Freedom of assembly</li> <li>○ Autonomy</li> <li>○ Ability to think one’s own thoughts and form one’s own opinions</li> <li>○ Ability to make one’s own choices</li> <li>○ Responsibility and accountability</li> <li>○ Informed consent</li> <li>○ Human dignity</li> <li>○ Privacy</li> <li>○ Information privacy</li> <li>○ Bodily privacy</li> <li>○ Relational privacy</li> <li>○ Property</li> <li>○ Right to property</li> <li>○ Intellectual property rights</li> <li>○ Other basic human rights as specified in human rights declarations (e.g., to life, to have a fair trial, to vote, to receive an education, to pursue happiness, to seek asylum, to engage in peaceful protest, to practice one’s religion, to work for anyone, to have a family, etc.)</li> <li>○ Animal rights and animal welfare</li> </ul> </li> </ul>
<ul style="list-style-type: none"> <li>● Justice (distributive) <ul style="list-style-type: none"> <li>○ Just distribution of primary goods, capabilities, risks and hazards</li> <li>○ Non-discrimination and equal treatment relative to age, gender, sexual orientation, social class, race, ethnicity, religion, disability, etc.</li> <li>○ North–South justice</li> <li>○ Intergenerational justice</li> <li>○ Social inclusion</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>● Well-being and the common good <ul style="list-style-type: none"> <li>○ Supportive of happiness, health, knowledge, wisdom, virtue, friendship, trust, achievement, desire-fulfilment, and transcendent meaning</li> <li>○ Supportive of vital social institutions and structures</li> <li>○ Supportive of democracy and democratic institutions</li> <li>○ Supportive of culture and cultural diversity</li> </ul> </li> </ul>

Munoko et al., (2020) summarise the 5 steps for the researcher to follow in ATE as:

- “First, at the *technology level*, the researcher considers the features of the technology of ethical concern, independent of its current or potential use. This level involves the identification of the inherent and consequential risks of the technology.
- Secondly, at the *artifact level*, the researcher considers the “physical configuration that, when operated in the proper manner and the proper environment, produces the desired result.” At this level, the researcher focuses on the artifacts independent of their actual applications and identifies the risks associated with the intended use of the artifacts.
- Third, at the *application level*, the actual use of an emerging technology’s artifact is studied. At this level, the researcher considers the unintended

consequences for the users of the applications and other stakeholders (for example, ChatGPT).

- Fourth, the researcher evaluates the potential importance of the issues identified.
- Finally, the fifth part of the ATE framework is optional, where the researcher can design a feedback stage.
- There are additional optional stages beyond the fifth step. One optional stage is the responsibility assignment stage, where “moral responsibilities are assigned to relevant actors for ethical outcomes at the artifact and application levels.” Another optional stage is the governance stage, which provides policy recommendations.”

Munoko et. al., ([2020](#)) then combine ATE with the ETICA approach ([Stahl, 2011](#)), as they feel that each of the methods, while closely linked, contributes something that the other does not. More recently, ATE has been cited as one example of ‘technology oriented assessment methods’, including eTA, eIA, as well as value-sensitive design (VSD), privacy for design, socially responsible design (SRD), eco-design, ethics by design ([Gurzawska, 2021](#)).

One critique of ATE is that trying to predict what might be the impact and outcomes of emerging technologies, will be problematic, as until people take up and use those technologies it is difficult to recognise what might be the unintended and emergent properties. However, it is still possible that likely outcomes can be conceptualised and recognised, within a framework such as ATE.

### **2.1.2 Ethical Impact Assessment (eIA)**

The framework identifies key social values and ethical issues, provides some brief explanatory contextual information which is then followed by a set of questions aimed at the technology developer or policymaker. The aim of this framework is to facilitate consideration of ethical issues, in consultation with stakeholders, which may arise in their undertaking. In addition to consultation with stakeholders, the framework includes a set of ethical tools and procedural practices which can be employed as part of the ethical impact assessment. The ethical tools help the technology developer to get a better idea of how the technology is perceived ethically by stakeholders. Furthermore, the framework provides a diagrammatic pathway which is useful to

technologists to check and review potential ethical challenges and to mitigate some of the risks. The EIA framework consists of the following steps:

- 1) conducting an EIA threshold analysis,
- 2) preparing an EIA plan,
- 3) identifying ethical impacts
- 4) evaluating the ethical impacts (step 3 and 4 are to be carried out in consultation with stakeholders),
- 5) formulating and implementing remedial actions,
- 6) reviewing and auditing the EIA.

The EIA framework does not account for emerging technologies in the future but investigates continuously the ethical implications of what is known about the technology under development. However, as there are often inherent privacy issues such as equality and human dignity etc. in new and emerging technology, research has also been carried out to integrate privacy impact assessment into EIA (Wright and Friedewald, 2013).

### **2.1.3 Future Studies**

Prediction, foreseeing of the future is a key feature of all human cultures and was traditionally expressed by oracles, and augurs who could gift the future in the present. In modern scientific societies, prediction moved from the professions of clairvoyants, fortune tellers and prophets to professionals, academics who would develop techniques and methodologies for 'seeing the future'. So, when thinking about the ethical implications of these emerging technologies, one is in a sense trying to predict the future impact of these technologies and their potential consequences.

Future Studies emerges as an interdisciplinary field, recognising that the 'future' is not produced by one agent, but a number of intersecting, often colliding and reacting processes, which is often also seen as technologies emerge. A critical problem for it is the role of *time* – not understood as linear and singular but, with the future seen as an outcome of gestures and properly studied as 'interval crossers' and 'interval openers' ([Schneider, 2019](#)). Future Studies also accounts for the role of imagination, and 'the



imaginary as resources for (re-) shaping our world and imagining new relations' and prioritising the role that stories play in constructing human existence, ([Spengler, 2019](#)).

Future Studies goes beyond prediction, as it aims to shape the future according to principles and values that are important to humans. But what is the future – is it anytime that is beyond the present, or a place that is always shaped by fictional imaginaries and any prediction must consequently be partly, a work of fiction.

Sociologist Zygmunt Bauman noted that the future is not always a desired goal, and he coined the term “*retrotopia*” as an umbrella term ([Bauman, 2017](#)) for those movements and trends that seek to get back to something, rather than moving somewhere else (cited in Paul ([2019](#))). Hence ideas of the future are intrinsically connected to the past and present, imagined and factual, as opportunities, and destruction are feasible outcomes of any process.

Future Studies is not without its critics, for to have a future must imply a desired or imagined state of existence, calling into question who decides this future? Who is left out or excluded from future imaginings? The question is whether technology innovation is the solution to the problems developed in tech-capitalist societies? Technology, as the engine of capitalist innovation, opening up the possibilities of creating new products, processes and practices, underlying a belief in unfettered creativity and flexibility of the human species to adapt to any technologically inspired living arrangement.

Höjer and Mattsson ([2000](#)) identified four critical problems with a Future Studies approaches: 1) identifying ‘cyclic behaviour in socio-technical changes’; 2) viewing one technology to be crucially reliant on the development of another (in their case it was transport and communication that entangled and connected), 3) interrogating basic assumptions about a field (in their case it was the ‘hypothesis of constant travel time’ as a stable), and 4) human and resource relationships (613). The future is a ‘fiction’ of sorts, shaped by practices, ideas and, extrapolated into some undefined future point – problematically producing a determinism – if this, then that – view. Moreover, they suggest that ‘backcasting’ as an alternative and better predictor than

‘forecasting’ in cases where future scenarios are seen as detrimental, and harmful. Sardar prefers the term ‘alternative futures’ due to the possibility of plurality, identity crises and meaning ([Sardar, 2010](#)).

Ethically speaking, the ‘future’, if it exists at all, is a contested domain, heterogenous, and diverse (and contested), while ethics also can be seen to propose a set of standards to be recognised and incorporated into technological practices and artefacts. Artificial Intelligence (AI) is a case in point, with a past littered with inaccurate accounts – and yet evidence of failed predictions are passed over, as new ones form and develop ([Sundvall, 2019](#)).

What all these approaches to ethical analysis of emerging technologies show is that it is difficult to predict the future. However, as techniques and approaches they each demonstrate that it is possible to develop some guidance on how to assess the possible ethical issues associated with a specific technology, so that developers and users may reflect on this and potentially incorporate those reflections into their design, development and use.

#### **2.1.4 Anticipatory Technology Ethics plus (ATE+)**

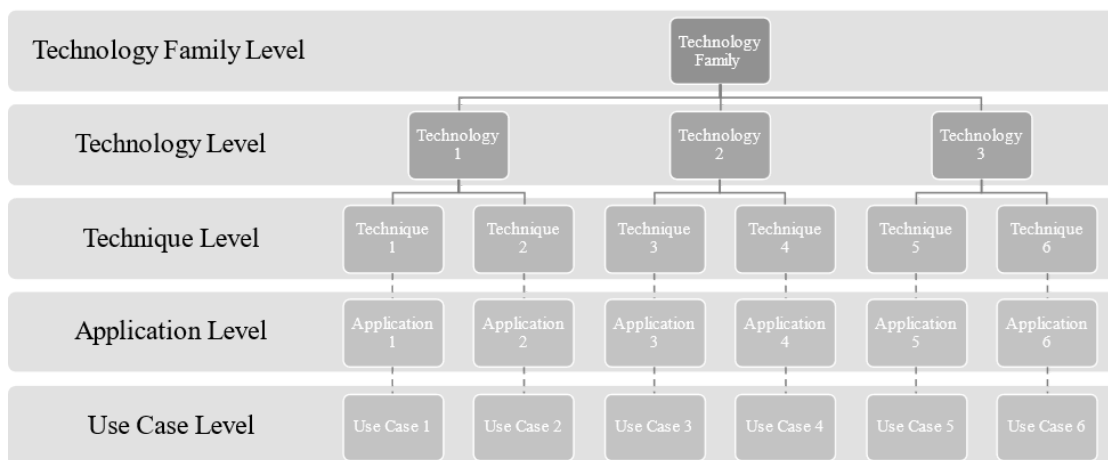
While ATE shows a lot of value in evaluating the ethics of emerging technologies, it does have some limitations when used in practice. While the TechEthos project adopted the ATE approach as the starting point for its ethical analysis ([Buchinger et al., 2022](#)), given the limitations, an expanded version of ATE, named ATE+, has been developed ([Umbrello et al., 2023](#)). This augments the analysis taxonomy, creating a more detailed framework that is less abstract and so aims to be more useful in applied settings, in particular complementing ethics-by-design approaches.

ATE+ begins by identifying 4 gaps in ATE that need to be addressed. These are: a) the identification stage begins with culturally and personally situated values of ‘what is good’ which is rather removed and abstract compared with looking at practice. b) The question of ‘whose values’, in terms of the values adopted as the starting point. c) Moving from the identification stage to the evaluation stage is complex/subjective and needs engineering and user expertise, plus contextual understanding. d) The main

focus of an ATE analysis is on what does not work/negative impacts, which means the potential to be gained from positive impacts can be missed. These shortcomings are observed across, “1) levels and objects of ethical analysis; 2) approach to foresight methods; 3) methods of ethical analysis” (Umbrello et. al., nd).

#### 2.1.4.1 Enhanced Levels and Objects of Ethical Analysis

Having identified the range of gaps in the original ATE formulation and noting the strengths of ATE, the authors of ATE+ propose some “modifications to the levels and objects of ethical analysis, the methods of foresight included; and the methods of the ethical analysis themselves” (Umbrello et al., 2023). First there is a revised version of levels and objects of ethical analysis (see Figure 2), to enable a more fine-grained analysis. This brings in a broader and more general picture, starting with a ‘technology family’ level (“collections of technologies sharing common goals, or formal or functional features”), then technology and finally a collection of techniques, which the specific technology could employ. The specific technique might be used in specific applications in the context of a specific use case.



**Figure 2: TechEthos level and objects of ethical analysis. Dotted lines from the Technique Level down to the Application Level and Use Case Level signify that not all technologies will have application or use cases as a condition of their readiness level**

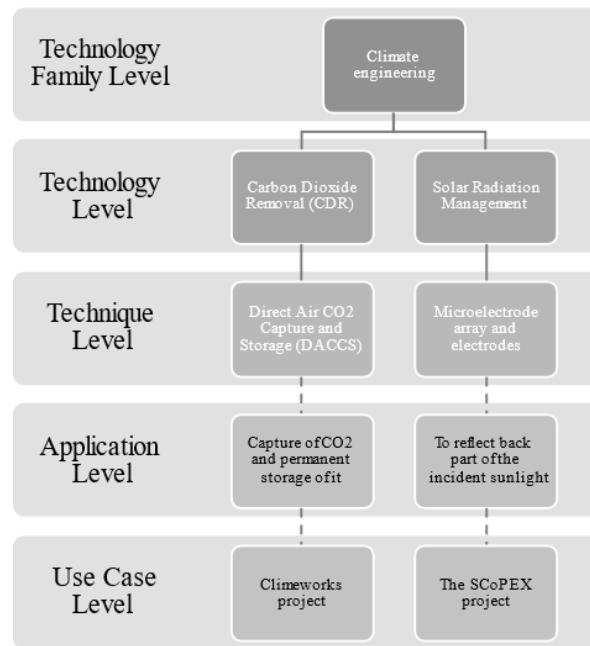
At the highest level, focus on families of technologies: collections of technologies sharing common goals, or formal or functional features. For example, a technology family bound by the term “climate engineering” might include technologies with the common goal of advancing carbon dioxide removal. At the middle, second and third levels, they propose focusing on specific technologies, which may combine various

techniques or domains sharing formal or functional features and goals. At the lowest two levels, focus is on techniques in specific applications and use cases. At these levels material artefacts, products and actual procedures come in.

Example applications of the ATE+ levels of ethical analysis applied to climate engineering is shown in Figure 3 ([Umbrello et al., 2023](#)). Further, to better identify the potential positive outcomes of a technology, and better engage with a wider range of stakeholders, a landscape of ethical concerns, relevant to both intended users and others was identified (see Table 2).

**Table 2: Landscape of ethical concerns in the ATE+ framework**

Ethical concerns related to...	Levels of analysis		
	High	Middle	Low
Main goals or features...	<i>...of technology families</i>	<i>... of technologies</i>	<i>...of technique (or application or use case, as appropriate)</i>
Desirable or undesirable unintended side-effects for intended users...	<i>...of technology families</i>	<i>... of technologies</i>	<i>...of technique (or application or use case, as appropriate)</i>
Potential contribution to enabling future morally controversial developments if...	<i>...technology families are purposed to different goals</i>	<i>...technologies are purposed to different goals</i>	<i>...techniques (or application or use case, as appropriate) are purposed to different goals or with different procedures</i>
Unintended side-effects for non-users (desirable or undesirable), when considering uncertainties and risk perceptions...	<i>...of technology families</i>	<i>...of technologies</i>	<i>...of techniques (procedures, actions, or goals) in application or use case, as appropriate)</i>



**Figure 3: ATE+ levels of ethical analysis applied to the climate engineering technology family**

To augment the original landscape of ethical concerns covered by ATE, reference to “likelihood” was removed and instead focused on “desirability.” Doing so allowed first, to give the levels and object of analyses a balance to the potentially implicitly negative/undesirable term “consequences.” In addition, the terms “desirable” and “undesirable” provided a more substantial warrant for engaging the diverse public, stakeholder, and expert groups and drawing explicit analytical attention to tensions related to potentially conflicting values of different parties. Such concerns can be surfaced for intended users and non-users alike, allowing for casting an even broader analytical net.

One of the issues noted for ATE+, is that in the reconceptualization of the levels of ethical analysis, direct reference to the concept of ‘artifact’ has been lost. However, one might argue that it is subsumed within the new levels, between Technique and Application.

#### 2.1.4.2 Enhanced Foresight Methods

Within the context of the TechEthos project, there was an element of future perspective, in which a variety of future social and ethical issues were surfaced by

creating contrasting future scenarios. Whereas ATE had focused on ‘likely futures’, ATE+ replaced this with a focus on questions of ‘plausible futures’, where plausible aims to stimulate reflection on social, ethical, environmental, economic and other impacts (and where plausible refers to something ‘able to be believed’, rather than likely which is about the expectation that something will happen<sup>2</sup>). This plausibility focus is combined with the narrative method (or narratological approach) which can then accommodate some lack of transparency in ethical argumentation (Grinbaum, 2020). This in turn reflects how the public perception of a technology (current or emerging) is shaped by the social debates around them (cf. the shaping of AI by the current public debate and media discussion around ChatGPT).

#### **2.1.4.3 Enhanced Methods of Ethical Analysis**

In response to the original ATE ethical analysis and based on results elicited within the TechEthos project (Adomaitis, Grinbaum and Lenzi, 2022), ATE+ proposes the following cross-cutting steps as their enhancement to the ATE ethical analysis method (which itself had only two stages, a. identify ethical issues, and b. evaluate ethics issues (Brey, 2012b)):

- 1) Describe objects of interest, procedures, techniques, approaches, applications, use cases of interest, etc. (e.g., natural language processing in health);
- 2) Investigate core philosophical notions and dilemmas that serve as conceptual scaffolding for the ethical issues (e.g., Is there an inherent preference for material reality over virtual reality?);
- 3) Identify values and principles (e.g., transparency, dignity) and return to step 4 for clarification if necessary;
- 4) Use narrative analysis to demarcate both transparent ethical considerations and morally opaque presuppositions in technological judgement concerning the values and principles identified in step 5 (e.g., “Be careful what you wish for”);
- 5) Ethnographically engage with critical stakeholders associated with technologies based on narratives instead of an addition to open-ended questions.
- 6) Formulate a set of operationalised design questions to be asked regarding the implementation of techniques (or applications and use cases) (e.g., does the

---

<sup>2</sup> <https://dictionary.cambridge.org/dictionary/english/>

XR system take stock of the potential changes of behaviour in its users? Who profits from the changes in behaviour and how are the changes incited?).

Given the importance of these issues, the TechEthos project proposes to further develop its approach to ethical analysis, using the five levels of ethical analysis (from ATE+) as guidance.

## **2.2 Integrating academic ethical frameworks with policy and primary data**

The approach proposed in this paper integrated the theoretical ethical frameworks with two types of ‘hands-on’ information: 1) policy documents, and 2) empirical data concerning ethical issues of the technologies, as drawn from industry and academic experts.

### **2.2.1 Integrating ethics with policy - scan of existing ethical codes, frameworks and guidelines**

#### **2.2.1.1 The search strategy**

Starting with integrating ethics with policy documents, the aim was to collect and map policy documents that would allow capturing of the policy landscape surrounding each technology family. Wright’s (2011) methodological considerations were taken as a starting point which explicitly included the work of policymakers as a means to shape technology’s ethical impact assessment. This approach includes policy documents in the form of a scan of existing ethical codes, frameworks and guidelines cited or proposed in the academic literature as well as in the grey literature. The methodology for this literature review began with a mixture of search contexts. First it identified published reports, academic journal articles, books, and working papers that examined guidelines, ethical codes, codes of conduct, and governance frameworks as used within the three TechEthos technology families (climate engineering, digital extended reality and neurotechnologies) on standard academic databases as JStore, ACM Digital Library, IEEE Xplore Digital Library, AIS eLibrary as well as Google scholar.

The key terms used for the literature review are:

- ‘ethical codes’
- ‘ethical frameworks’
- ‘ethical guidelines’

These terms were searched for in the abstract, as the aim was for the ethical guidelines’ to be the article’s key focus. To ensure relevance, the search included references to the technology family (e.g. digital Extended Reality) and specific instances of technological artefacts (e.g. NLP) in the body of the text.

Once the ethical frames were retrieved, the next step was to review the results further based on a number of fundamental ethical principles (namely the impact on Human rights, Freedom, Autonomy, Integrity, Responsibility, Privacy and Security (Brey, 2012b) to gain an idea of the extent to which these principles were covered in the documents gathered. For example, concerning the principle of integrity for Climate Engineering documents, we noted that “Integrity was an ethical issue mentioned within a number of research documents from both academic and research organisations. Hubert and Reichwein (2015) argue that there are limits to scientific freedom.” (Cannizzaro et al 2021). The resulting set of ethical principles for each technology family varied considerably and different results were returned.

The research also carried out a general Google search identifying reports from companies, or organisations that are traditionally excluded from academic databases (‘grey literature’).

The aim was to obtain a set of at least 20 documents per technology family, comprising both published academic literature and grey literature from industry, government, non-academic and non-governmental (NGO) research and policy organisations that would have ethical guidelines, codes and frameworks relating to the technology families as a key content in their text. The search did not seek to include texts which mentioned ethical principles in general without reference to specific guidelines, codes and texts.



Through this search, the study wanted the specific technology type and the keywords ‘guideline/code/framework’ to be present in the abstract or at least the keywords, to ensure it would obtain documents that were specifically *about* ethical guidelines for the technology families and their specific technologies.

### 2.2.1.2 Mapping of extracted codes, guidelines and frameworks

The next step was to map the characteristics of the extracted codes, guidelines and frameworks to make sure there was a sufficiently diverse variety of policy documents - particularly to ensure that a mix of academic as well as grey literature articles had been captured (Table 3). The strategy used in this mapping exercise follows Rothenberger, Fabian and Arunov (2019) who reviewed ethical guidelines for Artificial Intelligence. They included the type of organisation issuing the guideline and a definition for each.

**Table 3: Example of sample based on selected sources within the neurotechnologies technology family which illustrate instances of codes, frameworks, and guidelines, as well as the diversity of type of organisation covered, including academic sources and non academic sources.**

<b>Neurotechnologies: ethical guidelines, codes, frameworks and issues (based on Rothenberger, Fabian and Arunov (2019))</b>			
<b>Guideline</b>	<b>Type of organisation</b>	<b>Definition</b>	<b>Extract of source guideline</b>
Ethical code	Academia	Ethical codes set forth responsibilities to which individuals and groups or organisations hold themselves to account.	...professional self-regulation [...] should start within a company, institution or other work unit with a code of ethics or set of clearly articulated principles to which leadership adheres... (Chang et al 2019)
Ethical frameworks	Academia, other research organisation	Ethical frameworks set forth general or specific principles to which countries, organizations, or research communities hold themselves to account.	Australia currently lacks a clear regulatory framework for ensuring that individuals are informed about how their data are captured, stored, analyzed, and shared (Australian Brain Alliance 2019)  The degree of perturbation of advanced neurotechnology on the current ethical legal framework is quantitatively higher than non-computational techniques (Ienca and Andorno 2017).

Neurotechnologies: ethical guidelines, codes, frameworks and issues (based on Rothenberger, Fabian and Arunov <a href="#">(2019)</a> )			
Guideline	Type of organisation	Definition	Extract of source guideline
Ethical guidelines	Academia, other research organisation	Ethical guidelines collect general or specific principles specifying how a technology or field ought to develop	<p>The clinical research organizing team has also created guidance for any necessary modifications needed in ethical protocols due to revisions of research guidelines (Sadato et al 2019).</p> <p>To develop national guidelines for responsible neuroinnovation to assist neuroscientists, engineers, and developers to translate research into effective and ethical products. (ABA 2019).</p>

### 2.3 Integrating ethical frameworks with primary data - digital ethnographies and expert interviews

Collecting and mapping the policy documents allows for capturing of the policy landscape surrounding each technology family. Following this, the ethical frameworks could be integrated with primary data. The rationale for this primary data gathering was to “stay in contact with technology developers during the whole developmental process”, as prescribed in Ethical Technology Assessment (eTA) ([Palm & Hansson, 2006](#)). To do so involved extracting empirical data consisting of ethical speculations on future ethical issues that the technology families might bring. This was done through a methodology consisting of digital ethnographies and expert interviews. In addition Brey ([2012a](#)) argues in favour of future studies of technology. This is given that “the possibility for a viable future depends on the imagination and on the imaginary as resources for (re-)shaping our world and imagining new relations ([Spengler, 2019](#)). Therefore, both data-collection techniques - digital ethnographies and expert interviews - and analysis procedures were framed by a broad future ethics consideration.

Thus, in order to preliminary map out the socio-economic impact of the emerging technology families of climate engineering, extended reality and neurotechnologies, we carried out a preliminary, systematic literature review of ethical issues thought to affect these technologies. To do so, we aimed to find 10 academically influential papers on ethics of the said technology family and identify the key ethical issue outlined.

The ethical principles guiding the collection and analysis of primary data are supported by TechEthos' key objective to promote 'ethics by design' i.e. to bring ethical and societal values into the design and development of technology from the very beginning of the process.

### **2.3.1 From ethnography to digital ethnography**

A classic definition of traditional ethnography is provided by Ingold who defines it as an approach with the objective to “describe the lives of people other than ourselves, with an accuracy and sensitivity honed by detailed observation and prolonged first-hand experience” ([Pink & Morgan, 2013](#)). Hence *information*, *emotions*, *observership* and *subjectivity* appear to be key traits of this research methodology. Another key trait is the focus on *context* by means of reference to the concept of thick “description”, borrowed from anthropology ([Geertz, 1973](#)). Context here refers to the web of meanings, which constitute a culture and within which objects as cultural signs are situated. Prasad ([1997](#)) argues that it is the ethnographer's task to uncover and present these multiple meanings and their complex connections with each other in the course of analysing any social event. He reminds readers that meanings are sometime shared but other times contradictory and contested. Greenhalgh and Swinglehurst ([2011](#)) refer to three more concepts characterising ethnography, which they term key interpretive criteria i.e authenticity, plausibility, criticality. Authenticity is gained through immersion of the ethnographer within the culture, plausibility amounts to developing explanations, which make sense to participants and are arranged in a coherent narrative, and criticality refers to questioning assumptions.

Time is a central concept within ethnography. This approach to research is usually intense and long, for example it would require a one year of fieldwork immersion. However, in short-term ethnography the “immersion” of the ethnographer is for only a

short period. These ethnographies are characterised by research activities being undertaken in a shorter time frame ([Pink & Morgan, 2013](#)). This approach has also been dubbed “quick and dirty” as it recognises the impossibility of gathering a complete and detailed understanding of the setting at hand” ([Pink & Morgan, 2013](#)). Alongside a compressed notion of time, place and space are a key feature of short-term ethnography. Ethnographic places are not simply fieldwork localities, but rather entanglements through which ethnographic knowing emerges. This is significant for the purpose of this project because in times of Covid-19 pandemics, lockdown and remote working conditions, places, including ethnographic, have become virtual, hence the emergence of short term *digital* ethnographies. This type of ethnography considers how humans live in a digital sensory environment. Horst et al ([2015](#)) define digital ethnography as a way to research practices that are reported or demonstrated, for example through participants’ own digital media biographies and capturing the language that is used when speaking about their area of concern. Referring to Algorithmic ethnography, during and after COVID-19, Christin ([2020](#)) defines digital ethnography as a collection of methods that entail identifying, gathering, and analysing digital data.

### **2.3.2 Ethnography of technologies**

Ethnography tends to become shaped by the discipline it is being engaged through, and the research evolves in dialogue with theory rather than being led or structured a priori by it ([Pink & Morgan, 2013](#)). Technology and particularly, technological innovation (such as that which comes from emerging technologies), can be investigated “in-the-making” through ethnography ([Petschick, 2015](#)). Thus, the ethnography of technology Prasad ([1997](#)) explains how the anthropological tradition within which ethnography is situated treats technologies as a cultural artifact accomplishing specific social functions as well as both reflecting and structuring social practices. In other words, in the ethnographic approach, technologies are seen as more than merely functional instruments fit for specific purposes, but they are seen as cultural and symbolic object/artifact e.g., they may be ceremonial, embedding the myths of the culture in which they are situated or they may exert social control ([Prasad, 1997](#)). Ethnography can uncover the symbolic function of a technology within the context of the culture in which it is embedded, Because of this, this proposed method seeks to examine this symbolism under the perspective of ethics,

that is, by looking at what some of the elements of the ethnographic objects may mean in relation to ethics and specifically ethical issues brought about by the technology families in the future.

Table 4 shows a sample of ethnographic objects analysed, comprising the material for analysis. A search for businesses' proposing applications within the technology families has been made from the business platform LinkedIn. This was reputed to be a better source than Google for search thanks to its filters which helped to gauge the relevance of the results rather efficiently as it contains filters such as companies, people, region, industry and company size. A mixture of web pages and YouTube videos were selected to use as ethnographic objects of analysis. YouTube videos were selected when they included talks at a conference or interviews by media agents rather than solely promotional videos which are more staged and may have hindered the detection of any spontaneity of emotions triggered when talking about the future.

**Table 4: List of references to the digital ethnographic objects (video or company website) systematically- selected for the digital ethnographies. The names of the companies have been withheld to ensure anonymity.**

<b>Company Reference Number</b>	<b>Ethnographic Object Type</b>	<b>Technological application</b>	<b>Country in which the company is based</b>
1	Website page	Electroencephalography (EEG) and Brain-Computer Interface	Lithuania
2	Website Page	Wearable medical Device for monitoring Parkinson's disease	Greece-UK
3	YouTube Video	Neuromodulation through prismatic lenses	Italy
4	YouTube Video	Brain-Computer Interface (implant)	US
5	Website page	Carbon Dioxide Removal and Utilisation	Sweden
6	Website page	Carbon Dioxide Removal	US
7	YouTube Video	Carbon Dioxide Removal and Geological storage	Switzerland
8	YouTube Video	Carbon Dioxide Removal	US

<b>Company Reference Number</b>	<b>Ethnographic Object Type</b>	<b>Technological application</b>	<b>Country in which the company is based</b>
9	Website page	XR – holographic display	Denmark
10	Website page	XR - extended reality experience	UK
11	YouTube Video	VR social platform	US
12	YouTube Video	VR and AR	Portugal

### **2.3.3 Expert interviews**

The ethical frameworks were further integrated with empirical data coming from expert interviews, specifically, through semi-structured expert interviews (Table 5).

During the expert interviews, ethical dilemmas, questions informed by epistemological analysis, future studies, as well as the ‘guiding questions’ method suggested by Stahl, Timmermans and Flick (2017) have been used in order to open ethical reflection on new and emerging issues. In addition to this, the interviews have followed a similar structure to that of the literature review where questions around future ethical issues and impacts have been explored, as well as the ethical principles and values that arise when analysing each technology family.

The TechEthos project focuses on the ethical issues associated with the three technology families, therefore the criterion for interviewee selection was technical and ethical expertise associated with Climate engineering, Extended Digital Reality and Neurotechnology. During the first round of expert interviews, eight interviews took place online, using MS Teams.

The contact details of the interviewees was identified through collaboration with the TechEthos project partners. The interviewees were contacted via a template email and, after agreeing to an interview, each interviewee was sent a TechEthos information leaflet and a consent form to complete, sign and return as their acceptance to participate in the interview.

**Table 5: List of experts interviewed during the first round of interviews in the project.**

<b>Technology area</b>	<b>Expertise</b>	<b>Country</b>
Neurotechnology	Neurology, Ethics	France
Extended Digital Reality	Academic, Ethics	US
Extended Digital Reality	Academic, Ethics	UK
Climate Engineering	Academic	Netherlands
Neurotechnology	Academic, Neuroethics	Sweden
Climate Engineering	Academic	Netherlands
Extended Digital Reality	Policy maker, ethics	US
Climate Engineering	Academic	Norway

Follow-up email contact was made with all potential interviewees who had not responded by return of the completed and signed consent form, within seven days of the original email being sent.

The semi-structured but flexible interviews were approximately 30 minutes duration with anticipated scope for extension beyond, given interviewee active/engaged participation and willingness to continue. Accordingly, the interview protocol consisted of a minimum of eight essential, open questions (see Appendix 1). Much like the digital ethnographies, these questions were also framed by a broad future studies perspective. The interviews were audio and video recorded via MS Teams, and the insights have been captured as a summary of each question.

#### **2.3.4 Expert consultation and interviews**

The consultations with experts was conducted through qualitative interviews and workshops that were set up to receive feedback on the following questions:

- Clarity: Is the meaning of the value in the context of this technology family clear and comprehensible?
- Completeness: Is the main argument in the subsection complete? What should be added?

- Operationalization: Are the questions at the end of the subsection helpful operationally? Is anything missing in that aspect?
- What else do you find interesting and worth mentioning about this technology family?

The consultations took part as a form of an online workshop in June 2022, which began with a discussion of the overall structure, followed by three plenary sessions dedicated to each of the three TechEthos technology families. The workshop was attended by ten European technology ethics experts, drawn from Academia and Industry (see Table 6).

**Table 6: List of experts consulted during the consultation.**

<b>Ref No.</b>	<b>Role (academic, policymaker, industry)</b>	<b>Country</b>
1	Interdisciplinary research Centre	Switzerland
2	Networks	Lithuania
3	Academia	USA
4	Academia (former science funding org.)	Ireland
5	Standards	Italy
6	Academia	UK
7	Industry	Switzerland
8	Academia	Spain
9	Research Centre	Germany
10	Ex Academia	Russia

## **2.4 Discussion**

With awareness of existing policy documents, as well as empirical data concerning ethical issues coming from digital ethnographies and expert interviews, developers, policymakers and scholars, can act on those specific areas in the present, foregrounding an ethics by design approach in order to contribute to the safest possible outcome that is the responsible future development and deployment of the given technology families.



#### **2.4.1 Towards a ‘TechEthos Anticipatory ethics Matrix’ (TEAeM)**

Having reviewed a number of key ethical frameworks and the stages developed as part of the ethics analysis in the TechEthos project, these can now be combined to create a more developed ‘TechEthos anticipatory ethics matrix’ or TEAeM (Figure 6). This integrates the stages of the proposed ATE+ approach ([Umbrello et al., 2023](#)), which has also emerged from the research done within the TechEthos project, together with a number of elements derived from the analysis in this paper drawn from other established ethical frameworks, notably the link to policy via policy documents (drawn from ATE, Brey (2012b)) and ethical impact assessment eIA (drawn from previous research, ([Wright, 2011](#))), the link to empirical data (drawn from eTA, ([Palm & Hansson, 2006](#))) which maintains contact with the technology developers and hence the potential for ethical influence on them and the link to future ethics ([Angheloiu & Tennant, 2020](#); [Spengler, 2019](#)), which provides a future orientated approach, which is appropriate for these families of emerging technologies. While this can seem like a lot of elements to combine, it does allow for a very flexible and integrative approach.

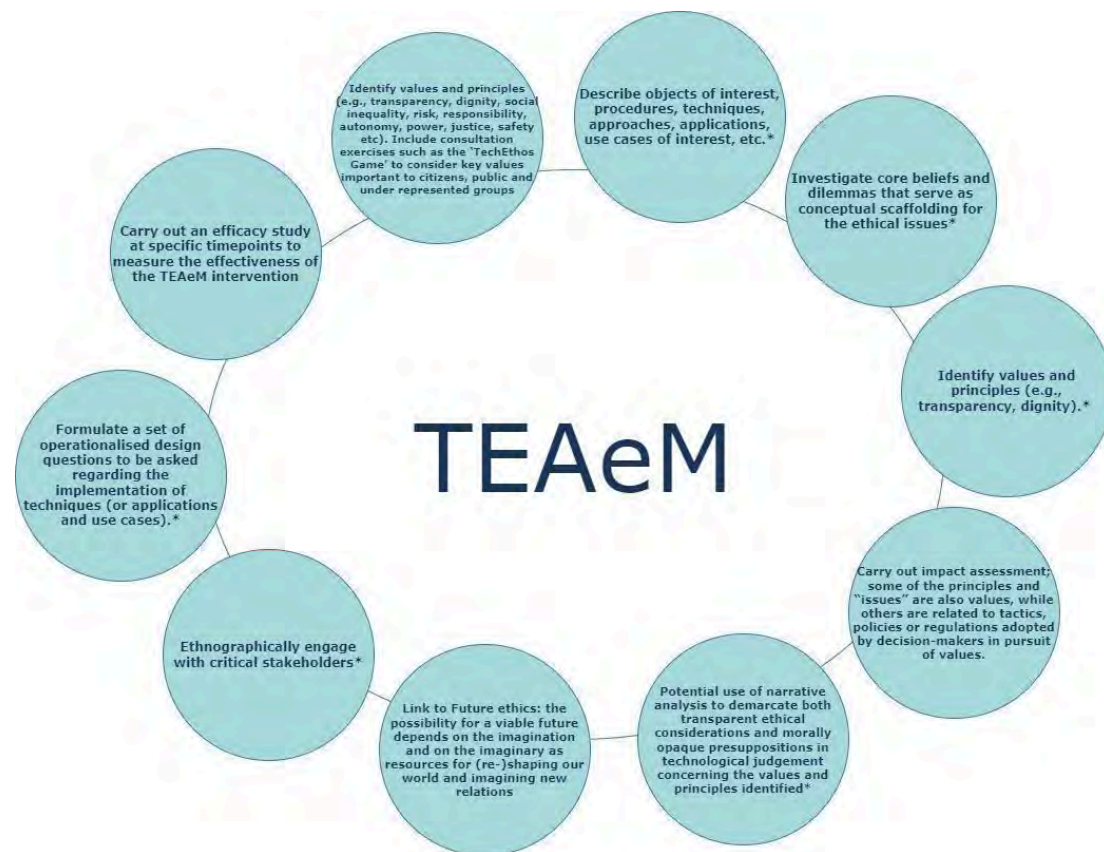


Figure 6: The TechEthos Anticipatory ethics Matrix (TEAeM). Note \* denotes a step detailed in ATE+ (Umbrello et al., 2023)

### 2.4.2 Testing TEAeM on Quantum Cryptography

The qualitative testing of the TEAeM framework, and to demonstrate its broad applicability (Table 7), was done by showing how it relates to a specific emerging technology. The chosen emerging technology is ‘quantum cryptography’, which is seen as a potentially transformational technology that could significantly impact society (Seskir *et al.*, 2023).

**Table 7: Application of TEAeM to Quantum Cryptography**

TEAeM	TEAeM applied to Quantum Cryptography
Describe objects of interest, procedures, techniques, approaches, applications, use cases of interest, etc.	“Based on quantum bits that can be zero and one simultaneously and on instantaneous correlations across the device, a quantum computer acts as a massively parallel device with an exponentially large number of simultaneous computations. There already exist algorithms overcoming the speed and capacity of any classical supercomputer.” Examples include, “communication links whose security is underwritten by unbreakable laws of physics (breakthrough compared to conventional encryption)”. <a href="#">(Porcari et al., 2021)</a>
Investigate core philosophical notions and dilemmas that serve as conceptual scaffolding for the ethical issues	As this is an essentially as yet un-developed technology, so the following are based on the ‘promise’ of what the technology might provide: Errors and misunderstanding Encryption Natural ethics Control on research <a href="#">(Wakunuma &amp; Stahl, 2014)</a>
Identify values and principles (e.g., transparency, dignity)	The development and deployment of quantum technology raise specific ethical issues, such as justice, fairness, equity, inclusion, respect for persons, ESG, human rights, etc. <a href="https://coruzant.com/quantum/impact-of-quantum-on-the-digital-economy-and-society/">(https://coruzant.com/quantum/impact-of-quantum-on-the-digital-economy-and-society/)</a>
Carry out impact assessment. Some of the principles and “issues” are also values, while other issues are related to tactics, policies or regulations adopted by decision-makers in pursuit of values (like data protection). The identification of values and policy design are two different needs <a href="#">(Wright, 2011)</a> . (Anticipatory Technology Ethics (ATE)).	Use of one of a range of impact assessment tools to identify what are the potential impacts of the technology, as it currently stands. Use of academic and grey literature, as well as potentially relevant policy documents, to establish the set of values that have been linked to quantum computing to date (for example, de Wolf <a href="#">(2017)</a> identifies ethical aspects of quantum computing including: cryptography, increased inequality and making the impact positive).
Use narrative analysis to demarcate both transparent ethical considerations and morally opaque presuppositions in technological judgement concerning the values and principles identified	For example, creation of scenarios around various quantum computing near and middle future contexts to help developers, users and others think about the range of issues, both transparent and opaque. For example, whether in the air or on the ground, quantum computers will help to streamline traffic control; they will be able to quickly calculate the optimal routes concurrently which allows for efficient scheduling and would reduce traffic congestion, and while generally positive, what other issues might this incur, such as who controls the traffic flow or aspects such as surveillance.
Link to Future ethics: the possibility for a viable future depends on the imagination and on the imaginary as resources for (re-)shaping our world and imagining new relations.	Use of future oriented analysis in the digital ethnography, to establish what kind of future is being envisioned by the quantum computing technology developers and application experts and organisations.

TEAeM	TEAeM applied to Quantum Cryptography
Ethnographically engage with critical stakeholders associated with technologies based on narratives.	Use LinkedIn to search for companies working in ‘quantum computing’ and then review websites/videos, etc., using digital ethnography approach, eg. <a href="https://www.quantumcomputinginc.com/">https://www.quantumcomputinginc.com/</a>
Link to empirical data: aim to stay in contact with technology developers during the whole developmental process and discuss different approaches to problems that arise. ... continuous dialogue and repeated assessments are preferable to one single large-scale assessment (Ethical Technology Assessment (eTA)).	Engage with Quantum Computing developers and users to engage in ongoing dialogue with them about problems that arise in the development and application processes. Use of databases, such as Cordis, to identify research projects in the appropriate field and contact them to establish a set of experts in Quantum Computing that can also be consulted with, for example a quick search of Cordis projects database ( <a href="https://cordis.europa.eu/projects/en">https://cordis.europa.eu/projects/en</a> ) shows 1,177 results for 'quantum computing'.
Formulate a set of operationalised design questions to be asked regarding the implementation of techniques (or applications and use cases).	Use the results from the various analysis carried out in the previous stages to create the set of quantum computing relevant design questions.

This was a first attempt at applying this version of the TEAeM framework to a specific emerging technology, to evaluate the ease of use and usefulness. Therefore, the decision was to apply this to a specific sub-field of quantum computing, namely ‘quantum cryptography’, which can be defined as “a technique that involves the use of the laws of quantum mechanics to enable the parties involved to exchange random strings of qubits with one another. These qubits may be used as a key to encrypt and decode messages that are being sent between the parties.” (Alhayani et al., 2023). The application of TEAeM was reasonably good, with some research needed to unearth specific information required by different elements. This was also reviewed by an expert in the field and it seems clear that one of the limitations of the approach is a deeper understanding of the conceptual limits and pragmatic limits of an emerging technology. In the case of quantum computing and quantum cryptography, while some elements are conceptually possible, many are still seen as ‘promises’ of what might come. Hence, looking for the ethical issues is even more difficult.

As noted above, due to the theoretical nature of quantum cryptography, there are strong limitations and challenges when applying a framework such as TEAeM to technologies that are more conceptual than realised. However, as technologies progress, then having the early insights into the ethical issue and considerations needed for that technology is a good starting point. We do feel that the further

development of the TEAeM framework could be a positive step towards guiding emerging technologies towards a more ethical road.

## **2.5 Conclusion**

This paper has presented the results of part of the EU funded TechEthos project, which deals with the ethics of new and emerging technologies anticipated to have high socio-economic impact. Combining the method that flows out of a review of existing ethical frameworks used for analysing technology, especially future oriented anticipatory ones, with existing policy documents and empirical data coming from digital ethnographies and expert interviews, developers, policymakers and scholars, has led to the first steps in the development of the TechEthos Anticipatory ethics Method (TEAeM). This was then applied to another high socio-economic impact emerging technology, namely quantum cryptography, which was used as a first test application.

In this way it is hoped to show that the paper has achieved what it set out to do. Namely to develop a first step towards a method that analyses the ethical issues involved that emerge as the technologies themselves are emerging, and which also allows for greater ethical sensitivity through the use of techniques such as scenarios and the development of questions that prompt deeper ethical insights to be surfaced.

## **Appendix 1**

The interviewees were asked the following questions in a semi-open format:

1. Can you tell us about your area of expertise, how many years have you worked in your field of interest?
2. As a result of technological innovation in the area of (technology family) how do think the world will change by 2045?
3. In your view, what do you think are the benefits associated with this technology by 2045?
4. Can you anticipate what risks and harms might arise?
5. Who are the main beneficiaries of this [technology family]? And who will be excluded in your view?
6. Considering the global interest in the issue of ethics, what do you predict to be the ethical issues that could arise by 2045?
7. Do you think we have gone past the point of reversibility & irreversibility of this technology? And please explain why?
8. Is there anything else you would like to add which we have not covered already?

## References

- Alhayani, B. A., AlKawak, O. A., Mahajan, H. B., Ilhan, H., & Qasem, R. M. (2023). Design of Quantum Communication Protocols in Quantum Cryptography. *Wireless Personal Communications*.  
<https://doi.org/10.1007/s11277-023-10587-x>
- Angheloiu, C., & Tennant, M. (2020). Urban futures: Systemic or system changing interventions? A literature review using Meadows' leverage points as analytical framework. *Cities*, *104*, 102808.
- Bauman, Z. (2017). *Retrotopia*. Polity.
- Bhalla, N. et al. (2023) *TechEthos D5.1: Enhancement of Ethical Frameworks and Outline of Detailed Ethics Framework*. Available at: [www.techethos.eu](http://www.techethos.eu).
- Boenink, M., Swierstra, T., & Stemerding, D. (2010). Anticipating the interaction between technology and morality: A techno-ethical scenario study of experimenting with humans in bionanotechnology. *Studies in Ethics, Law, and Technology*, *4*(2), 1–38.
- Boer, B., Hoek, J., & Kudina, O. (2018). Can the technological mediation approach improve technology assessment? A critical view from 'within'. *Journal of Responsible Innovation*, *5*(3), 299–315.
- Brey, P. A. E. (2012a). Anticipating ethical issues in emerging IT. *Ethics and Information Technology*, *14*(4), 305–317.  
<https://doi.org/10.1007/s10676-012-9293-y>
- Brey, P. A. E. (2012b). Anticipatory Ethics for Emerging Technologies. *NanoEthics*, *6*(1), 1–13. <https://doi.org/10.1007/s11569-012-0141-7>
- Buchinger, E., Kienegger, M., Zahradnik, G., Bernstein, M. J., Porcari, A., Gonzalez, G., Pimponi, D., & Buceti, G. (2022). *TechEthos technology portfolio: Assessment and final selection of economically and ethically high impact technologies. Deliverable 1.2 to the European Commission. TechEthos Project Deliverable*. Available at: [www.techethos.eu](http://www.techethos.eu).
- Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, *81*, 1–15.

- Cannizzaro, S., Brooks, L., Richardson, K., Umbrello, S., Bernstein, M., Adomaitis, L., (2021). *Methodology for ethical analysis, scan results of existing ethical codes and guidelines*. TechEthos Project Deliverable. Available at: [www.techethos.eu](http://www.techethos.eu).
- Christin, A. (2020). Algorithmic ethnography, during and after COVID-19. *Communication and the Public*, 5(3–4), 108–111. <https://doi.org/10.1177/2057047320959850>
- Criado-Perez, C. P. (2019). *Invisible Women: Exposing Data Bias in a World Designed by Men*. Random House.
- De Wolf, R. (2017). The potential impact of quantum computers on society. *Ethics and Information Technology*, 19, 271–276.
- Geertz, C. (1973). *Thick description: The interpretation of cultures. The Interpretation of Cultures*. London: Harper Collins.
- Greenhalgh, T., & Swinglehurst, D. (2011). Studying technology use as social practice: The untapped potential of ethnography. *BMC Medicine*, 9(1), 45. <https://doi.org/10.1186/1741-7015-9-45>
- Gurzawska, A. (2021). Responsible Innovation in Business: Perceptions, Evaluation Practices and Lessons Learnt. *Sustainability*, 13(4), 1826. <https://doi.org/10.3390/su13041826>
- Haessler, P., Giones, F., & Brem, A. (2022). The who and how of commercializing emerging technologies: A technology-focused review. *Technovation*, 102637. <https://doi.org/10.1016/j.technovation.2022.102637>
- Höjer, M., & Mattsson, L.-G. (2000). Determinism and backcasting in future studies. *Futures*, 32(7), 613–634. [https://doi.org/10.1016/S0016-3287\(00\)00012-4](https://doi.org/10.1016/S0016-3287(00)00012-4)
- Horst, H., Lewis, T., Hjorth, L., Postill, J., & Pink, S. (2015). Digital ethnography: Principles and practice. *Digital Ethnography*, 1–216.
- Hubert, A.M. and Reichwein, D., 2015. *An Exploration of a Code of Conduct for Responsible Scientific Research Involving Geoengineering: Introduction, Draft Articles and Commentaries*. Draft Articles and Commentaries (May 1, 2015). <https://ssrn.com/abstract=3513900> or <http://dx.doi.org/10.2139/ssrn.3513900>
- Kiran, A. H. (2012). Responsible design. A conceptual look at interdependent design–use dynamics. *Philosophy & Technology*, 25(2), 179–198.



- Kiran, A. H., Oudshoorn, N., & Verbeek, P. P. (2015). Beyond checklists: Toward an ethical-constructive technology assessment. *Journal of Responsible Innovation*, 2(1), 5–19.
- Le Guin, U. (1981). *The Left Hand of Darkness*. Orbit.
- Marak, S. (2019). *Anthropocene* (H. Paul, Ed.). Palgrave-MacMillan.
- Munoko, I., Brown-Liburd, H. L., & Vasarhelyi, M. (2020). The Ethical Implications of Using Artificial Intelligence in Auditing. *Journal of Business Ethics*, 167(2), 209–234. <https://doi.org/10.1007/s10551-019-04407-1>
- OECD. (2017). *The Next Production Revolution: Implications for Governments and Business*. OECD. <https://doi.org/10.1787/9789264271036-en>
- Palm, E., & Hansson, S. O. (2006). The case for ethical technology assessment (eTA). *Technological Forecasting and Social Change*, 73(5), 543–558. <https://doi.org/10.1016/j.techfore.2005.06.002>
- Paul, H. (Ed.). (2019). *Critical Terms in Futures Studies*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-28987-4>
- Petschick, G. (2015). Ethnographic panels for analyzing innovation processes. *Historical Social Research/Historische Sozialforschung*, 210–232.
- Pink, S., & Morgan, J. (2013). Short-Term Ethnography: Intense Routes to Knowing: Short-Term Ethnography. *Symbolic Interaction*, 36(3), 351–361. <https://doi.org/10.1002/symb.66>
- Porcari, A., Pimponi, D., Gonzalez, G., Buceti, G., Buchinger, E., Kienegger, M., Bernstein, M. J., & Zahradnik, G. (2021). *TechEthos DI.1 Description of selected high socio-economic impact technologies*. TechEthos. Available at: [www.techethos.eu](http://www.techethos.eu).
- Potts, J. (2018). Futurism, Futurology, Future Shock, Climate Change: Visions of the Future from 1909 to the Present. *PORTAL Journal of Multidisciplinary International Studies*, 15(1–2), 99–116. <https://doi.org/10.5130/portal.v15i1-2.5810>
- Prasad, P. (1997). Systems of Meaning: Ethnography as a Methodology for the Study of Information Technologies. In A. S. Lee, J. Liebenau, & J. I. DeGross (Eds.), *Information Systems and Qualitative Research* (pp. 101–118). Springer US. [https://doi.org/10.1007/978-0-387-35309-8\\_7](https://doi.org/10.1007/978-0-387-35309-8_7)
- Rothenberger, L., Fabian, B., & Arunov, E. (2019). RELEVANCE OF ETHICAL GUIDELINES FOR ARTIFICIAL INTELLIGENCE – A SURVEY AND

EVALUATION. *Research-in-Progress Papers*.

[https://aisel.aisnet.org/ecis2019\\_rip/26](https://aisel.aisnet.org/ecis2019_rip/26)

- Rotolo, D., Hicks, D., & Martin, B. R. (2015). What is an emerging technology? *Research Policy*, *44*(10), 1827–1843.  
<https://doi.org/10.1016/j.respol.2015.06.006>
- Sardar, Z. (2010). The Namesake: Futures; futures studies; futurology; futuristic; foresight—What’s in a name? *Futures*, *42*(3), 177–184.  
<https://doi.org/10.1016/j.futures.2009.11.001>
- Schneider, R. (2019). Gesture. In H. Paul (Ed.), *Critical Terms in Futures Studies* (pp. 145–149). Springer International Publishing.  
[https://doi.org/10.1007/978-3-030-28987-4\\_23](https://doi.org/10.1007/978-3-030-28987-4_23)
- Spengler, B. (2019). Imagination. In H. Paul (Ed.), *Critical Terms in Futures Studies* (pp. 163–169). Springer International Publishing.  
[https://doi.org/10.1007/978-3-030-28987-4\\_26](https://doi.org/10.1007/978-3-030-28987-4_26)
- Stahl, B. (2011). IT for a better future: How to integrate ethics, politics and innovation. *Journal of Information, Communication & Ethics in Society*, *9*(3), 140–156.
- Stahl, B., Timmermans, J., & Flick, C. (2017). Ethics of Emerging Information and Communication Technologies On the implementation of responsible research and innovation. *Science and Public Policy*, *44*(3), 369–381.  
<https://doi.org/10.1093/scipol/scw069>
- Sundvall, S. (2019). *Artificial Intelligence* (H. Paul, Ed.). Palgrave-MacMillan.
- Umbrello, S., Bernstein, M. J., Vermaas, P. E., Resseguier, A., Gonzalez, G., Porcari, A., Grinbaum, A., & Adomaitis, L. (2023). From speculation to reality: Enhancing anticipatory ethics for emerging technologies (ATE) in practice. *Technology in Society*, *74*, 102325.  
<https://doi.org/10.1016/j.techsoc.2023.102325>
- Wakunuma, K. J., & Stahl, B. C. (2014). Tomorrow’s ethics and today’s response: An investigation into the ways information systems professionals perceive and address emerging ethical issues. *Information Systems Frontiers*, *16*(3), 383–397. <https://doi.org/10.1007/s10796-014-9490-9>
- Wright, D. (2011). A framework for the ethical impact assessment of information technology. *Ethics and Information Technology*, *13*(3), 199–226.

Wright, D. and Friedewald, M. (2013) 'Integrating privacy and ethical impact assessments', *Science and Public Policy*, 40(6), pp. 755–766. Available at: <https://doi.org/10.1093/scipol/sct083>.

# Generative AI as a Tool for Thematic Analysis: An Exploratory Study with ChatGPT

**Sian Joel-Edgar & Yu-Chun Pan**

Northeastern University London, London, UK

sian.joel-edgar@nulondon.ac.uk; yu-chun.pan@nulondon.ac.uk

## Abstract

*Artificial intelligence (AI) has seen rapid development in recent years and it has increasingly applied to various fields. Research is no exception. However, there is much to be explored in this domain. This study aims to explore the suitability of current generative AI applications for research purposes. The focus is on the generative AI's capability to synthesise information as a potential alternative or supplement to human-based information synthesis. In order to evaluate the effectiveness of the thematic analysis produced by generative AI, this study compares the generative AI-produced results by ChatGPT with human-generated results, based on the same set of papers. The results show generative AI produced very similar results to humans, in terms of the topics themselves and the number of topics identified. However, there are also some minor mismatches between generative AI and human results.*

**Keywords:** AI, Artificial Intelligence, ChatGPT, ChatGPT4, Research Methods, Thematic Analysis

## 1.0 Introduction

Artificial Intelligence (AI) has made its way into different aspects of our lives and attracted attention from several domains. AI can be considered a system's capability to correctly interpret and learn from external data and to use the learning to achieve specific goals and tasks through flexible adaptation (Makarius et al, 2020). The ability of AI has come a long way since models based on decision trees, random forests and k-means clusterings. Generative AI, a field within artificial intelligence (AI), is responsible for generating fresh and potentially unique content (van Dis et al., 2023). Its application can be viewed as both a creative and rational tool, depending on its usage and the surrounding circumstances. With the capacity of natural language processing powered by supervised and unsupervised training, generative AI applications demonstrated a wide range of use from creative writing to business document generation (Metz, 2023). In November 2022, OpenAI introduced ChatGPT, which swiftly garnered acclaim for its innovative approach to generating AI-based content (Dwivedi et al, 2023). ChatGPT, as one of the most commonly used

generative AI applications, provides unique text in response to user queries by harnessing a huge collection of textual data. The outputs closely mimic human-generated content. There has been widespread usage of ChatGPT in a variety of fields, such as software development, poetry, essays, corporate communication, and legal agreements (Zhuo et al, 2023).

AI has significantly impacted societies and individuals. Organisations are implementing AI in their business process at a fast pace (De Cremer & Kaparov, 2021). It has been applied to marketing (Manis, et al., 2023), social media marketing (Liu et al., 2023), engineering design (Yüksel et al., 2023), healthcare (Mahdi et al., 2023), mental health (Thieme et al., 2023), banking (Rahman et al., 2023), human resource management (Chowdhury et al., 2023) and so on. When it comes to research, several research papers have listed ChatGPT as authors (Stokel-Walker, 2023) which caused debates on whether generative AI applications can be considered as credited authors. There is no universally agreed guidance on such stances, and some of the most prestigious academic journals, such as Nature and Science, have made it clear that they do not accept generative AI applications as authors. Whether generative AI applications should be listed as credited authors or not, researchers are increasingly adopting generative AI for research purposes. Therefore, there is a need to advance our understanding of such implementations.

This study aims to explore the suitability of generative AI as a tool for research purposes, thematic analysis in particular. The following literature review will provide an overview of the AI applications and generative AI for research purposes. An experiment that evaluates the effectiveness of a thematic analysis conducted by a generative AI application, namely ChatPGT, will be introduced. The findings will highlight the iterative steps this study takes and the restless, followed discussion and conclusion.

## **2.0 Relevant Work**

With the rapid advancement of AI algorithms, the content generated by AI, such as social media feeds, can be indistinguishable from human-generated content (Rossi et al., 2023). The availability of several AI applications, such as ChatGPT, Bard, Microsoft Copilot, and DALL-E, has sparked considerable interest and adoption of AI. People have been applying generative AI to a wide range of contexts. Academic

research is no exception. Within academia, AI has also drawn attention from researchers and educators (Dwivedi et al, 2023). Some of the key debates have centred around academic assessment integrity (Stokel-Walker, 2022; Eke, 2023). There are growing concerns around how ChatGPT and other generative AI applications could be used by students to produce assessments and consequently undermine academic integrity. While there is no one universally agreed solution and there is no validate tool to identify inappropriate use of generative AI , there is general consensus that a broader approach to integrating generative AI into pedagogy and assessment is required. It is important to note that while generative AI, such as ChatGPT, has been relatively new tool for research, other AI applications, asu cas Grammarly, has been utilised by a large number of users, indulging researchers, to improve professional and academic writing.

Due to the ability to process a large quantity of information and particularly processing natural language, generative AI has been experimented as a tool to research. Many researchers have utilised generative AI as part of the research methods in information systems research. Rossi et al. (2024) summarised that the current use of generative AI broadly fell into two categories, namely realistic experiments with generated content and using synthetic data. Many experiments require the use of text and images, and collecting naturally-occurred materials can be time-consuming and difficult, due to the need to be precise in controlling and measuring variables. Generative AI can help achieve realistic text and images to allow researchers to create materials efficiently. The other main use of generative AI for research is the creation of synthetic data. When it comes to the collection and storage of data related to human subjects, there are concerns about sensitive and identifiable information. Examples include synthetic profile pictures (Boyd at al., 2023) and synthetic medical images (Chambon et al., 2022). By using AI-generated synthetic data that closely resemble real-world data, researchers can mitigate privacy risks as well as address ethical and regulatory concerns.

Furthermore, generative AI has also been utilised to refine research questions and to check the completeness of the results (Burger et al., 2023). Going beyond text processing, AI has also been used as a tool for medical research (Castiglioni et al., 2021) such as recognising skin lesions with results matching or suppressing the accuracy of a dermatologist (Du-Harpur et al., 2020). Literature review is another area where generative AI has also been tested as a tool (Aydın and Karaarslan, 2022;

Pan et al., 2023). Literature review is an essential part of research. A higher-level synthesis is necessary for a literature review, and it must incorporate ideas from other fields to provide a comprehensive overview of a given subject (Watson and Webster, 2020). It has been suggested that such tasks can be assisted by AI by generating an ontological map of concepts (Li et al, 2020). For instance, ChatGPT has been applied to automate the process of systematic literature review in the field of water and wastewater management (Alshami et al., 2023). ChatGPT has also been applied to generate research ideas in finance research, although it was considered that the literature synthesis and proposed testing frameworks could be further refined (Dowling & Lucey, 2023).

While generative AI could be a useful tool, it was also shown that when using ChatGPT as a tool to identify the literature for review, it could produce inaccurate and even non-existing results that could not be found in other databases (Haman & Školník, 2023). Additionally, although threats to academic integrity have always existed such as plagiarism, there are limited means currently for publishers to effectively identify the authenticity of authorship and the inappropriate use of generative AI as they would for plagiarism. Unquestionably generative AI applications will become more and more advanced and more readily available with major technology companies such as Microsoft and Google investing heavily in this area. The use of generative AI will consequently become more common in the coming months and years. Generally, it is considered that the use of generative AI as a tool in information systems research is still in a very early stage and more clear guidelines should be carefully developed (Rossi et al., 2024).

### **3.0 Approach**

This study aims to explore the suitability of current generation AI applications commercially available for research purposes. The focus is on the generative AI's capability to synthesise information as a potential alternative or supplement to human-based information synthesis. Thematic analysis for systematic literature review (Crossan and Apaydin, 2010; Tranfield et al., 2003) requires a large amount of information to be synthesised, and therefore it is selected as the domain of this exploratory experiment. ChatGPT4 is selected as the generative AI application, since it is one of the most widely used generative AI applications currently.

In order to evaluate the effectiveness of the thematic analysis produced by generative AI, this study compares the generative AI produced results with human-generated results (the benchmark), based on the same set of papers. The benchmark is peer-reviewed and therefore considered an appreciated baseline to compare against the generative AI produced result. This study uses the human-generated systematic literature review results from *Unlocking the value of artificial intelligence in human resource management through AI capability framework* (Chowdhury et al., 2023) published at Human Resource Management Review as the benchmark. This research paper conducted a systematic literature review and identified 18 themes based on 29 papers. This exploratory study will analyse 29 papers (from the Chowdhury et al., 2023 paper) following the same steps of identifying AI applications, barriers and drivers in HRM, using ChatGPT4 to identify and refine the key topics which will then be compared with the 18 topics in the baseline paper. The following section will describe the steps taken by this exploratory study and evaluate the generative AI produced results.

## **4.0 Steps and Results**

The Chowdhury et al., 2023 paper used a systematic literature review protocol suggested in existing literature (Hopp et al., 2018; Tranfield et al., 2003). In their paper a topic modelling algorithm known as Latent Dirichlet Allocation was applied, resulting in 69 topics that were initially found from a Scopus search of relevant papers. After manual interpretation and text analysis were applied to the original 69 topics, 18 were then considered meaningful. Appendix B of the Chowdhury et al., 2023 paper provided the 18 topics and a sample of papers that demonstrated those topics.

### **4.1 Preparation**

The first step in the process of analysing the sample literature review was to create a long summary in a narrative form (without headings) of each paper (500 words). This would allow the summaries to be combined for later analysis. It is feasible the summarising process was not needed as each paper in its entirety could be analysed in ChatGPT. However, when testing whole paper thematic analysis, the results were based on the paper headings and subheadings. For example, in Cubrics (2020) paper, *'Drivers, Barriers and Social Considerations for AI Adoption in Business and*



*Management: A Tertiary Study*", the following "topics" were identified through ChatGPT:

- Introduction to the Study
- Rise of AI
- Research Gap
- Methodology
- Research Questions
- Importance of Social Factors
- Conclusion and References

Therefore, it is considered not suitable to use this prompt to identify topics, as ChatGPT focuses on the heading in the paper as topics. Consequently we used ChatGPT to summarise the paper to generate a narrative for each paper for next steps of analysis. The prompt "*summarise the paper into 500 words*" was used, as it generated an appropriate overview of the paper. The summaries were then used for the rest of the experiment.

#### **4.2 Applications of AI in HRM**

To repeat the analysis carried out by Chowdhury et al (2023) to understand the applications of AI in HMR, we imported the summary of all 29 sample papers into ChatGPT and provided the following prompt: "*provide a list of applications of AI in HRM reported in this literature*". This resulted in the following results (Table 1, column 2), which are shown against the results from Chowdbury et al (2023) in Table 1, column 1.

<b>Original list of applications of AI in HRM</b>	<b>Generative AI analysis of literature</b>
Candidate Experience (Job Applications)	Talent Acquisition
Candidate Recruitment	Employee Engagement Analysis
Onboarding	Performance Evaluation
Employee Engagement	Predictive Modelling
Career Development	Training and Development
Employee Performance Appraisal	Enhancing Decision Making
Compensation Packages	Cultural Shift Towards AI in SBMS
Employee Skills Development	Ethical Considerations in HR Practices

Workforce Management Analytics	Support For Administrative HR Tasks
HR Budget and Resource Allocation	Workforce Analytics

**Table 1. List of applications of AI in HRM**

### 4.3 Drivers of AI in HRM

The next step in our research was to look for the drivers to AI adoption. In the benchmark paper the identification of drivers was not presented as an example table or list. Consequently the list outlined in table 2 has been created through human interpretation of section 3.4 in the original paper. To conduct the same analysis with our sample data, the summaries of our sample articles were inputted into ChatGPT with the prompt: “provide a list of drivers of AI adoption in HRM reported in this literature”. This is shown in Table 2, column 2 against the benchmark list in column 1.

<b>Original list of drivers of AI in HRM</b>	<b>Generative AI analysis of literature</b>
Objectivity and Accuracy	Technological Advancements
Enhancing Creativity and Innovation	Demand for Data-Driven Decision-Making
Streamlining Organisational Processes	Enhanced Employee Experience and Engagement
Recruitment and Talent Acquisition	Efficiency and Automation of Administrative Tasks
Reducing Bias in Candidate Screening	Talent Acquisition and Management
Monitoring and Performance Measurement	Predictive Analytics in Workforce Planning
Predicting HR Development	Personalisation of Learning and Development
Employee Retention	Employee Performance Evaluation
Operational Optimisation	Cultural Transformation and Change Management
Enhancing Dynamic Capabilities	Global and Cross-Border HR Management
Employee Well-being	Compliance and Regulatory Requirements
	Strategic HRM and Decision Support

**Table 2. List of drivers of AI in HRM**

### 4.3 Barriers of AI in HRM

Finally, we looked at the barrier to AI adoption in HRM. The analysis on this topic occurred in section 3.5 of the benchmark paper and again this requires human interpretation of the findings to produce a list of barriers. This is shown in table 3 column 1. To conduct the same analysis with our sample data, the combined summaries of our sample of articles was inputted into ChatGPT with the prompt: “provide a list of barriers to AI adoption in HRM reported in this literature”. This is shown in Table 3 column 2.

<b>Original list of barriers of AI in HRM</b>	<b>Generative AI analysis of literature</b>
Complexity of HR Phenomena	Data Privacy Concerns
Small Data	Talent Gap in Analytics Skills
Ethical Constraints	Cultural Resistance to Change
Employee Reactions	Ethical Implications and Bias
Privacy and Data Protection	Integration with Existing Systems
Constant Tracking Issues	Limited Access to Technology
Potential Bias in Algorithms	Digital Divide
Data Quality Assessment	Cost and Resource Allocation
Training Dataset Optimisation	Lack of Clear Business Case
Technological Integration	Regulatory and Compliance Issues
Developing a Data-Centric Culture	Uncertainty about AI Capabilities and Outcomes
Technology Turbulence	Misalignment between AI Solutions and Organisational Needs
Transparency and Interpretability	Employee Privacy and Consent Concerns
Epistemological Issues in AI-Driven Recruitment	Need for Interdisciplinary Collaboration
AI's Limitations in Creative and Social Intelligence	Job Security Fears among Employees
External Environmental Variables	
Human-AI Synergy Requirement	

**Table 3. List of barriers of AI in HRM**

## 5.0 Findings and Conclusion

Based on the experiment, it is considered that the summaries generated by ChatGPT were suitable, as they tended to capture the essence of the papers while providing more information than the abstracts of the papers. This allowed the experiment to continue to the next step of using generative AI to identify topics prior to comparing generative AI and human results. When it comes to identifying topics based on the paper summaries, generative AI produces some results that are highly similar to the human-generated results. For example, 'Recruitment and Talent Acquisition' was one of the topics identified by humans and it can be closely matched to 'Talent Acquisition and Management' which was produced by ChatGPT. Similarly, 'Workforce Management Analytics' can be closely matched to 'Workforce Analytics', and 'Employee Engagement' to 'Employee Engagement Analysis'. This high level of similarity can be observed for all three sets of experiments. This is not a surprise, as generative AI has been found to produce content that can be indistinguishable from human-generated content.

On the other hand, there are also some mismatches between human and generative AI results. For instance, 'Predictive Modelling' was one of the topics as AI applications in HRM identified by ChatGPT that cannot be linked to topics identified by humans. We considered that the term 'Predictive Modelling' is too broad as it could overlap with 'Workforce Analytics', which was also identified by ChatGPT. It is likely that humans would synthesise elements of predictive modelling into the form where predictive modelling was applied to, e.g. sales prediction or workforce planning.

Additionally, another interesting finding from this study was that similar numbers of topics are identified, without guiding prompts of the expected number of topics. For Applications of AI in HRM, both humans and generative AI produced 10 topics. For drivers of AI in HRM, humans identified 11 topics and generative AI produced 12 topics. For the barriers of AI in HRM, humans identified 17 topics and generative AI identified 15 topics. Based on this exploratory study, generative AI was able to synthesise the papers and narrow down the topics into a very similar number of items to the human results.

It is also noted that the prompt word 'topics' did not generate suitable results when the study tried to use ChatGPT to identify the key topics of each paper. ChatGPT picked up the headings, e.g. methodology, as the topics. While ChatGPT was able to identify

topics from smaller text body, i.e., 500-word summary of a paper, it did not identify appropriate topics from a larger text body, e.g. a full research paper. It is possible that different prompt words or sets of prompt words/questions could lead to more effective results when analysing a larger text body. This should be further explored and evaluated.

The results presented here are the initial findings in the exploratory study. As an exploratory study of a relatively small scale, the findings cannot be over generalised. Future work would consider the inter-rater reliability between the human reviewers and ChatGPT on a more significant scale, with a more complex set of prompt questions. This comparison would be key to understanding the difference between human and generative AI in its categorisation of papers into topics, as well as how scholars could utilise generative AI to synthesise a large amount of literature, which could potentially accelerate the speed of systematic literature review.

## References

- Alshami, A., Elsayed, M., Ali, E., Eltoukhy, A. E., & Zayed, T. (2023). Harnessing the Power of ChatGPT for Automating Systematic Review Process: Methodology, Case Study, Limitations, and Future Directions. *Systems*, 11(7), 351.
- Aydın, Ö., & Karaarslan, E. (2022). OpenAI ChatGPT generated literature review: Digital twin in healthcare. Available at SSRN 4308687.
- Boyd, A., Tinsley, P., Bowyer, K., & Czajka, A. (2023, June). The value of ai guidance in human examination of synthetically-generated faces. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, No. 5, pp. 5930-5938).
- Burger, B., Kanbach, D. K., & Kraus, S. (2023). The role of narcissism in entrepreneurial activity: a systematic literature review. *Journal of Enterprising Communities: People and Places in the Global Economy*, (ahead-of-print).
- Castiglioni, I., Rundo, L., Codari, M., Di Leo, G., Salvatore, C., Interlenghi, M., ... & Sardanelli, F. (2021). AI applications to medical images: From machine learning to deep learning. *Physica Medica*, 83, 9-24.
- Chambon, P., Bluethgen, C., Langlotz, C. P., & Chaudhari, A. (2022). Adapting pretrained vision-language foundational models to medical imaging domains. arXiv preprint arXiv:2210.04133.
- Chowdhury, S., Dey, P., Joel-Edgar, S., Bhattacharya, S., Rodriguez-Espindola, O., Abadie, A., & Truong, L. (2023). Unlocking the value of artificial intelligence in human resource management through AI capability framework. *Human Resource Management Review*, 33(1), 100899.
- Crossan, M.M. and Apaydin, M., 2010. A multidimensional framework of organizational innovation: A systematic review of the literature. *Journal of Management Studies*, 47(6), pp.1154-1191.

- De Cremer, D., & Kasparov, G. (2021). AI should augment human intelligence, not replace it. *Harvard Business Review*, 18, 1.
- Du-Harpur, X., Watt, F. M., Luscombe, N. M., & Lynch, M. D. (2020). What is AI? Applications of artificial intelligence to dermatology. *British Journal of Dermatology*, 183(3), 423-430.
- Dowling, M., & Lucey, B. (2023). ChatGPT for (finance) research: The Bananarama conjecture. *Finance Research Letters*, 53, 103662.
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642.
- Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity?. *Journal of Responsible Technology*, 13, 100060.
- Haman, M., & Skolnik, M. (2023). Using ChatGPT to conduct a literature review. *Accountability in Research*, 1-3.
- Hopp, C., Antons, D., Kaminski, J., & Oliver Salge, T. (2018). Disruptive innovation: Conceptual foundations, empirical evidence, and research opportunities in the digital age. *Journal of Product Innovation Management*, 35(3), 446-457.
- Li, J., Larsen, K., & Abbasi, A. (2020). TheoryOn: A design framework and system for unlocking behavioral knowledge through ontology learning. *MIS Quarterly*, 44(4).
- Liu, R., Gupta, S., & Patel, P. (2023). The application of the principles of responsible AI on social media marketing for digital health. *Information Systems Frontiers*, 25(6), 2275-2299.
- Mahdi, S. S., Battineni, G., Khawaja, M., Allana, R., Siddiqui, M. K., & Agha, D. (2023). How does artificial intelligence impact digital healthcare initiatives? A review of AI applications in dental healthcare. *International Journal of Information Management Data Insights*, 3(1), 100144.
- Manis, K. T., & Madhavaram, S. (2023). AI-Enabled marketing capabilities and the hierarchy of capabilities: Conceptualization, proposition development, and research avenues. *Journal of Business Research*, 157, 113485.
- Makarius, E. E., Mukherjee, D., Fox, J. D., & Fox, A. K. (2020). Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization. *Journal of Business Research*, 120, 262-273.
- Metz, A. (2023). exciting ways to use ChatGPT—from coding to poetry. TechRadar.
- Pan, S. L., Nishant, R., Tuunanen, T., & Nah, F. F. H. (2023). Literature review in the generative AI era-how to make a compelling contribution. *The Journal of Strategic Information Systems*, 32(3).
- Rahman, M., Ming, T. H., Baigh, T. A., & Sarker, M. (2023). Adoption of artificial intelligence in banking services: an empirical analysis. *International Journal of Emerging Markets*, 18(10), 4270-4300.
- Rossi, S., Kwon, Y., Auglend, O.H., Mukkamala, R.R., Rossi, M., & Thatcher, J. (2023). Are Deep Learning-Generated Social Media Profiles Indistinguishable from Real Profiles? *Proceedings of the 56th Hawaii International Conference on System Sciences*, 134–143. <https://hdl.handle.net/10125/102645>
- Rossi, S., Rossi, M., Mukkamala, R. R., Thatcher, J. B., & Dwivedi, Y. K. (2024). Augmenting research methods with foundation models and generative AI. *International Journal of Information Management*, 102749.

- Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays-should academics worry?. *Nature*.
- Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: many scientists disapprove. *Nature*, 613(7945), 620-621.
- Thieme, A., Hanratty, M., Lyons, M., Palacios, J., Marques, R. F., Morrison, C., & Doherty, G. (2023). Designing human-centered AI for mental health: Developing clinically relevant applications for online CBT treatment. *ACM Transactions on Computer-Human Interaction*, 30(2), 1-50.
- Tranfield, D., Denyer, D. and Smart, P., 2003. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British journal of management*, 14(3), pp.207-222.
- van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: five priorities for research. *Nature*, 614(7947), 224-226.
- Watson, R. T., & Webster, J. (2020). Analysing the past to prepare for the future: Writing a literature review a roadmap for release 2.0. *Journal of Decision Systems*, 29(3), 129-147.
- Yüksel, N., Börklü, H. R., Sezer, H. K., & Canyurt, O. E. (2023). Review of artificial intelligence applications in engineering design perspective. *Engineering Applications of Artificial Intelligence*, 118, 105697.
- Zhuo, T.Y., Huang, Y., Chen, C. and Xing, Z., 2023. Exploring ai ethics of chatgpt: A diagnostic analysis. arXiv preprint arXiv:2301.12867.

# **Media Choice for Multi-motive Communication: Impersonal and Interpersonal Motives in Social Worker-Client Relationships**

## **Abstract:**

This paper explores the adaptation of social workers to digitally mediated communication with their clients, amidst the constraints imposed by the COVID-19 pandemic. Focusing on the centrality of interpersonal relationships to social work, we examine how social workers in the UK and Hong Kong adapted to the transition from face-to-face meetings with clients to remote interactions via various computer-mediated communication (CMCs) channels. Through interviews with 37 social workers, we analyse the tension between maintaining close interpersonal relationships, and satisfying the need for task efficiency and regulatory compliance. Our findings reveal that social workers employed a strategic mix of digital media to fulfil both interpersonal and impersonal communication motives. We propose an integrative theoretical framework which links the dimensions of interpersonality—impersonality and agency—communion to understand how social workers' communication media uses co-evolve with multiple communicational motives. This research contributes to understanding the transformative implications of digital media on professional practices in relationship-centric work.

Keywords: communication media, multi-motive communication, interpersonal relationship, technological affordances, social work, remote working

## **Introduction**

Maintaining interpersonal communication objectives when the interactions are digitally mediated presents a shared challenge for knowledge workers. In social work, a profession to which interpersonal relationships are central (Rollins 2020), this issue has become increasingly urgent during the COVID-19 pandemic. The pandemic's social distancing measures forced social workers to rapidly shift to remote working, communicating with service users via information and communication technologies (ICTs) out of necessity when conventional face-to-face (F2F) service delivery were suspended (Hacker, vom Brocke, Handali, Otto, & Schneider, 2020). This disrupted practices and created dilemmas such as balancing the effective communication and ethical application of technology.

To understand how social workers adapt in this circumstance, we investigated two case sites in the UK and Hong Kong, where social workers displayed high adaptability in mediated remote interactions. Our data demonstrates that social worker made use of a



variety of available digital media to maintain a meaningful connectivity with their clients while attempting multiple communicational goals. Media uses were tailored around both service user needs and work responsibilities to balance a range of co-existing and competing motives, such as relationship quality, pastoral control, regulative requirements, and task efficiency. While these digital media expanded service options, critical perspectives cautioned against aspects such as over-reliance, boundary crossing and ethical negligence (Nordesjö, Scaramuzzino and Ulmestig, 2022). Following this line of inquiry, we specify our research question around the communicational intricacies exposed in the phenomenon: how do social workers' communication media uses co-evolve with multiple communicational motives, and how does this together influence interpersonal relationship building? Our research draws on communication media choice theories and on social cognitive theories of interpersonal relationships to propose an integrative theoretical framework for technology-mediated multi-motive communications.

## **Literature review**

### **Social work and multi-motive communication**

Interaction-rich and relationship-centred professional contexts have yielded fruitful insights in the scholarship of computer-mediated communication (CMC) and information systems. Previous research has investigated how professions with unique communication goals adapt to CMC in their practice. For example, tele-nurses carry out adaptive practices in ICT-mediated service delivery to maintain a sense of belonging to their team and work (Hafermalz & Riemer, 2021).

Social work has shown a historical reluctance to adopt digital technologies (Goldkind, Wolf, & Jones, 2016). The therapeutic relationship between a social worker and client is a central aspect of enabling change and bringing about positive outcome for the client, and some social workers may find it challenging to establish this relationship through digital means (Grubb, 2022). In spite of such general cynicism of CMC for effective relationship building, human service professionals sometimes find that therapeutic relationship can be strengthened through digitalized communication, particularly informal ICT use, in comparison to in-person interactions (Mishna, Fantus, & McInroy, 2017; Nordesjö, Scaramuzzino and Ulmestig, 2022). Digitalised interactions may also enable connectivity when collocated service delivery is not possible, as well as bring efficiency benefits. In addition, while relational closeness to clients is considered a preferred outcome in communication, research found that social workers may appreciate the boundaries created by the ICT that impersonalised their communication with clients (Grubb, 2022). This diversity in communication goals renders media choice and use in social work a

complex phenomenon that require nuanced examination of both social and technological influences.

### **Theories of media choice**

Past research on CMC in work context has focused on how workers choose between different media to achieve communication goals. One important premise of these theories is that the goal-oriented communicants are motivated to seek a 'fit' between the features of a communication technology—or a medium, which includes face-to-face interactions—and the task in hand (Fulk, Steinfield, Schmitz, & Power, 1987, Jung & Lyytinen, 2014).

To consider the media, many influential theories synthesised the various features or perceived characteristics of a communication medium into a unidimensional 'trait', such as the medium's richness (Daft & Lengel, 1986), synchronicity (Dennis, Fuller, & Valacich, 2008) or naturalness (Kock, 2004). On the side of the tasks, communication motives are categorised into various information processes, such as reduction of equivocality (Daft & Lengel, 1986), information conveyance and convergence (Dennis, Fuller, & Valacich, 2008) or mutual understanding (Kock, 2004).

Following this vein, achieving communication goals are often considered as a process of task-closure, assuming the communicants as rational information processor, overlooking the social relationship aspects in the interactions. There is, thus, a need to take into account how communication via different media impacts on the interpersonal relations, particularly in interaction-heavy work environments.

Also, despite the importance of social relationship, there are circumstances in workplace where communicants do not engage in relational communications, even actively avoid them (as illustrated in Grubb, 2022). This could be due to various factors such as the need for work-life boundary, urgent task closure or neutralising conflict. Traditional analysis often considers these motivational mechanisms alternative to one and another, where actors seem to be presented with clear options to decide on media use and its inevitable trade-offs. However, as demonstrated above in the social work context, communicants' media use motives are often rooted in both interpersonal and impersonal communication goals, and the co-presence of these goals complicates analysis (Greene & Kate Magsamen-Conrad, 2010).

### **Interpersonal and impersonal motives in CMC**

In CMC theories, a dichotomy of relational tone of communication media is visible, namely interpersonal relationships and impersonal relationships (Walther, 1992, Walther,

2011). Interpersonal communication aims to establish rapport and foster a sense of intimacy between the communicators. Tools like video conferencing, which allow for multiple non-verbal cues, are often considered to facilitate this. Theorisation around interpersonal relationships often focus on mimicking F2F interaction by establishing emotional and social presence. The Social Presence Theory (Short, Williams, & Christie, 1976) exemplifies this, positing that media allowing a high level of social presence foster interpersonal relationships.

On the other hand, impersonal relationships are task-oriented and minimise relational and social elements. The focus is on the transfer of information rather than the establishment of social ties. Email or text-based chats, where emotional cues are reduced, often deemed to serve this type of relationship. To exemplify, the Social Identity Model of Deindividuation effects, or SIDE model (Lea & Spears, 1992; Reicher, Spears, & Postmes, 1995), identifies the fact that when communication is held with visual anonymity (e.g., via text-based media), the participants would experience an absence of awareness with their own and their counterpart's individuality (i.e., deindividuation), aligning with impersonal communication.

### **Agency and communion as essential social needs**

Agency and communion are often posited as fundamental social needs that underlie human interaction and motivation (Bakan, 1966). Agency refers to the need for mastery, achievement, and control over one's environment. It can be viewed to embody the "self-as-separate" principle, emphasising individual competence and the capacity for autonomous action (Deci & Ryan, 2000). In contrast, the notion of communion describes a need for "self-as-connected", highlighting interpersonal bonds, emotional intimacy, and a sense of belonging (Clark & Mills, 1979). These needs can also be conceptualised into self-oriented and other-oriented goal pursuit (Abele & Wojciszke, 2014). This dichotomy is reflected in different literatures with more specific conceptualisation. For example, interpersonal connectivity, defined by Kolb (2008 p.184) as "the person-to-person contact we make with others," provides a framework for understanding how individuals engage in and maintain social bonds. The dynamic of achieving a 'maximal grip' (Hafermalz, 2021) in communication suggests an optimization of these interpersonal connections, wherein individuals strive for a balance between intimacy and separation, a balance between empathetic resonance with others and maintaining professional boundaries and distance.

Collectively, these theories reveal a complex interplay between the interpersonal desire for connectedness and agency, and the impersonal inclination towards autonomy and deindividuation in communication dynamics.

## **Method and research setting**

We interviewed 14 UK and 23 Hong Kong social workers using purposive and snowball sampling. Interviews occurred remotely in 2021 and 2022.

A semi-structured interview format was used to probe the adaptation to pandemic conditions, technology use, and lessons learned. In Hong Kong, we interviewed social workers specializing in elderly, youth, family, and school services, with many of the working with children or within elderly care home settings. The UK participants, employed by Leeds local authority, served in adult social care, including general adult services, health, and rapid response teams. The interviews focused on technology's role before and after social distancing measures, shifts in work practices, digital collaboration, and skill development support.

Data from the interviews, lasting 30-60 minutes each, were transcribed, anonymised, and thematically analysed with NVivo software, identifying significant patterns and themes.

## **Preliminary Findings**

The sudden shift to digital methods during the pandemic was initially met with scepticism by social workers, who doubted the effectiveness of online interactions, stating, *"We were quite sceptical about the output quality of online group work."* Concerns also arose regarding ethical issues, with practitioners noting, *"We worried about the ethical consequences of using technology in care homes for older people."* Despite these obstacles, they developed proficiency with various technologies over time, with the encouraging outcome that *"Those positive feedback [from care home service users] stimulated several colleagues' enjoyment in using ICT to undertake various activities."* They learned to adapt their media use, realising emerging technological affordances in-situ, for instance, *"Our music therapy group uses zoom, but usually there are family members to assist."* Overall, adapting to new technologies and practices for remote social work has resulted in perceived benefits to both service providers and users. This adaptation period also highlighted the profession's gradual shift in communicational motives when they balanced between organisational resilience and task efficiency, with service quality and ethical commitments. *"If we keep doing this it would become simply 'chasing the numbers', but as the headmaster [of the local school] said, it's better this*

*than no service at all.” “In terms of respecting diversity, we need to make extra efforts to ensure disadvantaged children's access to information and resources by ICT.”*

Many narratives reflected the profession's journey towards integrating ICTs effectively while maintaining the core ethos of social work. For example:

*“I think one of the subtle difficulties with that kind of technology [technology-mediated communication] is it normalizes, not seeing people physically. [...] But there's that whole kind of visceral information about seeing someone being with someone, knowing what it feels like to be with them, and to be in their space with them all of that is now missing very often.”*

In terms of media choice, different technologies were deployed strategically, with interactive media for counselling and one-way communication for low-risk cases, increasing overall accessibility and options:

*“And I think it depends on the situation as well, because I always think depending on what you need to speak to somebody about, sometimes you would choose an e-mail over a phone call because you want things documented and a clear record of that. Sometimes if it's a difficult situation, it's better to speak to somebody over the phone so you can kind of have that sort of actual conversation.”*

*“But then I think in a way it [communicating via email] can be beneficial because you can sort of get all the information there and then and think, when people are in a crisis points and you sort of there with them physically and you sort of saying things at them, I think sometimes [...] they're just absolutely overloaded with information that I'm giving them. But actually, when they have it broken down, for example, in an e-mail [...] I think it's just easier to read and they can get their head around or reread it.”*

Interpersonal relationship building takes time, which often competes with the need for efficient task closure. Consequently, social workers may gravitate toward communication media based on an impersonal, task-oriented motive:

*“So, if I know or I've got a good sense from my conversations with an individual that they have capacity and there is a lot of it. I've enjoyed working with people in that way where it's definitely strengths-based, where they're very wanting to strive for independence [...] And so in that in those contexts I've got where someone has, like I said has capacity and is knowing what they want from the [social work] service. And I've got no questions in terms of, like, assessing over the phone or video calling”*

As social workers expand service delivery to a range of CMC media, their communicational motives seem to evolve alongside emergent affordances of different channels. Making strategic choices, social workers navigate their relationship goals by impersonalising some, and ‘inter-personalising’ other interactions with their client. For example, email is chosen over a phone call because of a need to document and have a clear record, here the impersonal motive dominates. Phone calls may facilitate more sensitive, emphatic exchanges where visual anonymity reduces agentic intensity. On the other hand, in-person visits often remain the preferred approach where an interpersonal relationship, high in both empathic and agentic motives, needs to be developed, such as in new relationships.

### Discussion and theoretical framework (exemplary)

Based on the combined theoretical perspectives from CMC theories and social psychology, we propose a typological framework, consists of a dimension of interpersonality—impersonality and a dimension of agency—communion. As agency and communion are also conceptualised as self-orientation and others-orientation (Abele & Wojciszke, 2007), when applied to impersonalising motives they manifest as passivity—activity.

Communication motives	Example communicational practices (affordances)	Communication media
<b>Interpersonal–Agentic</b>	Monitoring Coordinating Persuading Compelling	Physical space (institutional, e.g., care home) Video recording Video meeting (multiple connection)
<b>Interpersonal–Communal</b>	Phenomenological copresence (‘Being there’) Supporting Caring	Physical space (personal, e.g., clients’ home) Text message (high connectivity) Video meeting Live streaming Social media (individuated)
<b>Impersonal–Passive</b>	Asocial task closure Information conveying Parallelism	Text message (low connectivity) Phone call Group call Email Social media (de-individuated)
<b>Impersonal–Active</b>	Boundary management Anonymising De-individuating	Text message (low connectivity) Email Group call

Arguing from our conceptualisation, when a social worker changes from using phone calls to group video conferencing to update clients’ conditions with their family members, the socio-material *nature* of this communication shift. The material aspects of the media

channel (e.g., parallelism, boundary management) results in higher transactional efficiency but decreased interpersonal relationship quality. Such affordances, upon reflection, can gradually reset the social worker's original goal of relationship-oriented service delivery to transaction-oriented service delivery (i.e., 'chasing the numbers').

## Conclusion

In this research, we intend to delve into the interpersonal elements of communication, particularly the empathetic and agentic dimensions that have been previously undervalued in professional settings. This examination will be linked to the unique affordances of various CMC channels. By recognizing these shifts in communication practices, our study anticipates revealing how altering the bounds of traditional F2F interaction can redefine the meaning of social work itself. This theoretical exploration sets the stage for future empirical research to assess how changes in the materiality of professional communication may fundamentally transform the nature of work.

## References

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*(5), 751–763. doi: [10.1037/0022-3514.93.5.751](https://doi.org/10.1037/0022-3514.93.5.751)

Bakan, D. (1966). *The duality of human existence: An essay on psychology and religion* (p. 242). Oxford, England: Rand McNally.

Clark, M. S., & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of Personality and Social Psychology, 37*(1), 12–24. doi: [10.1037/0022-3514.37.1.12](https://doi.org/10.1037/0022-3514.37.1.12)

Daft, R. L., & Lengel, R. H. (1986). Organizational Information Requirements, Media Richness and Structural Design. *Management Science, 32*(5), 554–571. doi: [10.1287/mnsc.32.5.554](https://doi.org/10.1287/mnsc.32.5.554)

Deci, E. L., & Ryan, R. M. (2000). The 'What' and 'Why' of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry, 11*(4), 227–268. doi: [10.1207/S15327965PLI1104\\_01](https://doi.org/10.1207/S15327965PLI1104_01)

- Dennis, A., Fuller, R., & Valacich, J. (2008). Media, Tasks, and Communication Processes: A Theory of Media Synchronicity. *MIS Quarterly*, 32(3), 575. doi: [10.2307/25148857](https://doi.org/10.2307/25148857)
- Fulk, J., Steinfield, C. W., Schmitz, J., & Power, J. G. (1987). A Social Information Processing Model of Media Use in Organizations. *Communication Research*, 14(5), 529–552. doi: [10.1177/009365087014005005](https://doi.org/10.1177/009365087014005005)
- Goldkind, L., Wolf, L., & Jones, J. (2016). Late Adapters? How Social Workers Acquire Knowledge and Skills About Technology Tools. *Journal of Technology in Human Services*, 34(4), 338–358. doi: [10.1080/15228835.2016.1250027](https://doi.org/10.1080/15228835.2016.1250027)
- Greene, K. & Kate Magsamen-Conrad. (2010). Disclosure Decisions in Existing Relationships Online: Exploring Motivations for CMC Channel Choice. In J. Park & E. Abels (Eds.), *Interpersonal Relations and Social Patterns in Communication Technologies: Discourse Norms, Language Structures and Cultural Variables*. IGI Global. doi: [10.4018/978-1-61520-827-2](https://doi.org/10.4018/978-1-61520-827-2)
- Grubb, A. (2022). Avoiding Intimacy—An Ethnographic Study of Beneficent Boundaries in Virtual Voluntary Social Work. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 33(1), 72–82. doi: [10.1007/s11266-021-00350-w](https://doi.org/10.1007/s11266-021-00350-w)
- Hacker, J., vom Brocke, J., Handali, J., Otto, M., & Schneider, J. (2020). Virtually in this together – how web-conferencing systems enabled a new virtual togetherness during the COVID-19 crisis. *European Journal of Information Systems*, 29(5), 563–584. doi: [10.1080/0960085X.2020.1814680](https://doi.org/10.1080/0960085X.2020.1814680)
- Hafermalz, E., & Riemer, K. (2021). Productive and connected while working from home: What client-facing remote workers can learn from telenurses about ‘belonging through technology’. *European Journal of Information Systems*, 30(1), 89–99. doi: [10.1080/0960085X.2020.1841572](https://doi.org/10.1080/0960085X.2020.1841572)
- Jung, Y., & Lyytinen, K. (2014). Towards an ecological account of media choice: A case study on pluralistic reasoning while choosing email. *Information Systems Journal*, 24(3), 271–293. doi: [10.1111/isj.12024](https://doi.org/10.1111/isj.12024)
- Kock, N. (2004). The Psychobiological Model: Towards a New Theory of Computer-Mediated Communication Based on Darwinian Evolution. *Organization Science*, 15(3), 327–348. doi: [10.1287/orsc.1040.0071](https://doi.org/10.1287/orsc.1040.0071)



Mishna, F., Fantus, S., & McInroy, L. B. (2017). Informal Use of Information and Communication Technology: Adjunct to Traditional Face-to-Face Social Work Practice. *Clinical Social Work Journal*, 45(1), 49–55. doi: [10.1007/s10615-016-0576-3](https://doi.org/10.1007/s10615-016-0576-3)

Nordesjö, K., Scaramuzzino, G., & Ulmestig, R. (2022). The social worker-client relationship in the digital era: A configurative literature review. *European Journal of Social Work*, 25(2), 303–315. doi: [10.1080/13691457.2021.1964445](https://doi.org/10.1080/13691457.2021.1964445)

Orlikowski, W. J. (2007). Sociomaterial practices: Exploring technology at work. *Organization Studies*, 28(9), 1435–1448. doi: [10.1177/0170840607081138](https://doi.org/10.1177/0170840607081138)

Walther, J. B. (1992). Interpersonal Effects in Computer-Mediated Interaction: A Relational Perspective. *Communication Research*, 19(1), 52–90. doi: [10.1177/009365092019001003](https://doi.org/10.1177/009365092019001003)

Walther, J. B. (2011). Theories of Computer-Mediated Communication and Interpersonal Relations. In M. L. Knapp & J. A. Daly (Eds.), *The SAGE Handbook of Interpersonal Communication* (4th edition). London: SAGE Publications Ltd.

# Unpacking the ‘Black Boxes’ of Analytics: A Data Value Map Analysis

## Abstract

*The paper reports on a three-stage research design pursued to uncover the shared understanding of practices linked to the execution of a Data & Analytics (D&A) strategy, within the \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* AME (Africa and Middle East) region. This research uses the Data Value Map (DVM) to present a visualised representation of the “experiential stories” of ten key informants. These key informants represent a snapshot of the important roles constituting the D&A team within the \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* AME region. The DVM analysis reveals eight ‘black boxes’ that represent the most frequently occurring topics of conversation across the ten key informant interviews. A reflection (by the Head of Data & Analytics for the AME region) on the impact of these ‘black boxes’ on the execution of the D&A strategy is also presented. The paper concludes with a sense of the practical implications linked to this applied research approach.*

**Keywords:** Business Analytics, Strategy Execution, Traditional Organisation, Transformation, Capability, Reflective Practice

## 1 Introduction

In recent years there has been growing interest in Business Analytics (BA) amongst the researcher and practitioner communities (Delen & Ram, 2018; Hindle et al., 2020; Mikalef et al., 2020). Defined by Stubbs (2013) as the use of data-driven insight to generate value, BA has been associated with superior organisational performance (Cao & Duan, 2017) and recognised as a competitive differentiator for organisations across most industries (Akhtar et al., 2019; Bumblauskas et al., 2017; Davenport, 2006). It is no surprise therefore that senior executives are increasingly focused on transforming their organisations to become data-driven (Delen & Ram, 2018; LaValle et al., 2011).

However, there remains much work to do, as organisations are struggling to become data driven (Davenport & Redman, 2020; Klee et al., 2021; Smith et al., 2019), with the New Vantage Partners (2023) Data and AI Executive Survey highlighting that less than 25% of organisations consider themselves to have already become data-driven. Traditional organisations, which were established in the pre-digital era and whose success has been achieved through non-digital business models, organisational structures and culture, face additional challenges, as their data-driven transformation can be impeded by those traditional structures and processes which had worked for them in the past (Gust et al., 2017).

\*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* has a clearly defined Data & Analytics strategy to become a data-driven organisation. However, bringing strategies to life is a significant challenge for organisations, as they struggle to close the gap between strategy and execution, often as a result of differences in thinking between those who create the strategy and those tasked with implementation (Bonchek, 2017; Kenny, 2019; Wiita & Leonard, 2017). Teams which are more successful in closing the *strategy-execution gap* are more likely to spend additional time engaging in dialogue and involving the people responsible for the implementation (Bonchek, 2017; Wiita & Leonard, 2017). Therefore, the objective of this paper is *to explore the strategy-execution gap of a regional unit within a global multinational organisation which is implementing a Data and Analytics strategy*. In particular, the study aims to assess the level of shared understanding and identify the enablers and inhibitors to the execution of the strategy, as key determinants of the strategy-execution gap. The study is based on the “experiential stories” of ten D&A team members from the \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* AME region. Their respective *individual* stories are told through the lens of the Data Value Map (a discursive template to guide data/analytics conversations). These individual maps are then analysed (following an inductive open coding approach) to produce a *collective* story of the D&A strategy execution in the \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* AME region.

The remainder of this paper is structured as follows. In the next section we present the background to the case followed by an overview of the three-stage research design being pursued. This is followed by a presentation of our observations (eight ‘black boxes’) gained through coding the “experiential stories” of the ten \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* AME region key informants (members of the D&A team) in Stage Two of the research design. Furthermore, the reflections of the Head of D&A for the AME region, in response to this analysis, are then captured as part of Stage Three of the research design. Finally, the paper concludes with a reflection on the implications for practice.

## 2 Background to the Case

\*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* is a leading developer and marketer of premium \*PRODUCTS-ANONYMOUS-FOR-REVIEW\*, with a portfolio of more than 300 international, regional, local and speciality brands. It has grown from the small family-owned business established by \*FOUNDER-ANONYMOUS-FOR-REVIEW\* in \*\*\*\* in \*\*\*\* to being the largest \*\*\*\* in Europe and the number 2 in the world. It employs over 80,000 employees and operates production facilities in more than 70 countries. Since 2021 the lead author has been Head of Data & Analytics (D&A) for the Africa, Middle East (AME) region with responsibility for leading an organisational data-driven transformation and maturing the D&A capability.

### 2.1 \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* AME Data & Analytics Strategy

The \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* AME region comprises sixteen Operating Companies (OpCos) of various sizes ranging from large OpCos such as \*OPCO-ANONYMOUS-FOR-REVIEW\* with many production sites and thousands of employees to much smaller OpCos with a single \*PRODUCTION-PLANT-ANONYMOUS-FOR-REVIEW\* and just a few hundred employees.

The Data & Analytics (D&A) strategy of the \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* AME region strives to unlock the value of data to achieve *HIGHER* impact on business performance, deliver *FASTER* time to insight, and grow a *STRONGER* organisation by holistically developing capabilities across the four pillars of the BAR (Business Analytics Recipe). The BAR outlines four pillars of business analytics capability, namely Solid Foundations (Data & Technology), People & Process (Skills, Org Structure, WoWs), Actionable Business Value (Analytics Use Cases which are actionable, feasible and deliver a positive value impact to organisational performance) and Organizational Culture (Data Driven Mindset and Literacy).

Higher calls for achieving *HIGHER* impact by focusing on the business outcomes of BA rather than inputs such as technology deployments or building BA solutions. Faster is about delivering *FASTER* time to insights so that business decision makers have a data solution in their hands as soon as possible after an opportunity or business problem

is identified. There are two elements to growing STRONGER together, namely strengthening cross functional collaboration between data and business teams and also strengthening the data foundations of data availability and quality, along with robust technology platforms.

The strategy is brought to life through a series of initiatives linked to the BAR. For example, under the People & Process pillar an initiative was to establish above-OpCo capabilities in the form of Regional Data Management and Analytics Hubs to provide OpCos with the specialised D&A skills such as Data Engineers, Data Scientists, etc. The Data Management Hub (located in Egypt) and the Analytics Hub (located in South Africa) were both established in the second half of 2021. Other examples of initiatives include implementing a Governance process for Analytics Use Cases under the Actionable Business Value pillar and deploying a cloud data platform under the Solid Foundations pillar.

In 2023, two years after the launch of the AME D&A Strategy, significant progress had been made. AME had been unique amongst the four \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* regions as the only one with a Head of D&A role, D&A Strategy, Regional D&A Hubs, etc, and was outperforming in terms of realising value from data. In 2022, despite accounting for 13% of global revenue, the AME region delivered 60% of global value from analytics use cases. However, by mid-2023 the lead author as Head of D&A had formed the view that some elements of the strategy were at that time not being realised as well as others. This was in contravention of the Holistic principle of the BAR which states that all elements of the BAR are essential for success and need to be developed holistically to prevent a gap emerging between the organisations ability to produce and consume analytics. Areas of concern included the speed of deployment of analytics use cases was not accelerating as anticipated, value delivery while continuing to grow was starting to drift behind the ambitious targets, and embedding analytics solutions in long established business processes was proving challenging. To assess if the strategy was as well understood by the D&A Team (responsible for the strategy implementation) as had been perceived by the Head of D&A (lead author), it was decided to engage the teams by conducting a Data Value Map (DVM) analysis as outlined in the following sections. The outputs would then be used

to inform future actions such as amending / recommunicating the strategy or course corrections in the implementation.

### **3 Methodology**

This section presents the Data Value Map (DVM) and the use of the DVM discursive template as part of a three-stage research design.

#### **3.1 Data Value Map**

The mindset to transform information into a “*critical business asset*” (Laney, 2018, p. 10) does not always dominate the ways business manage their information assets. Therefore, weaponising information as opposed to just using it (c.f. Laney, 2018) can be a significant business challenge. In fact, Laney (2018, p. 12) comments that “*myths create cognitive roadblocks that hinder business leaders from realising anything near the full promise of information*”. Being able to see through the “*cerebral fog of these myths*” Laney (2018, p. 12) is needed and taking advantage of ways of seeing what we say (a concept popularised by Karl Weick in his work on organisational sense-making). Once such device that offers this visibility is the Data Value Map (DVM).

The Data Value Map (see Figure 1) is a structured discursive template (c.f. Sammon & Nagle, 2017) that positions the key components of the often referred to *information supply chain* (c.f. Laney, 2018) or *information chain*. (c.f. Redman, 2008) The four components of the information supply chain (ISC) are *acquisition*, *integration*, *analysis* and *delivery*. Each of the components serves a specific purpose, where *acquisition* details the gathering of data from business activities; *integration* describes the combining of datasets from numerous sources; *analysis* describes the processing of analytics on subsets of data; and finally, *delivery* focuses on supplying analytical insights in a suitable format. These four components are book ended by two human actors, namely: the *data creator* and the *data user*. The data that flows from the data creator to the data user must be (i) of high quality and (ii) put to use. This demands that a shared understanding of data behaviours and business value exists (between the data creators and the data users) in order to frame the four components of the information supply chain. To do so goes a long way to protecting against the much-bemoaned

concept of GIGO (garbage in, garbage out), where the quality of the output is determined by the quality of the input.

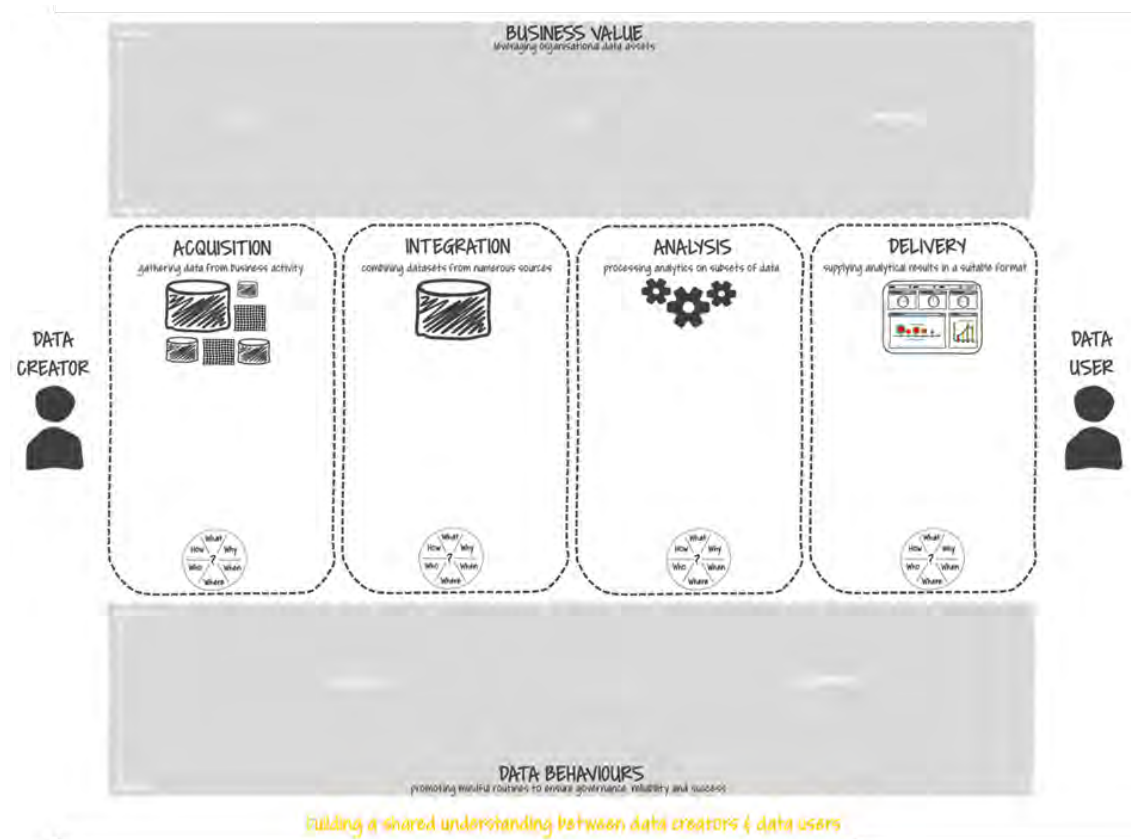


Figure 1: The Data Value Map (DVM) Discursive Template

The Data Value Map (DVM) enables a visual sense-making process of unfold where all too often competing multi-stakeholder conversations are taking place between business and technology. Therefore, through answering a series of simple probing questions along the four components of the ISC, all stakeholders are in a better position to see what they say. The alignment of *people*, *process*, and *technology* with the capability to organize, govern and share data to achieve business ends is fundamental to promoting data-driven conversations using the DVM (c.f. Sammon & Nagle, 2017). To aid these conversations, six very simple questions (*why*, *what*, *when*, *who*, *where* and *how*) ensure all underlying assumptions are at least questioned if not fully examined. Therefore, answers to questions can be placed on the DVM to create an elaborate visual to effectively communicate the sense that stakeholders make of the data. The very act of going through this rigorous process of questioning ensures that every implicit

assumption is questioned, therefore challenging the status quo, in the pursuit of unlocking a value driven data conversation.

### 3.2 The Three-Stage Research Design

Stage One of the research design involves gathering data from ten key informants. These data were gathered using semi-structured interviews. The questions invited the key informants to share their role in the analytics production lifecycle and their sense of ‘what works’/‘does not work’ in the execution of the D&A strategy. Those interviewed across the AME region share their “experiential stories” from various perspectives (local/region/global) and positions (e.g. Data Management Hub, Data Analytics Hub, OpCo) within (\*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* AME region). These key informants were selected not for their “representativeness” alone, but for their “informedness and ability to communicate” (Campbell, 1955, p. 339). Therefore, using the informant technique typically means that the researcher gathers data from a person who performs an organisational role and is well informed and well able to speak the language of the researcher.

This approach affords us the opportunity to “capture the meaning” from those practitioners “living the experience” (leading a D&A initiative in a traditional organisation) and “theorize about that experience” (Gioia et al., 2013, p. 17). Being inspired by features of the Gioia Methodology (as a “systematic inductive approach to concept development” and assumes that “the organisational world is socially constructed” (Gioia et al., 2013, p. 17)), we had an ambition to conceptualise the practitioner voice and to not “substitute practitioners’ understandings for theory” (Markus et al., 2021, p. 273). As a result, in our data collection we give “extraordinary voice to informants” where we view them as “knowledgeable agents”. As illustrated in Figure 2, these stories provide great coverage of the D&A strategy and reveal interesting observations right across the DVM (*acquisition, integration, analysis delivery, business value, and data behaviours*). All interviews took place between February and April 2023, and lasted between 45 and 90 minutes. The interviews were conducted by two members of the research team (excluding the lead author who is the Head of D&A for the AME region and was also interviewed as part of this research stage).



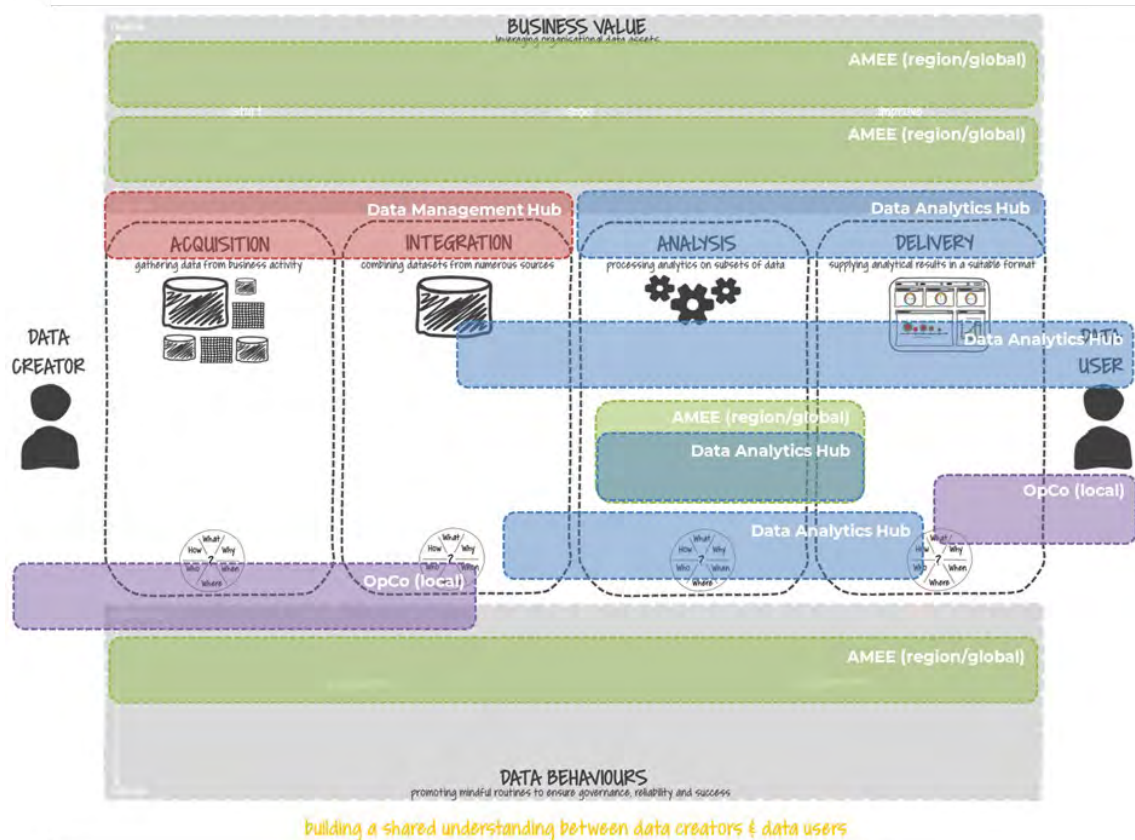


Figure 2: Distribution of Key Informants

Once all interview data was gathered, both team members also coded the data using an inductive open coding approach. Initially, we maintained “*the integrity of 1<sup>st</sup> order (informant-centric) terms*” when coding the ten interview transcripts during data analysis (Gioia et al., 2013, p. 26). Thereafter, as we progressed in our analysis of the data, we further organize the “*1<sup>st</sup>-order codes into 2<sup>nd</sup>-order (theory-centric) themes*” (Gioia et al., 2013, p. 26) using the Data Value Map as an analytical frame. This approach also afforded the opportunity to place an emerging theme (raised by a key informant) on the DVM (placing it where it is most relevant). Therefore, over the course of several rounds of coding, a rich picture of the analysis emerged (see Figure 3). This DVM rich picture showcases eight ‘black boxes’ that emerge as requiring further discussion. These eight DVM analysis ‘black boxes’ include the following: (i) *ERP System (data quality)*; (ii) *Commerce Systems (data quality)*; (iii) *Data Prime (Azure DB)*; (iv) *Harmonised Data Pipelines*; (v) *AVC Use Cases*; (vi) *Common Business Language*; (vii) *Translators (local, region, global)*; (viii) *POC Driven (faster, stronger, higher)*. This visualization of the eight ‘black boxes’ positioned on the DVM now

provided an opportunity to build a shared understanding amongst the D&A team members (which was the focus of Stage Two).

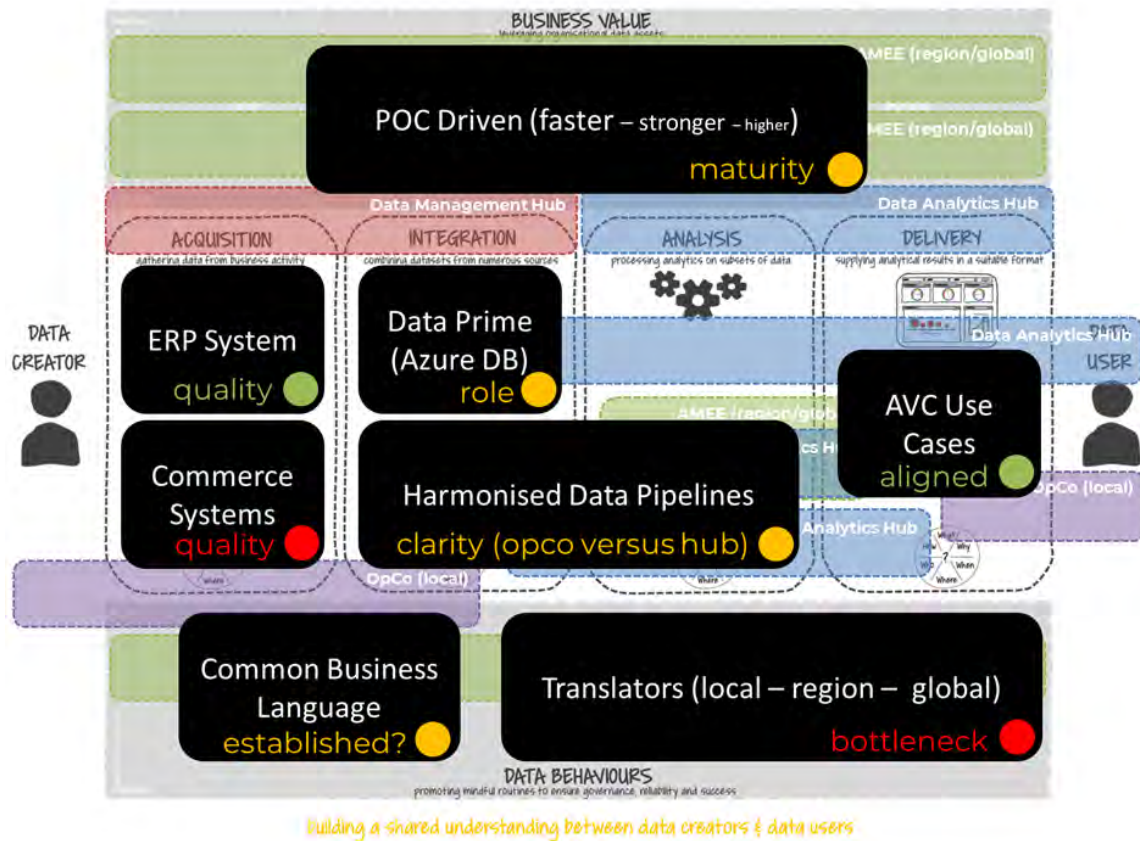


Figure 3: The Eight ‘Black Boxes’ of the DVM Analysis

For Stage Two, the two research team members (excluding the lead author) presented their analysis to the D&A cohort (within the AME region). This presentation took place on June 14<sup>th</sup>, 2023, and lasted for 90 minutes. This was viewed as a further opportunity to build a shared understanding within the D&A cohort. As part of this feedback presentation, each of the eight ‘black boxes’ were unpacked and the reason for the RAG (red/amber/green) status was explained. Finally, in Stage Three of the research design, the Head of D&A for the AME region (the lead author) reflected on the learnings from the DVM analysis. Specifically, they focused on the impact of the eight ‘black boxes’ on the execution of the D&A strategy within the AME region.

The outputs of Stage Two and Stage Three are now presented in the next section.

## **4 Findings & Discussion**

### **4.1 DVM Analysis (unpacking the ‘black boxes’)**

This section presents the work associated with Stage Two of the research design.

#### **4.1.1 Data Quality in the ERP System & Commerce Systems**

Both these systems contain data of interest to the analytics ambitions within the AME region. However, while key informants suggest that the data in the ERP System was of good quality, it was the data in the Commerce Systems that was of greater value (to analytics), but of lesser quality. There were several reasons for this reality within the region, however, historically, it seems that more effort has been focused on data quality within the ERP system (e.g. where data standards have been well defined over the past decade). It was also revealed that executing the analytics use cases has triggered the surfacing of data quality issues in the Commerce Systems, specifically. Therefore, as part of the presentation, the researchers concluded that the Commerce Systems data are just as important as the ERP System data for strategic success (e.g. data-driven insights to become the best-connected \*\*\*\*\*). As a result, they also questioned if a plan was in place to rectify these data quality issues? Data quality was viewed by several key informants as the biggest barrier to speed, therefore, the suggestion is that better data quality will make them faster (as per their D&A strategy).

#### **4.1.2 Data Prime**

Data Prime was mentioned by all key informants, however, their descriptions of what it is (and its role) vary considerably, with Data Prime being described as a data warehouse solution, a data lake, a cloud-based system, an ETL process to harmonise data, and so forth. Notwithstanding this, it is revealed throughout the course of the interviews that Data Prime is in fact a programme of work to create a platform (an Azure Database) to ingest data from different sources (e.g. move OpCo data to the cloud) in order to build analytics solutions. This would ensure that the required data is always available (to access) for faster use case deployment. Therefore, the ambition is that all AME region data (required for analytics) will reside on this platform, and one key informant revealed that 75% of data are currently on Data Prime. Therefore, as part

of the presentation, the researchers concluded that there is a lot of uncertainty amongst the key informant voices as to (i) what Data Prime actually is, and (ii) its role in delivering analytics use cases (e.g. does it deliver value).

### **4.1.3 Harmonised Data Pipelines**

The Harmonised Data Pipelines (HDP) was mentioned by several key informants and is described as a stepping-stone to the adoption of Data Prime (a global initiative) throughout the AME region. Therefore, it appears as if local OpCo data is loaded onto HDP, which is viewed as an AME region solution, that is architecturally similar to Data Prime. HDP is a solution to a Data Prime ‘bottleneck’ problem, where some analytics use cases demand more data than Data Prime might currently have available. Therefore, as part of the presentation, the researchers concluded that HDP connects integration and analysis (two components on the DVM) where HDP might be more like a ‘data mart’ (region solution) and Data Prime more like a ‘data warehouse’ (global solution). However, there is also a lack of clarity amongst the key informant voices as to the relationship between HDP and Data Prime in delivering analytics use cases (e.g. does it deliver value), and a lack of a shared understanding from a local to region level. As a result, the researchers also questioned if cleaning the “same” data (80% of time) is a never-ending prospect to deliver data-driven insights (20% of the time) as the best-connected \*\*\*\* (given the co-existence of HDP and Data Prime)?

### **4.1.4 Common Business Language**

A small number of key informants suggested that while Data Prime can make access to data faster, there is a pressing need for a common business language to make things stronger. One key informant suggests that people speak a different language, even within the same function (e.g. sales) across the region, and sometimes to describe the same things! Therefore, as part of the presentation, the researchers concluded that there is a need for a ‘shared’ common business language to support the delivery of analytics use cases (e.g. to deliver value). As a result, they also questioned if the AME region is structured appropriately to align with the data-driven ambitions? For example, when/where does the transformation of data (from source system to target system take place, and is the global/region/local legacy (processes, systems, data, people) costing

more than it should (to align with the strategic data vision: data-driven insights to become the best-connected brewer)?

#### 4.1.5 Translators and AVC Use Cases

The approach to proposing, evaluating, and prioritising use cases appears very robust (with an obvious alignment to the D&A strategy) within the AME region. In fact, key informants suggest it is an exemplar for other regions. Notwithstanding this, a point of difference did exist as to whether analytics use cases were ‘bottom-up’ versus ‘top-down’. The ongoing desire to deliver use cases faster also exists, and several reasons are provided as ‘bottlenecks’ to speed, from the quality of the data to the availability of the data, on Data Prime. However, a further reason is highlighted by some key informants and centres on a lack of ‘*Analytics Translators*’ within the region. As commented by a key informant, the translator (the region level – Analytics Hub) is viewed as the bridge between local (the daily reality – the practice) and global (the strategic ambition – the theory). A further appreciation of the role of the ‘*Analytics Translators*’ is provided by another key informant who suggests that they translate from local (OpCo) to region (Hub) and it is this relationship (between local and region) that generates value, in essence, relationship building makes things faster! Therefore, being “*close to the action*” and “*building domain knowledge*” enables “*assumptions about the reality to be challenged*” and a “*better sense of the data to be made*”. However, other key informants suggest that communication around the data landscape is the biggest challenge and clarifying the shape of this landscape is enabled by the translators, although there is a shortage of these resources. Therefore, the current small number of ‘*Analytics Translators*’ is a bottleneck to faster and higher, as they are the analytics use case owners. Therefore, as part of the presentation, the researchers concluded that a lack of ‘*Analytics Translators*’ is the biggest barrier to speed (e.g. data-driven insights to become the best-connected \*\*\*\*). As a result, they also questioned if a plan was in place to rectify this issue, suggesting that more translators will make the region faster (as per their D&A strategy).

#### **4.1.6 POC Driven**

The execution of the D&A strategy was conceptualised by several key informants as a journey. For example, the impact of a “*history of a decentralised data landscape*” is being exposed as part of the execution of the D&A strategy, and in many instances “*the analytics use cases triggered the surfacing of data quality issues*” (in Commerce Systems specifically). Therefore, as part of the presentation, the researchers concluded that the maturity of what is being done is aligned with a *Proof-of-Concept* approach. As a result, they also questioned the plan to move from shorter-term “POC driven” practices to longer-term delivery of sustainable value (as the best-connected \*\*\*\*\*) for stakeholders? In essence, the ‘value to feasibility’ trade-off (as regards the criteria used to accept/reject a use case) needs to be appreciated in the context of a ‘path to value’ approach for the next step in the journey of the D&A strategy execution.

#### **4.2 Head of D&A (Practitioner) Reflection on the eight ‘black boxes’**

This section presents the work associated with Stage Three of the research design. Despite having commissioned the DVM analysis, the Head of D&A for the AME region (lead author) was not provided with the outcomes in advance of the results sharing workshop (delivered by the two-member research team in June). This was intentional to ensure that the results shared with the wider D&A team represented the views of the teams implementing the D&A strategy as objectively captured by the external researchers, and unbiased by the views of the Head of D&A. Following the workshop, the identified key concepts (eight ‘black boxes’) along with the status of each was reflected upon by the Head of D&A (lead author) as outlined in Figure 4.

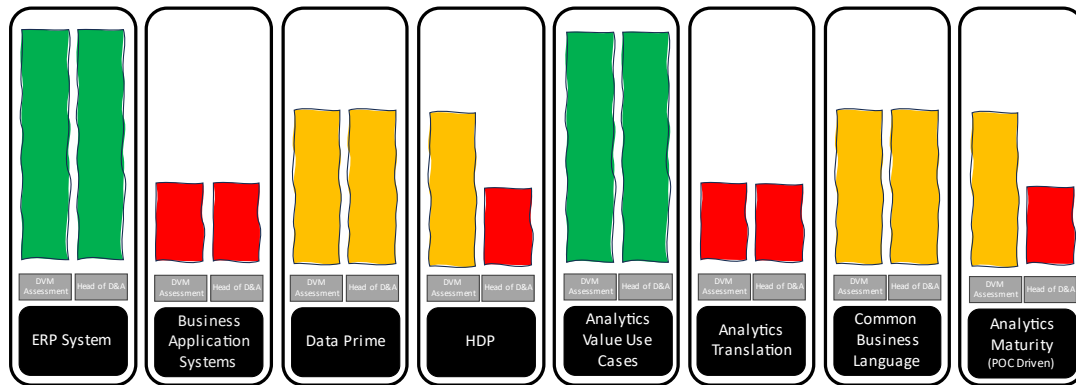


Figure 4: Comparing the Eight 'Black Boxes' from the DVM Analysis

#### 4.2.1 ERP System

##### Head of D&A RAG: Green

Most of the data currently used for BA is data from the ERP systems of the OpCos and is predominately transactional data such as sales invoices. The acquisition of data from ERP systems was assessed as Green by the DVM analysis and as Head of D&A, I would concur with the finding of the research team. In general, this data is of relatively good quality and can be relied upon with confidence to make decisions. There are several reasons for this, firstly the ERP systems are used to enable the business to operate, so any errors are highlighted in the course of daily operations. For example, a customer will not pay an invoice for goods which were not received, so if there are errors, they are quickly corrected. Financial reporting is also based on ERP data so historically there have been robust controls put in place to ensure the accuracy for business control purposes. Furthermore, in recent years a programme was completed in which all the small and medium OpCos in Africa standardised on a common ERP which has led to greater standards, consistency across OpCos and improved processes, with a data quality score in excess of 99%. Notwithstanding that there are certainly issues of data quality, particularly in some less mature OpCos, the standard of data quality is generally high.

#### 4.2.2 Business Application Systems (Commerce Systems)

##### Head of D&A RAG: Red

As \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* continues to digitise its business operations, several business application systems have been introduced in

recent years. As we focus on digitising our route to consumer, many of these applications are in the commerce area such as e-commerce B2B systems, call centre systems, distributor systems, but also include other areas such as warehouse management systems. In contrast to ERP data, the quality of data acquired from Business Application systems was rated as Red by the research team. As Head of D&A this was not unexpected as Business Application data quality is recognised as a challenge impeding business analytics and there are initiatives in place to address. It was the case in the past that some of these applications were deployed on a piecemeal basis in response to operational requirements and priorities in particular OpCos. Individually, these Business Applications were successfully implemented and met the needs they were initially designed to meet. For example, the distributor management system works as designed and allows distributors to place orders with \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\*, manage their own stocks, plan their routing, report their sales, and so on. However, when it comes to using this data for data driven decision making, challenges soon occur. Different distributors utilise the application in different ways with some fully utilising and others continuing to manage their business on paper and just using the system to place orders with \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\*. Another example is distributors might set up a new customer on the system even though the same customer might be already recorded in the \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* ERP and even the distributor system as other distributors might also be delivering to the same customer. These non-ERP applications contain a rich source of data which is essential for impactful analytics beyond traditional reporting and BI, so addressing the data quality and consistency issues across the business applications has been prioritised, with several initiatives across technology, process re-engineering, quality measurement, etc. initiated in the past 18 months.

#### **4.2.3 Data Prime**

##### **Head of D&A RAG: Amber**

Data Prime is a programme with several components including deploying an Azure Cloud database along with a suite of data management tools for data governance, data lineage and data quality. It also includes some basic training on the tools. In the area of data integration, the Data Prime initiative is assessed by the research team as amber,



which is an assessment which I as the Head of D&A concur with, but for different reasons. The narrative around Data Prime was that it would deliver data driven decision making, when it actually only delivers technology solutions which enable business analytics. There are many more elements required in a Business Analytics capability beyond just the technology. Therefore, from an OpCo perspective, stakeholders struggle to see the value of Data Prime and it is not surprising to hear it described as a solution looking for a problem. Recognising the fallacy of attempting to deliver a technology led business analytics capability, the approach in AME was changed. Data Prime in AME was set up to not only deliver the technology platform and the data from ERP, but also data from the top 3 Commerce Business Application Systems along with an initial analytics use case. The project approach was also adjusted to ensure the programme was completed five times faster than the next fastest region. From the perspective of the Head of D&A, the status of Data Prime has improved from Red to Amber as it has been successfully completed and is delivering value by enabling faster deployment of use cases in line with the regional Higher Faster Stronger D&A Strategy. It is not Green, because while the data is more accessible, it remains unharmonized so is still not fully analytics ready.

#### **4.2.4 Harmonised Data Pipelines**

##### **Head of D&A RAG: Red**

As previously mentioned, Data Prime is delivering accessibility to data from the key business applications across \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\*. However, the data remains unharmonized so is not ready to be used in business analytics use cases. To resolve this challenge, a relatively new initiative called Harmonised Data Pipelines has been established to build pipelines which harmonise the data required for business analytics use cases. While the data is sourced from the Data Prime cloud, it is technically possible for HDP to integrate data from applications which are not ingested to Data Prime. Given that the programme is relatively new and there have been many changes in the six months prior to the research interviews, it is understandable that the research team would rate it as Amber. However, as the Head of D&A, I rate the current status as Red, due to the slow progress to date, the challenges facing the ambitious aspirations of the programme due to the complicated data landscape across the region, and the insufficient level of the resources currently allocated by the HDP Team to

deliver on the aspirations. While the challenges are recognised, they are being addressed by the Global Teams responsible for HDP, so the expectation is that status will very quickly improve.

#### **4.2.5 Analytics Value Use Cases & AVC**

##### **Head of D&A RAG: Green**

The Higher element of the AME D&A Strategy calls for a focus on impact over inputs. Strengthening the governance of analytics use cases has been one of the most impactful and transformative strategic initiatives undertaken to date. The DVM research team assessed this element as Green and as Head of D&A I concur. Prior to establishing the Analytics Value Council, the Business Analytics teams were at full capacity delivering solutions, but the utilisation and value realised was mixed. The AVC brings together the senior leaders from the Data & Analytics and the business teams to jointly prioritise use cases, track progress and remove any blockers which are escalated by the teams to the AVC. Use cases are prioritised based on the value which they are likely to create, the technical feasibility given the available data, quality of the data, technology, and the skillset of the teams and the actionability (do the business decision makers have the capability to implement the solution). This has led to cross-functional alignment, ensured the prioritisation of business analytics solutions which are delivering a significant impact on business performance and provided clarity to the BA Teams on the value and purpose of their contributions. The implementation of robust and transparent governance for BA has been recognised across the business as a significant driver of the BA success to date.

#### **4.2.6 Analytics Translators**

##### **Head of D&A RAG: Red**

The role of the Analytics Translators, or Analytics Product Managers is recognised as possibly the most crucial role in the Analytics Hub, particularly as the Analytics Hub is shifting from a provider of technical analytics solutions to a BA centre of excellence focused on impacting business performance by leveraging the power of data. As Head of D&A, I assess the status of this element as Red which is the same as the assessment of the DVM Research. The existing Translator resource is performing very well in terms

of partnering with business to identify opportunities, refine business questions and partnering with the business and technical teams through the stages of data acquisition, integration, analysis, and delivery. The scope of the role is end-to-end and is playing a leading role in ensuring BA is actioned with value realised and measured. However, the challenge is that there is not enough resource, and this constraint has created a bottleneck and is impacting on the delivery of faster time to insight as called for by the D&A Strategy. Therefore, while there is alignment that the status of this element is Red, additional resource is currently being recruited so this status is expected to be short term. Furthermore, as currently structured, most of the translation is conducted by the Analytics Portfolio Managers (Previously known as Analytics Translators) which are attached to the Regional Analytics Hub or the Global Analytics Teams, which are developing the BA Use Cases. The Regional D&A Strategy calls for this translation activity to be located closer to the business decision makers, so at the OpCo level there is a requirement to further build this capability.

#### **4.2.7 Common Business Language**

##### **Head of D&A RAG: Amber**

The Common Business Language (CBL) is a collection of data related standards, terms and rules described in clear language that everyone across the organisation can understand. The CBL is required in order to deliver a solid data foundation as per the D&A Strategy, as it allows people across the organisation to better communicate and collaborate, drives consistency in interpretations, allows shared understanding of data and eliminates disconnected static views of data. The DVM Assessment rated this element as Amber, primarily due to a lack of clarity around the status of the CBL and as Head of D&A, this is a fair assessment. The creation of the Common Business Language is recognised as a priority and is being developed by the global data management team. Good progress has been made to date with the data governance and cataloguing tool (Collibra) now delivered as part of the Data Prime programme, but it is not yet being widely utilised beyond the Data Management Teams. Furthermore, work continues to define the Data Standards, Business Terms and Business Rules, so while these are in place and available for some domains, the CBL is not yet fully developed.

#### **4.2.8 Analytics Maturity Stage - POC Driven (Realising D&A Strategy Faster, Stronger, Higher)**

##### **Head of D&A RAG: Red**

The D&A Strategy is to deliver higher impact through scalable value enhancing use cases which are actioned by decision makers. This calls for analytics solutions to be embedded in the business processes. The DVM Assessment rated this element as Amber, but as Head of D&A, this element would currently be defined as Red. It is the case that some analytics use cases have been scaled across multiple OpCos and are embedded in decision making with proven value. However, these use cases are in the minority to date, and are concentrated in just one domain. A significant number of use cases still struggle to move beyond the POC or pilot phase, even when the POCs are successful and deliver value. The reasons are many and varied and range from lack of consistent data quality and standards to a lack of capability and / or engagement amongst the business decision maker community. Moving the maturity beyond the current status is not a quick fix and requires a holistic approach with capability building across the entire D&A Ecosystem including data, technology, people competencies, ways of working, overall organisational culture and capability. It is for this reason, i.e., the complexity of the solution requiring advances across the entire DVM, that this element is currently rated red by the Head of D&A.

## **5 Conclusions: So What?**

The DVM analysis outlined in this paper set out with the objective of providing \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* AME with an assessment of the D&A strategy-execution gap, both in terms of how well the D&A strategy had landed with the people tasked with implementation, and in identifying enablers and inhibitors. It was expected that this DVM analysis would inform any course-corrections to ensure a successful D&A strategy execution. The strategy was found to be well understood by the teams, thereby validating the strategy communication and alignment. The DVM analysis also identified, from the perspective of those implementing the strategy, the areas requiring attention and ‘course correction’ (see Table 1). Overall, the approach taken (undertaking a DVM analysis) also provided a framework to structure observations by mapping the key areas (black boxes) to the DVM components.

The biggest shifts in the focus areas of the AME D&A Strategy following the DVM Analysis are in terms of breaking from the POC Driven approach and building capability at the end user (decision makers in OpCos) level. The observation of the lead author is that while a good D&A strategy was developed, the organisation moved too quickly to over focusing on developing and deploying analytics use cases and did not focus enough on building the organisational wide capability to embed analytics in a sustainable way. It was described as starting to build a house with good plans but then trying to move into and live in the house before construction was completed.

The three primary criteria for prioritising use cases had been Value (the impact of the use case on business revenue growth or cost savings), Scalability (potential to deploy use case across multiple OpCos), Actionability (end users ability to do something different resulting from the use case). Delivering Scalability in practice proved to be more difficult to achieve than had been envisioned in the strategy. Differing data foundations and levels of end user capability meant that it was not easy to deploy use cases developed in one OpCo across multiple OpCos. While a significant business transformation programme is in progress to simplify and standardise ERP, business applications and processes across all \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* OpCos, which will enable scalability in analytics use cases, this has not yet been completed. In the meantime, following the DVM analysis, the criteria for prioritising use cases have been amended with feasibility replacing scalability in the top 3 criteria. Feasibility includes assessing how feasible it is for the OpCo to implement and embed the analytics solution in the everyday business process rather than as a one-off ad hoc piece of analysis.

Table 1: Summary of 'Black Boxes' and resultant actions

<b>Black Box</b>	<b>DVM Component</b>	<b>Strategy-Execution Reality (So What?)</b>	<b>Inhibitor / Enabler</b>	<b>Action Plan (What Now?)</b>
ERP System (data quality)	Acquisition	Clarity and alignment on role of ERP with good level of data quality enabling analytics	Enabler	Maintain
Commerce Systems (data quality)	Acquisition	Recognised pain-point impeding strategy implementation	Inhibitor	Continued prioritisation of initiatives to address data quality
Data Prime (Azure DB)	Integration	Lack of clarity on role of Data Prime and the value delivered	Inhibitor	Reassess strategic position of Data Prime and role in overall D&A Strategy
Harmonised Data Pipelines	Integration & Analysis	Lack of clarity on role of HDP / difference to Data Prime and challenged with HDP delivery	Inhibitor	Re-establish alignment on role of HDP, and plan to accelerate development
AVC Use Cases	Delivery	Recognition that analytics value governance is working well	Enabler	Maintain
Common Business Language	Data Behaviours	Recognised that implementation of Common Business Language lagging other elements of D&A Strategy	Inhibitor	Continue to accelerate plans to implement
Translators (local, region, global)	Data Behaviours	Availability of resource recognised pain-point impeding strategy implementation	Inhibitor	Continue to accelerate plans to put resources in place at Region and OpCo level
POC Driven	Business Value	Too much focus in demonstrating feasibility of small-scale analytics use cases	Inhibitor	Shift in approach to selecting use cases with increased emphasis on impact, path to value and actionability.

In addition to amending the criteria for use case selection, the resources of the AME D&A team were also reoriented to accelerating capability building at the OpCo level. The profile of D&A roles in the OpCos was changed to move more towards transformation type activities such as Change Management and Analytics Product Management while further consolidating technical roles such as Data Scientists in the Regional Hub. Other initiatives included executive upskilling programmes, OpCo focused D&A Playbooks, training courses for technical and non-technical employees, Maturity Assessments and engagement initiatives such as newsletters and competitions.

The DVM outputs along with the resultant shifts in the D&A Strategy were reviewed by the Senior Leadership Team of the Region at a strategy off-site during which it was noted that the DVM Analysis was a useful exercise and provided reassurance that the course-corrections being implemented in the D&A Strategy were addressing the real implementation pain points to becoming a data driven organisation. In the second half of the year, progress with the implementation of the strategy was observed by the Head of D&A (lead author) to have picked up again with a renewed vigour and enthusiasm from the implementation teams playing a role.

The objective of this research was to determine if the D&A strategy was landing with the people responsible for implementing it and how implementable it was in its current format. On December 19th, 2023, the Digital & Technology Director for \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\* in the AME Region addressed the year end Global Townhall being broadcast to employees across \*MULTINATIONAL-ANONYMOUS-FOR-REVIEW\*. She proudly exclaimed that for the second year running the AME Region (smallest of the four regions) had successfully generated the most value from Data & Analytics, had been the first to complete Data Prime, (the company wide programme to implement a cloud based analytics platform), and the success had been achieved by realizing the holistic Business Analytics strategy focused on achieving Higher business impact, Faster time to insight and growing Stronger as a data driven collaborative organization with solid data and technology foundations. This level of success could only have been achieved by ensuring that there was no strategy-execution gap in the bringing to life of the D&A strategy.

## References

- Akhtar, P., Frynas, J. G., Mellahi, K., & Ullah, S. (2019). Big Data-Savvy Teams' Skills, Big Data-Driven Actions and Business Performance. *British journal of management*, 30(2), 252-271. <https://doi.org/10.1111/1467-8551.12333>
- Bonchek, M. (2017). Is execution where good strategies go to die. *Harvard Business Review*, 7.
- Bumblauskas, D., Nold, H., Bumblauskas, P., & Igou, A. (2017). Big data analytics: transforming data to action. *Business process management journal*, 23(3), 703-720. <https://doi.org/10.1108/BPMJ-03-2016-0056>
- Campbell, D. T. (1955). The Informant in Quantitative Research. *The American journal of sociology*, 60(4), 339-342. <https://doi.org/10.1086/221565>
- Cao, G., & Duan, Y. (2017). How do top- and bottom-performing companies differ in using business analytics? *Journal of enterprise information management*, 30(6), 874-892. <https://doi.org/10.1108/JEIM-04-2016-0080>
- Davenport, T. H. (2006). COMPETING ON ANALYTICS [Article]. *Harvard Business Review*, 84(1), 98-107.
- Davenport, T. H., & Redman, T. C. (2020). Digital transformation comes down to talent in 4 key areas. *Harvard Business Review*, 21.
- Delen, D., & Ram, S. (2018). Research challenges and opportunities in business analytics. *Journal of Business Analytics*, 1(1), 2-12.
- Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational research methods*, 16(1), 15-31. <https://doi.org/10.1177/1094428112452151>
- Gust, G., Flath, C. M., Brandt, T., Ströhle, P., & Neumann, D. (2017). How a Traditional Company Seeded New Analytics Capabilities. *MIS quarterly executive*, 16(3), 215-230.
- Hindle, G., Kunc, M., Mortensen, M., Oztekin, A., & Vidgen, R. (2020). Business analytics: Defining the field and identifying a research agenda. *European journal of operational research*, 281(3), 483-490. <https://doi.org/10.1016/j.ejor.2019.10.001>
- Kenny, G. (2019). 5 Simple Rules for Strategy Execution. In: Harvard Business Review.
- Klee, S., Janson, A., & Leimeister, J. M. (2021). How Data Analytics Competencies Can Foster Business Value- A Systematic Review and Way Forward. *Information systems management*, 38(3), 200-217. <https://doi.org/10.1080/10580530.2021.1894515>
- Laney, D. B. (2018). *Infomatics*. Gartner Inc.
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT sloan management review*, 52(2), 21.
- Markus, M. L., Rowe, F., Nantes, U., & Bentley, U. (2021). Guest Editorial: Theories of Digital Transformation: A Progress Report. *Journal of the Association for Information Systems*, 22(2), 273-280. <https://doi.org/10.17705/1jais.00661>
- Mikalef, P., Pappas, I. O., Krogstie, J., & Pavlou, P. A. (2020). Big data and business analytics: A research agenda for realizing business value. *Information & management*, 57(1), 103237. <https://doi.org/10.1016/j.im.2019.103237>
- NewVantagePartners. (2023). *Data and Analytics Leadership Annual Executive Survey 2023 - Executive summary of findings*. [https://www.newvantage.com/files/ugd/e5361a\\_247885043758499ba090f7a5f510cf7c.pdf](https://www.newvantage.com/files/ugd/e5361a_247885043758499ba090f7a5f510cf7c.pdf)



- Redman, T. C. (2008). *Data Driven : profiting from your most important business asset*. Harvard Business Press.
- Sammon, D., & Nagle, T. (2017, 2017). The Data Value Map: A framework for developing shared understanding on data initiatives.
- Smith, T., Stiller, B., Guszczka, J., & Davenport, T. (2019). Analytics and AI-driven enterprises thrive in the Age of With. *Deloitte Insights*.
- Stubbs, E. (2013). *Delivering business analytics: practical guidelines for best practice*. Wiley. <https://doi.org/10.1002/9781119203674>
- Wiita, N., & Leonard, O. (2017). How the most successful teams bridge the strategy execution gap. *Harvard Business Review*, 23.

UKAIS 2024 Proceedings  
KEYWORD INDEX


---


Index of author-provided keywords.



Abstract Data	<a href="#"><u>Understanding Artificial Intelligence For Data With High Level Of Abstraction: Beyond Pixel Importance</u></a>
Academic Workload	<a href="#"><u>An Exploration of Contemporary Factors that Influence Adoption of Instant Messaging by Academics in Higher Education</u></a>
Accountability	<a href="#"><u>Accountability and legitimacy in implementation of electronic patient records in hospitals: The Jamaican case</u></a>
action learning	<a href="#"><u>MOVING ACTION LEARNING SETS ONLINE: REFLECTING ON PRIVACY, INTERSECTIONALITY AND GROUP FAILURE</u></a>
affective experience	<a href="#"><u>On the Temporal Experiences of Remote Workers</u></a>
affective norms	<a href="#"><u>Context-Aware Emotional State Inference: Extending the Affective Norms of Valence, Arousal, and Dominance for N-grams</u></a>
affordances	<a href="#"><u>Digital Platforms and Value Creation - conceptualising the link</u></a>
Agile Development	<a href="#"><u>Leveraging Generative AI in Information Systems Development</u></a>
AI	<a href="#"><u>Generative AI as a Tool for Thematic Analysis: An Exploratory Study with ChatGPT</u></a>
AI Divide	<a href="#"><u>Conceptualising the Artificial Intelligence Divide: A Systematic Literature Review and Research Agenda</u></a>
algorithm aversion	<a href="#"><u>The Delegation Paradox: A Study on How Uncertainty Impacts Experts' Willingness to Delegate Decisions to Algorithmic Solutions</u></a>
Algorithmic Fairness	<a href="#"><u>Operationalizing Algorithmic Fairness: Empirical Study and Sociotechnical Framework Proposal</u></a>

Analytics	<a href="#"><u>Project Materiality and Integrated Analytics in Sustainability Augmented Portfolio Management</u></a>
Apparel	<a href="#"><u>Apparel Matters? A New First-Impression Static Appearance Dataset</u></a>
Artificial Intelligence	<a href="#"><u>Generative AI as a Tool for Thematic Analysis: An Exploratory Study with ChatGPT</u></a>
	<a href="#"><u>Conceptualising the Artificial Intelligence Divide: A Systematic Literature Review and Research Agenda</u></a>
	<a href="#"><u>Artificial Intelligence and Blockchain Technology</u></a>
Artificial Intelligence Ethics	<a href="#"><u>Uncovering the Landscape of Ethical Artificial Intelligence</u></a>
Augmented Reality	<a href="#"><u>The Role of Augmented Reality in Enhancing Sustainable Decision-Making in the Fashion Industry</u></a>
	<a href="#"><u>Sustainability of business lies in its ability to be resilient; AR technology resilience in frontline employee, the case of retail sector</u></a>
Autism Spectrum Disorder	<a href="#"><u>Online Health Communities for Parents of Children with Autism Spectrum Disorder</u></a>
<b>B</b>	
barriers	<a href="#"><u>Factors influencing the career decisions of women software entrepreneurs: perspectives from India and Ireland</u></a>
Bibliometric Analysis	<a href="#"><u>Mapping and Visualising the Digital Economy in The Context of Developing Countries: A Bibliometric Analysis</u></a>
Blended learning	<a href="#"><u>The Technological Challenges of Blended Learning in Higher Education: A case study of Blackboard</u></a>
Blockchain	<a href="#"><u>The Role of Middle Managers in Steering Digital Transformation: A Dynamic Capabilities Perspective</u></a>
Blockchain Technology	<a href="#"><u>Mind the gap: using threshold concept theory to advance blockchain education</u></a>
	<a href="#"><u>Artificial Intelligence and Blockchain Technology</u></a>
boundary objects	<a href="#"><u>Virtual Representations as Boundary Objects: a Case of Building Information Modeling</u></a>

Building Information Modelling	<a href="#"><u>Virtual Representations as Boundary Objects: a Case of Building Information Modeling</u></a>
Business Analytics	<a href="#"><u>Unpacking the 'Black Boxes' of Analytics: A Data Value Map Analysis</u></a>
business models	<a href="#"><u>Business models for digital sustainability: The case of Fintech startups</u></a>
	
Capability	<a href="#"><u>Unpacking the 'Black Boxes' of Analytics: A Data Value Map Analysis</u></a>
carbon footprint	<a href="#"><u>Building Sustainable Software - Measuring the Effects of Multithreading on the Carbon Footprint of Software Systems</u></a>
Care	<a href="#"><u>The impact of the digitalisation of care on older, unpaid carers</u></a>
Carers	<a href="#"><u>The impact of the digitalisation of care on older, unpaid carers</u></a>
case method	<a href="#"><u>Developing the foundations for an inclusive IS education – a case method approach for promoting female digital leaders</u></a>
challenges	<a href="#"><u>Digital Ethics: Resolving “Wicked” Problems and Dilemmas</u></a>
ChatGPT	<a href="#"><u>Generative AI as a Tool for Thematic Analysis: An Exploratory Study with ChatGPT</u></a>
ChatGPT4	<a href="#"><u>Generative AI as a Tool for Thematic Analysis: An Exploratory Study with ChatGPT</u></a>
Citizen Science	<a href="#"><u>Data-driven Sustainable Cities in the UK: The Case of Public Space and Social Cohesion</u></a>
Clinical Decision Making	<a href="#"><u>Temporal Aspects in Cancer Clinical Decision-Making Through Machine Learning</u></a>
Clinical decision support systems	<a href="#"><u>A Conceptual Framework and Design Principles for Decision Support in Clinical Practice: Managing Knowledge and Uncertainty</u></a>
Co-creation	<a href="#"><u>Responsible digital: co-creating safe, secure and wise digital interventions with vulnerable groups</u></a>

collaboration	<a href="#"><u>From Connections to Classrooms: Utilizing Social Media platforms for Enhanced Engagement: The case of LinkedIn.</u></a>
communication media	<a href="#"><u>Media Choice for Multi-motive Communication: Impersonal and Interpersonal Motives in Social Worker-Client Relationships</u></a>
conceptual modelling	<a href="#"><u>A Conceptual Framework and Design Principles for Decision Support in Clinical Practice: Managing Knowledge and Uncertainty</u></a>
Confidentiality	<a href="#"><u>Sharing of personal health information for secondary use: A scoping review from the perspective of trust</u></a>
constraints	<a href="#"><u>Digital Platforms and Value Creation - conceptualising the link</u></a>
Consumer Behaviour	<a href="#"><u>The Role of Augmented Reality in Enhancing Sustainable Decision-Making in the Fashion Industry</u></a>
Context-aware	
Covid	<a href="#"><u>The Digital Divide and Learning Loss in India During COVID: The Role of Home Internet Access</u></a>
Critical Infrastructures	<a href="#"><u>Smart maintenance at offshore wind farms: A digital System of Systems approach</u></a>
cross-boundary collaboration	<a href="#"><u>Virtual Representations as Boundary Objects: a Case of Building Information Modeling</u></a>
crowdfunding	<a href="#"><u>Self-Determination Theory in Crowdfunding: Data from Vietnam</u></a>
Crowdsourcing Information	<a href="#"><u>Data-driven Sustainable Cities in the UK: The Case of Public Space and Social Cohesion</u></a>
Curriculum Design	<a href="#"><u>Becoming Analytical Champions: A Simple Hermeneutics-inspired Learning Analytics Model for Modules</u></a>
 Data Markets	<a href="#"><u>Artificial Intelligence and Blockchain Technology</u></a>
data-driven approach	<a href="#"><u>Driving Student Success through a Data-Driven Approach in Higher Education</u></a>

Data-driven sustainable cities	<a href="#"><u><i>Data-driven Sustainable Cities in the UK: The Case of Public Space and Social Cohesion</i></u></a>
Data-Sharing	<a href="#"><u><i>Smart maintenance at offshore wind farms: A digital System of Systems approach</i></u></a>
Dataset	<a href="#"><u><i>Apparel Matters? A New First-Impression Static Appearance Dataset</i></u></a>
decision delegation	<a href="#"><u><i>The Delegation Paradox: A Study on How Uncertainty Impacts Experts' Willingness to Delegate Decisions to Algorithmic Solutions</i></u></a>
decision theory	<a href="#"><u><i>A Conceptual Framework and Design Principles for Decision Support in Clinical Practice: Managing Knowledge and Uncertainty</i></u></a>
Design Science	<a href="#"><u><i>Design and Development of a Digital Transformation Canvas for SMEs in Developing Countries: a case study of Oman</i></u></a>
design science research	<a href="#"><u><i>A Conceptual Framework and Design Principles for Decision Support in Clinical Practice: Managing Knowledge and Uncertainty</i></u></a>
Design Thinking	<a href="#"><u><i>Leveraging Generative AI in Information Systems Development</i></u></a>
Developing Countries	<a href="#"><u><i>Mapping and Visualising the Digital Economy in The Context of Developing Countries: A Bibliometric Analysis</i></u></a>
Differences-in-difference	<a href="#"><u><i>The Digital Divide and Learning Loss in India During COVID: The Role of Home Internet Access</i></u></a>
digital adoption	<a href="#"><u><i>Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs</i></u></a>
digital artefacts	<a href="#"><u><i>Virtual Representations as Boundary Objects: a Case of Building Information Modeling</i></u></a>
Digital Divide	<a href="#"><u><i>Conceptualising the Artificial Intelligence Divide: A Systematic Literature Review and Research Agenda</i></u></a>

	<u><i>The Digital Divide and Learning Loss in India During COVID: The Role of Home Internet Access</i></u>
Digital Economy	<u><i>Mapping and Visualising the Digital Economy in The Context of Developing Countries: A Bibliometric Analysis</i></u>
digital ethics	<u><i>Digital Ethics: Resolving “Wicked” Problems and Dilemmas</i></u>
digital inclusion	<u><i>Digital inclusion network building: a network weaving analysis</i></u>
digital innovation	<u><i>Design and Development of a Digital Transformation Canvas for SMEs in Developing Countries: a case study of Oman</i></u>
	<u><i>A soft ecosystems methodology of digital innovation through a case study of the insurance industry’s response to connected cars</i></u>
	<u><i>Responsible Digital Innovation with Micro-Businesses: A Resource Orchestration Perspective</i></u>
Digital Intervention	<u><i>Responsible digital: co-creating safe, secure and wise digital interventions with vulnerable groups</i></u>
digital leaders	<u><i>Developing the foundations for an inclusive IS education – a case method approach for promoting female digital leaders</i></u>
digital platform	<u><i>Digital Platforms and Value Creation - conceptualising the link</i></u>
digital poverty	<u><i>Digital inclusion network building: a network weaving analysis</i></u>
digital sustainability	<u><i>Business models for digital sustainability: The case of Fintech startups</i></u>
Digital Technologies	<u><i>Internalization of Digital Technologies: Adapting to Organizational Inertia</i></u>
Digital Transformation	<u><i>The Underlying Practices of Digital Transformation Leadership: Theorising the Practitioner Voice</i></u>
	<u><i>Internalization of Digital Technologies: Adapting to Organizational Inertia</i></u>
	<u><i>Mapping and Visualising the Digital Economy in The Context of Developing Countries: A Bibliometric Analysis</i></u>

	<u><i>The Role of Middle Managers in Steering Digital Transformation: A Dynamic Capabilities Perspective</i></u>
	<u><i>Digital transformation in the public sector: Critical factors of Digital Justice Portal</i></u>
Digital Transformation Canvas	<u><i>Design and Development of a Digital Transformation Canvas for SMEs in Developing Countries: a case study of Oman</i></u>
Digital trust	<u><i>Digital Ethics: Resolving “Wicked” Problems and Dilemmas</i></u>
Digitalisation	<u><i>Sourcing Trends for the Digitalisation of the Health Sector: English National Health System</i></u> <u><i>The impact of the digitalisation of care on older, unpaid carers</i></u>
Digitalised Workplace	<u><i>An Exploration of Contemporary Factors that Influence Adoption of Instant Messaging by Academics in Higher Education</i></u>
Digitalization	<u><i>Smart maintenance at offshore wind farms: A digital System of Systems approach</i></u>
dilemmas	<u><i>Digital Ethics: Resolving “Wicked” Problems and Dilemmas</i></u>
disengagement	<u><i>Understanding Consumers’ Reactance of Technology-Enabled Personalization: Perspectives on Stress and Fatigue</i></u>
disruption	<u><i>Sustainability of business lies in its ability to be resilient; AR technology resilience in frontline employee, the case of retail sector</i></u>
Dynamic Capabilities	<u><i>The Role of Middle Managers in Steering Digital Transformation: A Dynamic Capabilities Perspective</i></u>
 E-justice system	<u><i>Digital transformation in the public sector: Critical factors of Digital Justice Portal</i></u>
ecosystem-as-is	<u><i>A soft ecosystems methodology of digital innovation through a case study of the insurance industry’s response to connected cars</i></u>



ecosystem-as-was	<u><i>A soft ecosystems methodology of digital innovation through a case study of the insurance industry's response to connected cars</i></u>
ecosystem-to-be	<u><i>A soft ecosystems methodology of digital innovation through a case study of the insurance industry's response to connected cars</i></u>
Ecosystems	<u><i>A soft ecosystems methodology of digital innovation through a case study of the insurance industry's response to connected cars</i></u>
Education	<u><i>The Digital Divide and Learning Loss in India During COVID: The Role of Home Internet Access</i></u>
emerging markets	<u><i>Self-Determination Theory in Crowdfunding: Data from Vietnam</i></u>
emerging technologies	<u><i>A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach (TEAeM)</i></u>
emotion	<u><i>Context-Aware Emotional State Inference: Extending the Affective Norms of Valence, Arousal, and Dominance for N-grams</i></u>
Emotion regulation	<u><i>Emotional regulation for improved student learning: can smartwatches be utilized to navigate an increasingly digital student context?</i></u>
Emotional Support	<u><i>Online Health Communities for Parents of Children with Autism Spectrum Disorder</i></u>
Emotions	<u><i>Emotional regulation for improved student learning: can smartwatches be utilized to navigate an increasingly digital student context?</i></u>
ethics	<u><i>A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach (TEAeM)</i></u>
ethics issues	<u><i>A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach (TEAeM)</i></u>
Evidence-Based Nudges	<u><i>Becoming Analytical Champions: A Simple Hermeneutics-inspired Learning Analytics Model for Modules</i></u>
Explainable AI	<u><i>Uncovering the Landscape of Ethical Artificial Intelligence</i></u>

Explainable Artificial Intelligence	<a href="#"><u>Understanding Artificial Intelligence For Data With High Level Of Abstraction: Beyond Pixel Importance</u></a>
Exploratory Research	<a href="#"><u>Teachers' Perceptions of Students' Use of Generative AI in Summative Assessments at Higher Education Institutions: An Exploratory Study</u></a>
extrinsic motivation	<a href="#"><u>Self-Determination Theory in Crowdfunding: Data from Vietnam</u></a>
<b>F</b>	
Factor Analysis of Information Risk	<a href="#"><u>Cybersecurity Investments Metrics using FAIR-ROSI</u></a>
FAIR-ROSI model	<a href="#"><u>Cybersecurity Investments Metrics using FAIR-ROSI</u></a>
Fashion Industry	<a href="#"><u>The Role of Augmented Reality in Enhancing Sustainable Decision-Making in the Fashion Industry</u></a>
Feedback	<a href="#"><u>Becoming Analytical Champions: A Simple Hermeneutics-inspired Learning Analytics Model for Modules</u></a>
female IT professionals	<a href="#"><u>Developing the foundations for an inclusive IS education – a case method approach for promoting female digital leaders</u></a>
female IT students	<a href="#"><u>Developing the foundations for an inclusive IS education – a case method approach for promoting female digital leaders</u></a>
Fintech	<a href="#"><u>Business models for digital sustainability: The case of Fintech startups</u></a>
First impression	<a href="#"><u>Apparel Matters? A New First-Impression Static Appearance Dataset</u></a>
Flow-oriented approach	<a href="#"><u>Sourcing Trends for the Digitalisation of the Health Sector: English National Health System</u></a>
Foundation Models	<a href="#"><u>Leveraging Generative AI in Information Systems Development</u></a>
frameworks	<a href="#"><u>A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach (TEAeM)</u></a>
frontline employee	<a href="#"><u>Sustainability of business lies in its ability to be resilient; AR technology resilience in</u></a>

frontline employee, the case of retail sector

**G**

Gamification

Metaverse, as a virtual learning platform: the application of gamification to enhance user engagement

GenAI

An Exploration of Contemporary Factors that Influence Adoption of Instant Messaging by Academics in Higher Education

gender

Factors influencing the career decisions of women software entrepreneurs: perspectives from India and Ireland

Generative AI

Leveraging Generative AI in Information Systems Development

Teachers' Perceptions of Students' Use of Generative AI in Summative Assessments at Higher Education Institutions: An Exploratory Study

Generative Artificial Intelligence

Generative AI in the Workplace: A Dive into Employees' Acceptance in China

Governance

Smart maintenance at offshore wind farms: A digital System of Systems approach

Grounded Approach

The Underlying Practices of Digital Transformation Leadership: Theorising the Practitioner Voice

**H**

Health Information Systems

Sharing of personal health information for secondary use: A scoping review from the perspective of trust

Health sector

Sourcing Trends for the Digitalisation of the Health Sector: English National Health System

Healthcare

Understanding Artificial Intelligence For Data With High Level Of Abstraction: Beyond Pixel Importance



Hermeneutics


Becoming Analytical Champions: A Simple Hermeneutics-inspired Learning Analytics Model for Modules

Higher Education

Teaching and Learning through the pandemic; the effects of remote work on women academics

	<p><u><i>Teachers' Perceptions of Students' Use of Generative AI in Summative Assessments at Higher Education Institutions: An Exploratory Study</i></u></p> <p><u><i>Metaverse, as a virtual learning platform: the application of gamification to enhance user engagement</i></u></p> <p><u><i>Impact of WhatsApp Groups on Teaching and Learning in Higher Education: A Review</i></u></p> <p><u><i>Driving Student Success through a Data-Driven Approach in Higher Education</i></u></p>
human-algorithm interactions	<u><i>The Delegation Paradox: A Study on How Uncertainty Impacts Experts' Willingness to Delegate Decisions to Algorithmic Solutions</i></u>
Human-centered IA	<u><i>Uncovering the Landscape of Ethical Artificial Intelligence</i></u>
<b>I</b>	
ICT	<u><i>Impact of WhatsApp Groups on Teaching and Learning in Higher Education: A Review</i></u>
imagining	<u><i>A soft ecosystems methodology of digital innovation through a case study of the insurance industry's response to connected cars</i></u>
Inclusion	<u><i>Fostering inclusion in open strategy: three IT-enabled mechanisms</i></u>
inclusive pedagogy	<u><i>Developing the foundations for an inclusive IS education – a case method approach for promoting female digital leaders</i></u>
India	<u><i>Factors influencing the career decisions of women software entrepreneurs: perspectives from India and Ireland</i></u>
Inertia	<u><i>Internalization of Digital Technologies: Adapting to Organizational Inertia</i></u>
Information infrastructure sourcing	<u><i>Sourcing Trends for the Digitalisation of the Health Sector: English National Health System</i></u>
Information Needs	<u><i>Online Health Communities for Parents of Children with Autism Spectrum Disorder</i></u>

Information Sharing	<a href="#"><u>Sharing of personal health information for secondary use: A scoping review from the perspective of trust</u></a>
Instant Messaging	<a href="#"><u>An Exploration of Contemporary Factors that Influence Adoption of Instant Messaging by Academics in Higher Education</u></a>
Inter-organisational trust interpersonal relationship	<a href="#"><u>Media Choice for Multi-motive Communication: Impersonal and Interpersonal Motives in Social Worker-Client Relationships</u></a>
intersectionality	<a href="#"><u>MOVING ACTION LEARNING SETS ONLINE: REFLECTING ON PRIVACY, INTERSECTIONALITY AND GROUP FAILURE</u></a>
intrinsic motivation	<a href="#"><u>Self-Determination Theory in Crowdfunding: Data from Vietnam</u></a>
Ireland	<a href="#"><u>Factors influencing the career decisions of women software entrepreneurs: perspectives from India and Ireland</u></a>
IS education	<a href="#"><u>Developing the foundations for an inclusive IS education – a case method approach for promoting female digital leaders</u></a>
IT Professionals	<a href="#"><u>Digital Ethics: Resolving “Wicked” Problems and Dilemmas</u></a>
IT strategy	<a href="#"><u>Digital transformation in the public sector: Critical factors of Digital Justice Portal</u></a>
IT-enabled organisational change	
 Key Informant	<a href="#"><u>The Underlying Practices of Digital Transformation Leadership: Theorising the Practitioner Voice</u></a>
knowledge work	<a href="#"><u>On the Temporal Experiences of Remote Workers</u></a>
 Leadership	<a href="#"><u>The Underlying Practices of Digital Transformation Leadership: Theorising the Practitioner Voice</u></a>


learning	<u><i>Teaching and Learning through the pandemic; the effects of remote work on women academics</i></u> <u><i>Emotional regulation for improved student learning: can smartwatches be utilized to navigate an increasingly digital student context?</i></u>
Learning Analytics	<u><i>Becoming Analytical Champions: A Simple Hermeneutics-inspired Learning Analytics Model for Modules</i></u>
Learning Barriers	<u><i>Mind the gap: using threshold concept theory to advance blockchain education</i></u>
Learning loss	<u><i>The Digital Divide and Learning Loss in India During COVID: The Role of Home Internet Access</i></u>
Legitimacy	<u><i>Accountability and legitimacy in implementation of electronic patient records in hospitals: The Jamaican case</i></u>
LinkedIn	<u><i>From Connections to Classrooms: Utilizing Social Media platforms for Enhanced Engagement: The case of LinkedIn.</i></u>
live sales	<u><i>The Power of Streamers' Speech: A Signalling Approach to Live Streaming Commerce</i></u>
live-streaming commerce	<u><i>The Power of Streamers' Speech: A Signalling Approach to Live Streaming Commerce</i></u>
LLMs	<u><i>An Exploration of Contemporary Factors that Influence Adoption of Instant Messaging by Academics in Higher Education</i></u>
 Machine autonomy	<u><i>Uncovering the Landscape of Ethical Artificial Intelligence</i></u>
Machine Learning	<u><i>Temporal Aspects in Cancer Clinical Decision-Making Through Machine Learning</i></u> <u><i>Uncovering the Landscape of Ethical Artificial Intelligence</i></u>
machine learning algorithm	<u><i>Driving Student Success through a Data-Driven Approach in Higher Education</i></u>
Matrix factorisation	

mediated communication	<a href="#"><u>Virtual Representations as Boundary Objects: a Case of Building Information Modeling</u></a>
Metaverse	<a href="#"><u>Metaverse, as a virtual learning platform: the application of gamification to enhance user engagement</u></a>
method	<a href="#"><u>A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach (TEAeM)</u></a>
Methodology in ISD	<a href="#"><u>Leveraging Generative AI in Information Systems Development</u></a>
micro-businesses	<a href="#"><u>Responsible Digital Innovation with Micro-Businesses: A Resource Orchestration Perspective</u></a>
Middle Manager	<a href="#"><u>The Role of Middle Managers in Steering Digital Transformation: A Dynamic Capabilities Perspective</u></a>
Multi-Feedback	
multi-motive communication	<a href="#"><u>Media Choice for Multi-motive Communication: Impersonal and Interpersonal Motives in Social Worker-Client Relationships</u></a>
multithreading	<a href="#"><u>Building Sustainable Software - Measuring the Effects of Multithreading on the Carbon Footprint of Software Systems</u></a>

## N

NDCG	
Netnography	<a href="#"><u>Online Health Communities for Parents of Children with Autism Spectrum Disorder</u></a>
network building	<a href="#"><u>Digital inclusion network building: a network weaving analysis</u></a>
network weaving	<a href="#"><u>Digital inclusion network building: a network weaving analysis</u></a>
NeuroIS	<a href="#"><u>Context-Aware Emotional State Inference: Extending the Affective Norms of Valence, Arousal, and Dominance for N-grams</u></a>
NFTs Community	<a href="#"><u>What are the Drivers of Non-Fungible Tokens (NFTs) Community Engagement?</u></a>
Non-Fungible Tokens (NFTs)	<a href="#"><u>What are the Drivers of Non-Fungible Tokens (NFTs) Community Engagement?</u></a>

## O

Older	<a href="#"><u><i>The impact of the digitalisation of care on older, unpaid carers</i></u></a>
Oncology	<a href="#"><u><i>Temporal Aspects in Cancer Clinical Decision-Making Through Machine Learning</i></u></a>
Online Health Community	<a href="#"><u><i>Online Health Communities for Parents of Children with Autism Spectrum Disorder</i></u></a>
Online Learning and Teaching	<a href="#"><u><i>Metaverse, as a virtual learning platform: the application of gamification to enhance user engagement</i></u></a>
Online learning challenges	<a href="#"><u><i>The Technological Challenges of Blended Learning in Higher Education: A case study of Blackboard</i></u></a>
online learning processes	<a href="#"><u><i>MOVING ACTION LEARNING SETS ONLINE: REFLECTING ON PRIVACY, INTERSECTIONALITY AND GROUP FAILURE</i></u></a>
online studies	<a href="#"><u><i>From Connections to Classrooms: Utilizing Social Media platforms for Enhanced Engagement: The case of LinkedIn.</i></u></a>
Online Technology	<a href="#"><u><i>Impact of WhatsApp Groups on Teaching and Learning in Higher Education: A Review</i></u></a>
Open Strategy	<a href="#"><u><i>Fostering inclusion in open strategy: three IT-enabled mechanisms</i></u></a>
Operationalized RAI	<a href="#"><u><i>Responsible AI Principles: Findings from an Empirical Study on Practitioners' Perceptions</i></u></a>
Organisational Transformation	<a href="#"><u><i>Design and Development of a Digital Transformation Canvas for SMEs in Developing Countries: a case study of Oman</i></u></a>
Overtime	<a href="#"><u><i>Generative AI in the Workplace: A Dive into Employees' Acceptance in China</i></u></a>
 pandemic	<a href="#"><u><i>Teaching and Learning through the pandemic; the effects of remote work on women academics</i></u></a>
Perceptions	<a href="#"><u><i>Responsible AI Principles: Findings from an Empirical Study on Practitioners' Perceptions</i></u></a>



platform value	<a href="#"><u>Digital Platforms and Value Creation - conceptualising the link</u></a>
post purchase	<a href="#"><u>The Power of Streamers' Speech: A Signalling Approach to Live Streaming Commerce</u></a>
Practices	<a href="#"><u>The Underlying Practices of Digital Transformation Leadership: Theorising the Practitioner Voice</u></a>
Prediction Markets	<a href="#"><u>Artificial Intelligence and Blockchain Technology</u></a>
Principles	<a href="#"><u>Responsible AI Principles: Findings from an Empirical Study on Practitioners' Perceptions</u></a>
Privacy	<a href="#"><u>Sharing of personal health information for secondary use: A scoping review from the perspective of trust</u></a>
privacy fatigue	<a href="#"><u>Understanding Consumers' Reactance of Technology-Enabled Personalization: Perspectives on Stress and Fatigue</u></a>
Process	<a href="#"><u>Internalization of Digital Technologies: Adapting to Organizational Inertia</u></a>
Product Return	<a href="#"><u>The Role of Augmented Reality in Enhancing Sustainable Decision-Making in the Fashion Industry</u></a>
productivity	<a href="#"><u>Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs</u></a>
productivity puzzle	<a href="#"><u>Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs</u></a>
Project Materiality	<a href="#"><u>Project Materiality and Integrated Analytics in Sustainability Augmented Portfolio Management</u></a>
Project portfolio management	<a href="#"><u>Project Materiality and Integrated Analytics in Sustainability Augmented Portfolio Management</u></a>
 qualitative	<a href="#"><u>Business models for digital sustainability: The case of Fintech startups</u></a>

qualitative case study

*Responsible Digital Innovation with  
Micro-Businesses: A Resource  
Orchestration Perspective*

*The Role of Middle Managers in  
Steering Digital Transformation: A  
Dynamic Capabilities Perspective*

Quality of life

*Generative AI in the Workplace: A Dive  
into Employees' Acceptance in China*

R

Radiotherapy

*Understanding Artificial Intelligence For  
Data With High Level Of Abstraction:  
Beyond Pixel Importance*

reading

*Context-Aware Emotional State Inference:  
Extending the Affective Norms of Valence,  
Arousal, and Dominance for N-grams*

Recommender systems

Reflective Practice

*Unpacking the 'Black Boxes' of Analytics:  
A Data Value Map Analysis*

remote work

*Teaching and Learning through the  
pandemic; the effects of remote work on  
women academics*

*On the Temporal Experiences of Remote  
Workers*

remote working

*Media Choice for Multi-motive  
Communication: Impersonal and  
Interpersonal Motives in Social Worker-  
Client Relationships*

Research Methods

*Generative AI as a Tool for Thematic  
Analysis: An Exploratory Study with  
ChatGPT*

resilience

*Sustainability of business lies in its ability  
to be resilient; AR technology resilience in  
frontline employee, the case of retail  
sector*

resource orchestration theory

*Responsible Digital Innovation with  
Micro-Businesses: A Resource  
Orchestration Perspective*

Responsible AI

*Responsible AI Principles: Findings from  
an Empirical Study on Practitioners'  
Perceptions*

*Operationalizing Algorithmic Fairness:  
Empirical Study and Sociotechnical  
Framework Proposal*

	<u><i>Teachers' Perceptions of Students' Use of Generative AI in Summative Assessments at Higher Education Institutions: An Exploratory Study</i></u>
Responsible Digital	<u><i>Responsible digital: co-creating safe, secure and wise digital interventions with vulnerable groups</i></u>
responsible innovation	<u><i>Responsible Digital Innovation with Micro-Businesses: A Resource Orchestration Perspective</i></u>
retail	<u><i>The Role of Augmented Reality in Enhancing Sustainable Decision-Making in the Fashion Industry</i></u> <u><i>Sustainability of business lies in its ability to be resilient; AR technology resilience in frontline employee, the case of retail sector</i></u>
Return on Security Investment	<u><i>Cybersecurity Investments Metrics using FAIR-ROSI</i></u>
Risk Assessment	<u><i>Cybersecurity Investments Metrics using FAIR-ROSI</i></u>
	
Safe Secure and Wise Use	<u><i>Responsible digital: co-creating safe, secure and wise digital interventions with vulnerable groups</i></u>
Satisfaction	<u><i>Digital transformation in the public sector: Critical factors of Digital Justice Portal</i></u>
Scoping Literature Review	<u><i>Sharing of personal health information for secondary use: A scoping review from the perspective of trust</i></u>
self-determination theory	<u><i>Self-Determination Theory in Crowdfunding: Data from Vietnam</i></u>
self-efficacy	<u><i>From Connections to Classrooms: Utilizing Social Media platforms for Enhanced Engagement: The case of LinkedIn.</i></u>
self-regulation	<u><i>From Connections to Classrooms: Utilizing Social Media platforms for Enhanced Engagement: The case of LinkedIn.</i></u>
semiotics	<u><i>Virtual Representations as Boundary Objects: a Case of Building Information Modeling</i></u>

Sensemaking	<u><i>Another tool for the toolbox? How strategists are adopting social media for sensemaking</i></u>
Signalling Theory	<u><i>The Power of Streamers' Speech: A Signalling Approach to Live Streaming Commerce</i></u>
Smart contract	
Smart Maintenance	<u><i>Smart maintenance at offshore wind farms: A digital System of Systems approach</i></u>
Smartwatches	<u><i>Emotional regulation for improved student learning: can smartwatches be utilized to navigate an increasingly digital student context?</i></u>
Social Capital Theory	<u><i>What are the Drivers of Non-Fungible Tokens (NFTs) Community Engagement?</i></u>
Social Exchange Theory	<u><i>What are the Drivers of Non-Fungible Tokens (NFTs) Community Engagement?</i></u>
Social media	<u><i>Another tool for the toolbox? How strategists are adopting social media for sensemaking</i></u> <u><i>Impact of WhatsApp Groups on Teaching and Learning in Higher Education: A Review</i></u>
social work	<u><i>Media Choice for Multi-motive Communication: Impersonal and Interpersonal Motives in Social Worker-Client Relationships</i></u>
Socio-Technical	<u><i>Conceptualising the Artificial Intelligence Divide: A Systematic Literature Review and Research Agenda</i></u>
Sociotechnical Systems	<u><i>Operationalizing Algorithmic Fairness: Empirical Study and Sociotechnical Framework Proposal</i></u>
software carbon intensity	<u><i>Building Sustainable Software - Measuring the Effects of Multithreading on the Carbon Footprint of Software Systems</i></u>
software development	<u><i>Building Sustainable Software - Measuring the Effects of Multithreading on the Carbon Footprint of Software Systems</i></u>

Software entrepreneur	<a href="#"><u>Factors influencing the career decisions of women software entrepreneurs: perspectives from India and Ireland</u></a>
speaking charisma	<a href="#"><u>The Power of Streamers' Speech: A Signalling Approach to Live Streaming Commerce</u></a>
Strategizing	<a href="#"><u>Another tool for the toolbox? How strategists are adopting social media for sensemaking</u></a>
Strategy Execution	<a href="#"><u>Unpacking the 'Black Boxes' of Analytics: A Data Value Map Analysis</u></a>
Strategy practice	<a href="#"><u>Another tool for the toolbox? How strategists are adopting social media for sensemaking</u></a>
streamers	<a href="#"><u>The Power of Streamers' Speech: A Signalling Approach to Live Streaming Commerce</u></a>
Structure	<a href="#"><u>Internalization of Digital Technologies: Adapting to Organizational Inertia</u></a>
student engagement	<a href="#"><u>From Connections to Classrooms: Utilizing Social Media platforms for Enhanced Engagement: The case of LinkedIn.</u></a>
student engagment	<a href="#"><u>Driving Student Success through a Data-Driven Approach in Higher Education</u></a>
student support strategy	<a href="#"><u>Driving Student Success through a Data-Driven Approach in Higher Education</u></a>
Success	<a href="#"><u>Digital transformation in the public sector: Critical factors of Digital Justice Portal</u></a>
support	<a href="#"><u>Teaching and Learning through the pandemic; the effects of remote work on women academics</u></a>
Sustainability	<a href="#"><u>Project Materiality and Integrated Analytics in Sustainability Augmented Portfolio Management</u></a> <a href="#"><u>The Role of Augmented Reality in Enhancing Sustainable Decision-Making in the Fashion Industry</u></a>
Sustainability Reporting Standards	<a href="#"><u>Project Materiality and Integrated Analytics in Sustainability Augmented Portfolio Management</u></a>

System of Systems [Smart maintenance at offshore wind farms: A digital System of Systems approach](#)

Systematic Literature Review [Conceptualising the Artificial Intelligence Divide: A Systematic Literature Review and Research Agenda](#)

Systems thinking [A soft ecosystems methodology of digital innovation through a case study of the insurance industry's response to connected cars](#)



tactics [Factors influencing the career decisions of women software entrepreneurs: perspectives from India and Ireland](#)

Teachers' Perception [Teachers' Perceptions of Students' Use of Generative AI in Summative Assessments at Higher Education Institutions: An Exploratory Study](#)

teaching [Teaching and Learning through the pandemic; the effects of remote work on women academics](#)

Teaching and Learning [Impact of WhatsApp Groups on Teaching and Learning in Higher Education: A Review](#)

TEAeM [A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach \(TEAeM\)](#)


techno-stress [Understanding Consumers' Reactance of Technology-Enabled Personalization: Perspectives on Stress and Fatigue](#)

technological affordances [Media Choice for Multi-motive Communication: Impersonal and Interpersonal Motives in Social Worker-Client Relationships](#)

Technology [Fostering inclusion in open strategy: three IT-enabled mechanisms](#)

Technology adoption [Generative AI in the Workplace: A Dive into Employees' Acceptance in China](#)  
[Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs](#)

technology leadership [Exploring the Productivity Puzzle: The Relationship between Technology](#)

	<u><i>Investment and Organisational Productivity in UK SMEs</i></u>
technology-enabled personalization	<u><i>Understanding Consumers' Reactance of Technology-Enabled Personalization: Perspectives on Stress and Fatigue</i></u>
Technostress	<u><i>An Exploration of Contemporary Factors that Influence Adoption of Instant Messaging by Academics in Higher Education</i></u>
Temporal	<u><i>Temporal Aspects in Cancer Clinical Decision-Making Through Machine Learning</i></u>
temporality	<u><i>On the Temporal Experiences of Remote Workers</i></u>
Thematic Analysis	<u><i>Generative AI as a Tool for Thematic Analysis: An Exploratory Study with ChatGPT</i></u>
Third Level Teaching and Learning	<u><i>Mind the gap: using threshold concept theory to advance blockchain education</i></u>
Threshold Concepts	<u><i>Mind the gap: using threshold concept theory to advance blockchain education</i></u>
time	<u><i>Temporal Aspects in Cancer Clinical Decision-Making Through Machine Learning</i></u> <u><i>On the Temporal Experiences of Remote Workers</i></u>
Traditional Organisation	<u><i>Unpacking the 'Black Boxes' of Analytics: A Data Value Map Analysis</i></u>
Transformation	<u><i>Unpacking the 'Black Boxes' of Analytics: A Data Value Map Analysis</i></u>
Transformative Learning	<u><i>Mind the gap: using threshold concept theory to advance blockchain education</i></u>
Transparency	<u><i>Accountability and legitimacy in implementation of electronic patient records in hospitals: The Jamaican case</i></u>
Trust	<u><i>Sharing of personal health information for secondary use: A scoping review from the perspective of trust</i></u>
 UK	<u><i>The impact of the digitalisation of care on older, unpaid carers</i></u>
uncertainty	<u><i>A Conceptual Framework and Design Principles for Decision Support in Clinical</i></u>



Unpaid [Practice: Managing Knowledge and Uncertainty](#)  
[The impact of the digitalisation of care on older, unpaid carers](#)

User Engagement [What are the Drivers of Non-Fungible Tokens \(NFTs\) Community Engagement?](#)

User experience [The Technological Challenges of Blended Learning in Higher Education: A case study of Blackboard](#)



valence-arousal-dominance [Context-Aware Emotional State Inference: Extending the Affective Norms of Valence, Arousal, and Dominance for N-grams](#)

value capture [Business models for digital sustainability: The case of Fintech startups](#)

value creation [Business models for digital sustainability: The case of Fintech startups](#)

value delivery [Business models for digital sustainability: The case of Fintech startups](#)

Virtual Reality [Metaverse, as a virtual learning platform: the application of gamification to enhance user engagement](#)

virtual representation [Virtual Representations as Boundary Objects: a Case of Building Information Modeling](#)

Visualisation [Mapping and Visualising the Digital Economy in The Context of Developing Countries: A Bibliometric Analysis](#)

Volunteered Geographic Information [Data-driven Sustainable Cities in the UK: The Case of Public Space and Social Cohesion](#)

VOSviewer [Mapping and Visualising the Digital Economy in The Context of Developing Countries: A Bibliometric Analysis](#)

Vulnerability [Responsible digital: co-creating safe, secure and wise digital interventions with vulnerable groups](#)



WhatsApp Group [Impact of WhatsApp Groups on Teaching and Learning in Higher Education: A Review](#)



wicked problem	<u><i>Digital Ethics: Resolving “Wicked” Problems and Dilemmas</i></u>
women	<u><i>Factors influencing the career decisions of women software entrepreneurs: perspectives from India and Ireland</i></u>
women academics	<u><i>Teaching and Learning through the pandemic; the effects of remote work on women academics</i></u>
word embedding	<u><i>Context-Aware Emotional State Inference: Extending the Affective Norms of Valence, Arousal, and Dominance for N-grams</i></u>

# UKAIS 2024 Proceedings

## AUTHOR INDEX

---

### A

- [Abrham, Itoro](#) *[Uncovering the Landscape of Ethical Artificial Intelligence](#)*
- [Acikgoz, Fulya](#) *[What are the Drivers of Non-Fungible Tokens \(NFTs\) Community Engagement?](#)*
- [Akbarighatar, Pouria](#) *[Responsible AI Principles: Findings from an Empirical Study on Practitioners' Perceptions](#)*
- [Alhammad, Muna M.](#) *[Teachers' Perceptions of Students' Use of Generative AI in Summative Assessments at Higher Education Institutions: An Exploratory Study](#)*
- [Alhassan, Ibrahim](#) *[The Underlying Practices of Digital Transformation Leadership: Theorising the Practitioner Voice](#)*
- [Ali, Mohammed](#) *[Conceptualising the Artificial Intelligence Divide: A Systematic Literature Review and Research Agenda](#)*  
*[Mapping and Visualising the Digital Economy in The Context of Developing Countries: A Bibliometric Analysis](#)*
- [Alizamani, Amirreza](#) *[Metaverse, as a virtual learning platform: the application of gamification to enhance user engagement](#)*
- [Angelopoulos, Spyros](#) *[The Delegation Paradox: A Study on How Uncertainty Impacts Experts' Willingness to Delegate Decisions to Algorithmic Solutions](#)*  
*[Internalization of Digital Technologies: Adapting to Organizational Inertia](#)*
- [Awais Shakir, Muhammad](#) *[The Role of Augmented Reality in Enhancing Sustainable Decision-Making in the Fashion Industry](#)*
- [Aylward, Bethany](#) *[Digital inclusion network building: a network weaving analysis](#)*

### B

- [Bahl, Aadya](#) *[Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs](#)*

<u>Baptista, Joao</u>	<u><i>Fostering inclusion in open strategy: three IT-enabled mechanisms</i></u>
<u>Bhalla, Nitika</u>	<u><i>A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach (TEAeM)</i></u>
<u>Bhatti, Zeeshan</u>	<u><i>The Role of Augmented Reality in Enhancing Sustainable Decision-Making in the Fashion Industry</i></u>
<u>Black, Rachael</u>	<u><i>The impact of the digitalisation of care on older, unpaid carers</i></u>
<u>Brooks, Laurence</u>	<u><i>A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach (TEAeM)</i></u>
<u>Brown-Houston, Mildred</u>	<u><i>Accountability and legitimacy in implementation of electronic patient records in hospitals: The Jamaican case</i></u>
<u>Buckley, Patrick</u>	<u><i>Artificial Intelligence and Blockchain Technology</i></u>
<u>Budu, Joseph</u>	<u><i>Digital Platforms and Value Creation - conceptualising the link</i></u>
<u>Busalim, Abdelsalam</u>	<u><i>What are the Drivers of Non-Fungible Tokens (NFTs) Community Engagement?</i></u>
	
<u>Cannizzarro, Sara</u>	<u><i>A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach (TEAeM)</i></u>
<u>Chang, Rayen Jui-Yen</u>	<u><i>Responsible Digital Innovation with Micro-Businesses: A Resource Orchestration Perspective</i></u>
<u>Chen, Po-Wen</u>	<u><i>The Role of Middle Managers in Steering Digital Transformation: A Dynamic Capabilities Perspective</i></u>
<u>Chen, Tzuyu</u>	<u><i>Understanding Consumers' Reactance of Technology-Enabled Personalization: Perspectives on Stress and Fatigue</i></u>
<u>Chen, Yun</u>	<u><i>Driving Student Success through a Data-Driven Approach in Higher Education</i></u> <u><i>Design and Development of a Digital Transformation Canvas for SMEs in Developing Countries: a case study of Oman</i></u>
<u>Cheung, Johnson Chun-Sing</u>	<u><i>Media Choice for Multi-motive Communication: Impersonal and</i></u>

- Chopde, Sumant *Interpersonal Motives in Social Worker-Client Relationships*  
*A Novel Multi-Feedback Hybrid Recommender System Based On Matrix Factorisation*
- Choudhary, Utkarsh *The Digital Divide and Learning Loss in India During COVID: The Role of Home Internet Access*
- Clohessy, Trevor *Mind the gap: using threshold concept theory to advance blockchain education*
- Conboy, Kieran *Temporal Aspects in Cancer Clinical Decision-Making Through Machine Learning*

**D**

- Das, Jishnu *A Conceptual Framework and Design Principles for Decision Support in Clinical Practice: Managing Knowledge and Uncertainty*
- Daud, Mohamed *The Technological Challenges of Blended Learning in Higher Education: A case study of Blackboard*
- Day, Nicholas *An Exploration of Contemporary Factors that Influence Adoption of Instant Messaging by Academics in Higher Education*
- Ding, Chuying *Internalization of Digital Technologies: Adapting to Organizational Inertia*
- Dixit, Gaurav *A Novel Multi-Feedback Hybrid Recommender System Based On Matrix Factorisation*
- Do, Minh Ngoc *From Connections to Classrooms: Utilizing Social Media platforms for Enhanced Engagement: The case of LinkedIn.*
- Dzvapatsva, Godwin Pedzisai *Impact of WhatsApp Groups on Teaching and Learning in Higher Education: A Review*

**E**



- Engesmo, Jostein *Emotional regulation for improved student learning: can smartwatches be utilized to navigate an increasingly digital student context?*

**F**

- Fagbamigbe, Omon *The Role of Augmented Reality in Enhancing Sustainable Decision-Making in the Fashion Industry*

- Fang, Linhao *Virtual Representations as Boundary Objects: a Case of Building Information Modeling*  
*Media Choice for Multi-motive*  
*Communication: Impersonal and*  
*Interpersonal Motives in Social Worker-Client Relationships*
- Fei, Chiaodeng *Understanding Consumers' Reactance of Technology-Enabled Personalization: Perspectives on Stress and Fatigue*
- Fletcher, Gordon *Mapping and Visualising the Digital Economy in The Context of Developing Countries: A Bibliometric Analysis*
- G**
- Galal-Edeen, Galal *Leveraging Generative AI in Information Systems Development*
- Gante, Stefanie *The Delegation Paradox: A Study on How Uncertainty Impacts Experts' Willingness to Delegate Decisions to Algorithmic Solutions*
- Gheitasy, Ali *The Technological Challenges of Blended Learning in Higher Education: A case study of Blackboard*
- Giaver, Fay *Emotional regulation for improved student learning: can smartwatches be utilized to navigate an increasingly digital student context?*
- Gill-Simmen, Lucy *Developing the foundations for an inclusive IS education – a case method approach for promoting female digital leaders*
- Griffiths, Marie *Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs*  
*Design and Development of a Digital Transformation Canvas for SMEs in Developing Countries: a case study of Oman*
- Griva, Anastasia *Temporal Aspects in Cancer Clinical Decision-Making Through Machine Learning*
- Gupta, Anushri *A soft ecosystems methodology of digital innovation through a case study of the insurance industry's response to connected cars*

**H**

<u>Hamblin, Kate</u>	<u><i>The impact of the digitalisation of care on older, unpaid carers</i></u>
<u>Han, Kate</u>	<u><i>Driving Student Success through a Data-Driven Approach in Higher Education</i></u>
<u>Harindranath, G.</u>	<u><i>Responsible digital: co-creating safe, secure and wise digital interventions with vulnerable groups</i></u>
<u>Hausvik, Geir Inge</u>	<u><i>A Conceptual Framework and Design Principles for Decision Support in Clinical Practice: Managing Knowledge and Uncertainty</i></u>
<u>He, Ying</u>	<u><i>Cybersecurity Investments Metrics using FAIR-ROSI</i></u>
<u>Heaslip, Graham</u>	<u><i>Mind the gap: using threshold concept theory to advance blockchain education</i></u>
<u>Heavin, Ciara</u>	<u><i>Factors influencing the career decisions of women software entrepreneurs: perspectives from India and Ireland</i></u> <u><i>Digital Ethics: Resolving “Wicked” Problems and Dilemmas</i></u>
<u>Heising, Luca</u>	<u><i>Understanding Artificial Intelligence For Data With High Level Of Abstraction: Beyond Pixel Importance</i></u>
<u>Honary, Mahsa</u>	<u><i>Uncovering the Landscape of Ethical Artificial Intelligence</i></u>
<u>Husamaldin, Laden</u>	<u><i>Project Materiality and Integrated Analytics in Sustainability Augmented Portfolio Management</i></u> <u><i>The Technological Challenges of Blended Learning in Higher Education: A case study of Blackboard</i></u>
	
<u>Irizarow, Aleksandra</u>	<u><i>Media Choice for Multi-motive Communication: Impersonal and Interpersonal Motives in Social Worker-Client Relationships</i></u>
<u>Irshad, Waseem</u>	<u><i>Sustainability of business lies in its ability to be resilient; AR technology resilience in frontline employee, the case of retail sector</i></u>
	
<u>Jackson, Jonathan</u>	<u><i>An Exploration of Contemporary Factors that Influence Adoption of Instant Messaging by Academics in Higher Education</i></u>

Joel-Edgar, Sian *Generative AI as a Tool for Thematic Analysis: An Exploratory Study with ChatGPT*

**K**

Kamariotou, Maria *Digital transformation in the public sector: Critical factors of Digital Justice Portal*

Kehoe, Lory *Mind the gap: using threshold concept theory to advance blockchain education*

Kitsios, Fotis *Digital transformation in the public sector: Critical factors of Digital Justice Portal*

Klyshbekova, Maira *Digital inclusion network building: a network weaving analysis*

Krishnan-Harihara, Subrahmaniam *Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs*

Kutar, Maria *Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs*  
*Conceptualising the Artificial Intelligence Divide: A Systematic Literature Review and Research Agenda*

**L**

Lee, Joyce Yi-Hui *The Power of Streamers' Speech: A Signalling Approach to Live Streaming Commerce*  
*Online Health Communities for Parents of Children with Autism Spectrum Disorder*  
*The Role of Middle Managers in Steering Digital Transformation: A Dynamic Capabilities Perspective*

Li, Honglei *Apparel Matters? A New First-Impression Static Appearance Dataset*

Lin, Hua-Xuan *The Power of Streamers' Speech: A Signalling Approach to Live Streaming Commerce*

Lin, Yi-Ling *The Power of Streamers' Speech: A Signalling Approach to Live Streaming Commerce*

Lorini, Maria Rosa *Responsible digital: co-creating safe, secure and wise digital interventions with vulnerable groups*

Luo, Cunjin *Cybersecurity Investments Metrics using FAIR-ROSI*

**M**

Matobobo, Courage *Impact of WhatsApp Groups on Teaching and Learning in Higher Education: A Review*

<u>Mazumdar, Suvodeep</u>	<u><i>Data-driven Sustainable Cities in the UK: The Case of Public Space and Social Cohesion</i></u>
<u>McCarthy, Pat</u>	<u><i>The Underlying Practices of Digital Transformation Leadership: Theorising the Practitioner Voice</i></u>
<u>Mo, Kitty Yuen-Han</u>	<u><i>Media Choice for Multi-motive Communication: Impersonal and Interpersonal Motives in Social Worker-Client Relationships</i></u>
<u>Moe, Carl Erik</u>	<u><i>A Conceptual Framework and Design Principles for Decision Support in Clinical Practice: Managing Knowledge and Uncertainty</i></u>
<u>Moeini, Mo</u>	<u><i>Fostering inclusion in open strategy: three IT-enabled mechanisms</i></u>
<u>Mogles, Nataliya</u>	<u><i>An Exploration of Contemporary Factors that Influence Adoption of Instant Messaging by Academics in Higher Education</i></u>
<u>Mohammad, Javed</u>	<u><i>Project Materiality and Integrated Analytics in Sustainability Augmented Portfolio Management</i></u>
<u>Mohammed, Aishatu Lawan</u>	<u><i>Conceptualising the Artificial Intelligence Divide: A Systematic Literature Review and Research Agenda</i></u>
<u>Morton, Josh</u>	<u><i>Another tool for the toolbox? How strategists are adopting social media for sensemaking</i></u>
<u>Moussa, Hanan</u>	<u><i>Leveraging Generative AI in Information Systems Development</i></u>
<u>Mozaffar, Hajar</u>	<u><i>Sourcing Trends for the Digitalisation of the Health Sector: English National Health System</i></u>
<u>Murugaraj, Pon Rahul</u>	<u><i>Building Sustainable Software - Measuring the Effects of Multithreading on the Carbon Footprint of Software Systems</i></u> <u><i>Context-Aware Emotional State Inference: Extending the Affective Norms of Valence, Arousal, and Dominance for N-grams</i></u>
<u>Müller, Sune</u>	<u><i>Sharing of personal health information for secondary use: A scoping review from the perspective of trust</i></u>
 <u>Nagle, Tadhg</u>	<u><i>Unpacking the 'Black Boxes' of Analytics: A Data Value Map Analysis</i></u>



*Becoming Analytical Champions: A Simple Hermeneutics-inspired Learning Analytics Model for Modules*

Nguyen, Hai

*From Connections to Classrooms: Utilizing Social Media platforms for Enhanced Engagement: The case of LinkedIn. Self-Determination Theory in Crowdfunding: Data from Vietnam*

**O**

O'Connor, Yvonne

*Digital Ethics: Resolving “Wicked” Problems and Dilemmas*

O'Driscoll, Kieran

*Unpacking the ‘Black Boxes’ of Analytics: A Data Value Map Analysis*

O'Flaherty, Brian

*Factors influencing the career decisions of women software entrepreneurs: perspectives from India and Ireland*

Ojutiku, Toluwanimi

*Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs*

Omon, Ubongabasi Kingsley

*Mapping and Visualising the Digital Economy in The Context of Developing Countries: A Bibliometric Analysis*

Onofrei, George

*Mind the gap: using threshold concept theory to advance blockchain education*

Ou, Carol

*Understanding Artificial Intelligence For Data With High Level Of Abstraction: Beyond Pixel Importance*

*Smart maintenance at offshore wind farms: A digital System of Systems approach*

Ou, Su Hua

*Business models for digital sustainability: The case of Fintech startups*

**P**

Pai, Peiyu

*Understanding Consumers’ Reactance of Technology-Enabled Personalization: Perspectives on Stress and Fatigue*

Pan, Liucen

*Fostering inclusion in open strategy: three IT-enabled mechanisms*

Pan, Yu-Chun

*Generative AI as a Tool for Thematic Analysis: An Exploratory Study with ChatGPT*

Panteli, Niki

*Developing the foundations for an inclusive IS education – a case method approach for promoting female digital leaders*

*Emotional regulation for improved student learning: can smartwatches be utilized to navigate an increasingly digital student context?*

*Responsible Digital Innovation with Micro-Businesses: A Resource Orchestration Perspective*

Pappas, Ilias

*Operationalizing Algorithmic Fairness: Empirical Study and Sociotechnical Framework Proposal*

*Responsible AI Principles: Findings from an Empirical Study on Practitioners' Perceptions*

Parsons, Micheal

*From Connections to Classrooms: Utilizing Social Media platforms for Enhanced Engagement: The case of LinkedIn.*

Pedersen, Keld

*Sharing of personal health information for secondary use: A scoping review from the perspective of trust*

Peng, Chih-Hong

*The Power of Streamers' Speech: A Signalling Approach to Live Streaming Commerce*

Platt, Jodyn

*Sharing of personal health information for secondary use: A scoping review from the perspective of trust*

Pujadas, Roser

*A soft ecosystems methodology of digital innovation through a case study of the insurance industry's response to connected cars*

Purao, Sandeep

*Responsible AI Principles: Findings from an Empirical Study on Practitioners' Perceptions*

**Q**

Qi, Jie

*Data-driven Sustainable Cities in the UK: The Case of Public Space and Social Cohesion*

Qin, Tiantian

*Another tool for the toolbox? How strategists are adopting social media for sensemaking*

**R**

Rehman, Ikram

*Project Materiality and Integrated Analytics in Sustainability Augmented Portfolio Management*

Richardson, Kathleen

*A Method for Analysing Ethical Issues in Emerging Technologies – the TechEthos Approach (TEAeM)*

Rousaki, Anastasia

*The impact of the digitalisation of care on older, unpaid carers*

**S**

Saadati, Parisa

*The Technological Challenges of Blended Learning in Higher Education: A case study of Blackboard*

Sammon, David

*Unpacking the 'Black Boxes' of Analytics: A Data Value Map Analysis*  
*The Underlying Practices of Digital Transformation Leadership: Theorising the Practitioner Voice*  
*Becoming Analytical Champions: A Simple Hermeneutics-inspired Learning Analytics Model for Modules*

Sammour, Yazan

*Design and Development of a Digital Transformation Canvas for SMEs in Developing Countries: a case study of Oman*

Sbaffi, Laura

*The impact of the digitalisation of care on older, unpaid carers*

Scheidel, Timo

*Understanding Artificial Intelligence For Data With High Level Of Abstraction: Beyond Pixel Importance*

Singh, Aditi

*Factors influencing the career decisions of women software entrepreneurs: perspectives from India and Ireland*

Song, Lynda

*Media Choice for Multi-motive Communication: Impersonal and Interpersonal Motives in Social Worker-Client Relationships*

Stelmaszak Rosa, Marta

*Teachers' Perceptions of Students' Use of Generative AI in Summative Assessments at Higher Education Institutions: An Exploratory Study*

Suseno, Yuliani

*Business models for digital sustainability: The case of Fintech startups*

Syme, Tony

*Exploring the Productivity Puzzle: The Relationship between Technology Investment and Organisational Productivity in UK SMEs*

**T**

Tan, Chekfoung

*Teachers' Perceptions of Students' Use of Generative AI in Summative Assessments at Higher Education Institutions: An Exploratory Study*

<u>Tapanainen, Tommi</u>	<u><i>Self-Determination Theory in Crowdfunding: Data from Vietnam</i></u>
<u>Tarhini, Ali</u>	<u><i>What are the Drivers of Non-Fungible Tokens (NFTs) Community Engagement?</i></u>
<u>Trinh, Quoc Dat</u>	<u><i>Self-Determination Theory in Crowdfunding: Data from Vietnam</i></u>
<u>Tseng, Shu-Fen</u>	<u><i>Online Health Communities for Parents of Children with Autism Spectrum Disorder</i></u>
<u>Tsirozidis, Georgios</u>	<u><i>Sharing of personal health information for secondary use: A scoping review from the perspective of trust</i></u>
<u>Tzouramani, Eleni</u>	<u><i>Moving Action Learning Sets Online: Reflecting on Privacy, Intersectionality and Group Failure</i></u>

**U**


<u>Ul Hassan, Umair</u>	<u><i>Temporal Aspects in Cancer Clinical Decision-Making Through Machine Learning</i></u>
<u>Unwin, Tim</u>	<u><i>Responsible digital: co-creating safe, secure and wise digital interventions with vulnerable groups</i></u>

**V**

<u>Vannini, Sara</u>	<u><i>Digital inclusion network building: a network weaving analysis</i></u>
<u>Vardaki, Maria</u>	<u><i>Teaching and Learning through the pandemic; the effects of remote work on women academics</i></u>
<u>Vasconcelos, Ana</u>	<u><i>Data-driven Sustainable Cities in the UK: The Case of Public Space and Social Cohesion</i></u>
<u>Vassilakopoulou, Polyxeni</u>	<u><i>Operationalizing Algorithmic Fairness: Empirical Study and Sociotechnical Framework Proposal</i></u> <u><i>Responsible AI Principles: Findings from an Empirical Study on Practitioners' Perceptions</i></u>
<u>Venters, Will</u>	<u><i>A soft ecosystems methodology of digital innovation through a case study of the insurance industry's response to connected cars</i></u>

**W**

<u>Wagg, Sharon</u>	<u><i>Digital inclusion network building: a network weaving analysis</i></u>
---------------------	--

<u>Wang, Fredrik</u>	<u><i>Operationalizing Algorithmic Fairness: Empirical Study and Sociotechnical Framework Proposal</i></u>
<u>Wang, Jialou</u>	<u><i>Apparel Matters? A New First-Impression Static Appearance Dataset</i></u>
<u>Wang, Jing</u>	<u><i>Virtual Representations as Boundary Objects: a Case of Building Information Modeling</i></u>
<u>Wang, Pin-Ran</u>	<u><i>Online Health Communities for Parents of Children with Autism Spectrum Disorder</i></u>
<u>Wang, Yiyu</u>	<u><i>Temporal Aspects in Cancer Clinical Decision-Making Through Machine Learning</i></u>
<u>Wodak, David</u>	<u><i>Smart maintenance at offshore wind farms: A digital System of Systems approach</i></u>
<u>Woo, Wai Lok</u>	<u><i>Apparel Matters? A New First-Impression Static Appearance Dataset</i></u>
	
<u>Xiao, Ling</u>	<u><i>Developing the foundations for an inclusive IS education – a case method approach for promoting female digital leaders</i></u>
<u>Xin, Tong</u>	<u><i>Cybersecurity Investments Metrics using FAIR-ROSI</i></u>
	
<u>Yuan, Lei</u>	<u><i>Generative AI in the Workplace: A Dive into Employees' Acceptance in China</i></u>
	
<u>Zamani, Efpraxia</u>	<u><i>The impact of the digitalisation of care on older, unpaid carers</i></u> <u><i>On the Temporal Experiences of Remote Workers</i></u> <u><i>Digital inclusion network building: a network weaving analysis</i></u>
<u>Zhu, Ruilin</u>	<u><i>Uncovering the Landscape of Ethical Artificial Intelligence</i></u>