Proceedings of 2025 International Conference on Information Resources Management

23-25 June 2025

Conference theme:

Embracing an Intelligent Future

Hosted by:

Bogazici University, Istanbul, Türkiye

Editors: Lech Janczewski, Barbara Krumay, Bilgin Metin

Conference sponsors:

Publisher:

ISBN 978-0-473-74090-0 ISSN: 2744-6220

Blank page

Table of Content

Welcome Message from Conf-IRM 2025 Conference Co-Chairs

Dear conference attendees

We are thrilled that you will join us at the 2025 International Conference on Information Resources Management (Conf-IRM). Since its inaugural edition in 2008, the conference series has travelled to several countries across the globe, including Canada, the United Arab Emirates, Jamaica, South Korea, Austria, Brazil, Vietnam, South Africa, Chile, China, Egypt and New Zealand. This year's edition takes place in the dynamic city of Istanbul, Türkiye—a historic crossroads of East and West, where centuries of cultural heritage meet forward-looking innovation and technology.

This year marks the 17th edition of Conf-IRM. The conference theme "Embracing an Intelligent Future" refers to the idea of actively accepting and integrating advancements in artificial intelligence (AI) and related technologies into our lives, societies, and industries, recognizing the potential for significant positive transformation while proactively addressing potential challenges that may arise from this rapid technological evolution.

The conference keynote speeches, panels, and sessions will address innovation, collaboration, and knowledge convergence. Whether you are a seasoned professional or senior scholar or just starting your journey in the tech world or early in your academic career, this conference promises to be an enriching experience. Our lineup of keynote speakers will delve into timely and thought-provoking topics, including the human-centred challenges of working with intelligent technologies, strategic leadership in digital transformation, and innovations in AI-driven financial services. Get ready to expand your horizons and gain fresh insights! Also, take advantage of this unique opportunity to connect with fellow tech enthusiasts, industry leaders, and potential research collaborators. Share ideas, swap contact details and forge lasting relationships.

Join us in embracing the spirit of innovation that propels the global IT community into the future. We are excited to welcome you to Conf-IRM 2025, hosted at the historic and scenic Boğaziçi University in Istanbul, Türkiye. Stay connected and be part of the conversation by following #Conf-IRM25 on social media for the latest updates, behind-the-scenes moments, and highlights from across the conference.

Let us together make this conference an unforgettable experience!

Nazim Taskin and Felix B Tan

Conference Committees

Conference Steering Committee Conference Co-Chairs

Nazım Taşkın
Bogazici Universty, Turkiye

Felix B Tan
Auckland University of Technology, New Zealand

Conference Program Co-Chairs

Bilgin Metin
Bogazici University, Turkiye

Barbara Krumar
Johannes Kepler University Linz, Austria

Publications and Proceedings Chair

Lech Janczewski
Auckland University, New Zealand

Conf-IRM Executive Committee

Keynote Speakers

David J. Pauleen (PhD) is an honorary adjunct professor in the School of Management at Massey University, New Zealand and a Distinguished Research Fellow, Graduate Institute and Department of Business Administration at National Chung Cheng University in Taiwan. His current areas of research interest include management wisdom, personal knowledge management, knowledge management, and emerging work practices His work has appeared in numerous journals including Journal of Business Ethics, Information Systems Journal, Journal of Knowledge Management, Behavior and Information Technology, Sloan Management Review, Journal of Management Information Systems, and Journal of Information Technology. He is also editor of the books, Virtual Teams: Projects, Protocols and Processes and Cross-Cultural Perspectives on Knowledge Management and co-editor of Personal Knowledge Management: Individual, Organizational and Social Perspectives and the Handbook of Practical Wisdom: Leadership, Organization and Integral Business Practice His co-authored books include, Wisdom, Analytics and Wicked Problems: Integral decision-making in the data age (Routledge, 2019) and Management Decision Making, Big Data and Analytics (Sage, 2021, second edition in 2026). He is a founding co-editor of the Routledge's Practical Wisdom in Leadership and Organization Series.

Athanasia (Nancy) Pouloudi is Professor of Information Systems Management and serves as Vice-Rector of International Cooperation and Growth at the Athens University of Economics and Business (AUEB), Greece since September 2024. She holds a first degree in Informatics (AUEB, Greece), and an MSc and PhD in Information Systems (both from the London School of Economics, UK). Her studies were supported by a scholarship by the Greek State Scholarship Foundation. Her academic career started as Lecturer at Brunel University, UK (1997-2001). In 2001, she joined the Department of Management Science and Technology at AUEB, initially as Assistant Professor in Information Systems Management, where she was promoted in 2006 to Associate Professor and in 2019 to Full Professor. She has served as Department Chair for two terms, in the academic years 2014-2016 and 2020-2022.

Her research focuses on organizational and social issues in IS adoption and implementation. She has published extensively in top journals including the Journal of the Association for Information Systems, the European Journal of Information Systems, the Information Systems Journal, the Journal of Information Technology, the Journal of Business Ethics and others. She has worked in a number of European Union funded projects on e-business, e-government, digital health and e-society, most recently studying children's digital maturity (digymatex.eu) and gender equality in business schools (targeted-mpi.eu). She is the Deputy Director of the ISTLab research lab of AUEB.

She is the Immediate Past President of the global Association for Information Systems (AIS), following her service as the 29th President in the academic year 2023-2024. She previously served as Region 2

(Europe/Middle East/Africa) Representative on AIS Council (2010-2013) and regularly serves on the AIS Region 2 Board and the Executive Committee of the Mediterranean Conference on Information Systems (MCIS).

She is a Senior Editor of the European Journal of Information Systems (EJIS) and a member of the Editorial Boards of the Journal of the Association for Information Systems (JAIS), IT for Development, Health Policy & Technology, and the International Journal of Society, Information, Communication & Ethics.

Her international academic service has been recognized by the 2016 AIS Sandra Slaughter Service Award and the 2020 AIS Technology Challenge Award.

Tracks and Track Chairs

Track 1: Embracing a Sustainable Intelligent Future
Andreja Pucihar, University of Maribor, Slovenia
Barbara Krumay, Johannes Kepler University Linz, Austria

Track 2: IS and AI Adoption and Digital Transformation
Kasuni Weerasinghe, Auckland University of Technology, NZ
Manuel Muehlburger, Johannes Kepler University Linz, Austria

Track 3: Information Security, Privacy, and Risk Management
Bilgin Metin, Bogazici University
Krassie Petrova, Auckland University of Technology, NZ

Track 4: Future of Education, Health and the Public Sector
Eusebio Scornavacca, Arizona State University, Tempe, US
Hamed Jafarzadeh, Macquarie University, Sydney, Australia

Track 5: Enterprise Information Systems and Industry 4.0
Jairo Gutierrez, Auckland University of Technology, Auckland, New Zealand
Mehmet Aydın, Bogazici University, İstanbul, Türkiye

Track 6: Social Media and Social Networks
Shafiq Alam, Massey University
Ziya Perdahci, Bogazici University

Track 7: IS for Development, Diversity, Inclusion and Equality in Digital Futures
Malcolm Garbutt, University of Western Cape, South Africa
Maryam Mirzaei, Auckland University of Technology, New Zealand

Track 8: Regional Perspectives on IS Management
Aslı Sencer, Boğaziçi University, Istanbul, Turkiye

Track 9: Workshops, Tutorials, and Panels
Felix B Tan - Auckland U of Technology NZ

Track 10: Doctoral Consortium
Mirjana Kljajić Borštnar, University of Maribor
Nuri Başoğlu, Izmir Institute of Technology

Conference Reviewers

- Adil Al-Busaidi
- Amarolinda Klein
- Anita Neumanova
- Annette Mills
- Bernd Schenk
- Brian Cusack
- Charuni Wijethunga
- David Rückel
- Deshanjali Diyasena
- Eranjana Kathriarachchi
- Faheem Hussain
- Farhaan Mirza
- Fei Liu
- G. Harindranath
- Gabriel Cogo
- Ghada Refaat
- Görkem Akdur
- Hamed Jafarzadeh
- Hossein Mosafer
- Ines Janusch
- Iris Groher
- Jairo Gutierrez
- Jalal Sarabadani
- Johannes Sametinger
- Kasuni Weerasinghe
- Katrin Mayrhofer-Schmirl
- Martin Stabauer
- Maryam Mirzai
- Masoud Afshari Mofrad
- Masoud Talebian
- Mee Loong Yang
- Melanie Pareera
- Melika Soleimani
- Michael Trierweiler
- Novy N.R.A. Mokobombang
- Paula Lutui
- Pegah Mavaie
- Saeed Rehman
- Sayan Kumar Ray
- Shafiq Alam
- Shirekha Charls
- Shoba Tegginmath
- Stefan Klikovits
- Sylvia Geyer
- Thomas Fischer
- Thuy Tien Nguyen
- Yassine. Dguidegue
- Yavuz Gunalay
- Ziya Perdahci

Doctoral Consortium

Participants

- Alaa Rabea, *Bogazici University, Turkiye*
- Albina Gaisina, *Istanbul Technical University, Turkiye*
- Ecem Akar Ertem, *University of Turku, Finland*
- Hasan Soydan, *Bogazici University, Turkiye*
- Yunus Emre Bulut, *Bogazici University, Turkiye*

Mentors

- Felix B Tan, *Auckland University of Technology, New Zealand*
- Gerald Grant, *Carleton University, Canada*
- Kasuni Weerasinghe, *Auckland University of Technology, New Zealand*
- Nancy Pouloudi, *Athens University of Economics and Business, Greece*
- Shafiq Alam, *Massey University, New Zealand*

Conference Program

Day 1 (Monday, 23rd June) - Doctoral Consortium	
10.30 - 11.00	**Registration**
11.00 - 11.15	**Welcome and Introduction**
11.15 - 12.00	**Keynote: The Importance of Developing Perspective** *David J. Pauleen, Massey University*
12.00 - 13.00	**Doctoral Consortium** *Nuri Başoğlu, Izmir Institute of Technology, Turkiye* *Mirjana Kljajić Borštnar, University of Maribor, Slovenia*
	Students Presentations
13.00 - 14.00	**Lunch Break**
14.00 - 15.30	**Students Presentations**
15.30 - 15.45	**Coffee/Networking Break**
15.45 - 16.00	**Concluding remarks**
18.00 - 21.00	**Welcome Reception**
Day 2 (Tuesday, 24th June) - Track Sessions	
09.00 - 09.45	**Registration**
09.45 - 10.00	**Opening Speech** *Nazım Taşkın, Bogazici University, Turkiye* *Felix B Tan, Auckland University of Technology, New Zealand*
10.00 - 10.15	**Rectorate Speech**
10.15 - 10.45	**Keynote** Abdulgadir Uyar, *Vakıfbank, Turkiye*
10.45 - 11.00	**Coffee/Networking Break**
11.00 - 11.45	**Keynote** *Nancy Pouloudi, Athens University of Economics and Business, Greece*
11.45 - 12.15	**Sponsor Speech** *Cenk Erkin, BeosinTR, Turkiye,*
12.15 - 13.15	**Lunch Break**
13.15 - 13.55	**Track 01 - Embracing a Sustainable Intelligent Future** *Andreja Pucihar, University of Maribor, Slovenia* *Barbara Krumay, Johannes Kepler University Linz, Austria*
	14 - Swift Crisis Response: The Roles of Organizational Mindfulness, Citizenship Behavior and Digital Infrastructure *Edward Bernroider, Everist Limaj,Nikolaus Obwegeser*
	2 - A Methodological Process Model for Recognizing Digital Transformation Opportunities *Manuel Muehlburger, Barbara Krumay, Thomas Deutsch*
13.55 - 14.50	**Track 03 - Security, Privacy, and Risk Management** *Bilgin Metin, Bogazici University, Turkiye* *Krassie Petrova, Auckland University of Technology, New Zealand*
	10 - Mitigating the Risk of Fake User-Generated Content *Christopher Gusenbauer, Michaela K. Trierweiler*

	7 - Explaining Perceived Quality of a Tertiary Education Recommendation System through user studies *Unathi September*
	21R - Towards a Holistic Conceptual Framework for Supply Chain and Third-Party Cybersecurity Risk Management (Research-in-Progress) *Muhammed Yusuf Akçakaya, Meltem E Mutlutürk, Bilgin Metin*
14.50 - 16.05	**Coffee/Networking Break**
16.05 - 17.20	**Track 05 - Enterprise Information Systems and Industry 4. 0** *Jairo Gutierrez, Auckland University of Technology, Auckland, New Zealand* *Mehmet Aydın, Bogazici University, Istanbul, Turkiye*
	11 - Perceptions and Governance of Emerging Technologies in New Zealand: Preparing the Next Generation for an IT-Driven Future *Shafiq Alam, Kasuni Weerasinghe, David J Pauleen,Hamed Jafarzadeh, Nazim Taskin, Ji Yu*
	5 - Corporate data Diogenes: Analyzing companies that behave like digital accumulators *Ariel I. La Paz, Josue A. Salinas, David Lopez*
	9 - Looking Beyond Technology: Hierarchical List of People and Culture Capabilities for Advanced Data Analytics *Tina Afshar Ghochan, Özden Özcan-Top, Banu Aysolmaz*
	19R - Exploring the Human-Centric Perspectives of Smart Warehouses Technology Adoption - the case of New Zealand *Maryam Mirzaei, Benjamin Dehe*
17.20 - 17.30	**Closing remarks**
Day 3 (Wednesday, 25th June) - Track Sessions	
09.00 - 09.30	**Registration**
09.30 - 11.10	**Track 02 - IS Adoption and Digital Transformation** *Kasuni Weerasinghe, Auckland University of Technology, NZ* *Manuel Muehlburger, Johannes Kepler University Linz, Austria*
	1 - A Framework for data space adoption integrating data sovereignty as a technical factor *Andreas Hutterer, Barbara Krumay*
	13 - Recommendations for teleconsultation implementation for healthcare providers: A systematic review *Denise Selina Y. Tan, Felix B Tan*
	8 - Generative Artificial Intelligence and Digital Transformation in Contact Center Businesses *Shafiq Alam; Lorraine Skelton; Muhammad Sohaib Ayub; Windy Dharmawan*
	16 - The Role of AI-Powered Personalisation throughout the Purchase Decision-Making Process on Online Marketplaces *Nodumo Temban, Maureen Tanner,Papama Mtambeka*
	17 - Transforming Customer Service in E-Commerce: The Case of Hepsiburada's Bespoke Platform *Yavuz Selim Hindistan,Handan Derya Ercan, Tayfun İskender, Hüseyin Özdemir,Rıdvan Yıldırım*
11.10 - 11.25	**Coffee/Networking Break**

11.25 - 12.50	**Track 04 - Future of Education, Health and the Public Sector** *Eusebio Scornavacca, Arizona State University, Tempe, US* *Hamed Jafarzadeh, Macquarie University, Sydney, Australia*
	12 - Predicting Learning Styles with AI: Toward Adaptive and Personalized Education *Md Mehedi Hasan, Gobinda Chandra Sarker, Md Rakibul Hoque, Mohammad Ahmed*
	3 - Adopting DevSecOps: A Framework for IT Governance and Culture Change Based on a Plan-Do-Check-Act (PDCA) Approach *Bilgin Metin, Timur Demir, Ali Can Keserel, Barışcan Güngör, Martin Wynn*
	18R - Rethinking Doctoral Research and Supervision in the Generative AI Era: A Research Agenda *Melika Soleimani, Kasuni Weerasinghe, Maryam Mirzai, David J Pauleen*
	22R - AI's Limits in Generating Audio-Visual Media for Communication of Serious Content *Ines Janusch, Simon Currle*
	20R - WeChat as a Digital Health Tool: Understanding the Role of Digital Health Literacy in its Effective Use among Elderly Diabetic Patient in China *Ruilin Zhu, Zhiqin Wu*
12.50 - 13.50	**Lunch Break**
13.50 - 14.50	**Track 06 - Social Media and Social Networks** *Shafiq Alam, Massey University, NZ* *Ziya Perdahci, Bogazici University, Turkiye*
	15 - The Pulse of the Paris Olympics: Sentiment and Topic Modelling of Opening Ceremony *Yunus Emre Bulut,Nazım Taskın*
	6 - Cyberbullying Detection With Machine Learning & Psychology: A Systematic Review *Lubaina Navaid, Faisal Iradat, Nazim Taskin*
	4 - Bridging the Two Sciences: Tapping into the Confluence of Network Science and Design Science Research *Samrat Gupta, Thorsten Schoormann, Mukul Gupta, Leona Chandra Kruse*
14.50 - 15.05	**Coffee/Networking Break**
15.05 - 15.20	**Closing Remarks**
17.30	Shuttle from Campus to Dinner Venue
18.00 - 22.00	**Gala Dinner & Award Ceremony**

Conference Papers Index

Notes:

- Number in brackets: page number of the paper text.
- Letter **R** indicates Research-in-Progress paper
- All these papers are also uploaded to https://aisel.aisnet.org/confirm2025/

Authors Index

Conference Papers

1. A FRAMEWORK FOR DATA SPACE ADOPTION INTEGRATING DATA SOVEREIGNTY AS A TECHNICAL FACTOR

Andreas Hutterer
Johannes Kepler University Linz
andreas.hutterer@jku.at

Barbara Krumay
Johannes Kepler University Linz
barbara.krumay@jku.at

Abstract
Digital transformation allows organizations to increase data transparency, but unfortunately, organizations still encounter obstacles to data sharing as they fear the loss of control over data. Data spaces as cross-organizational data infrastructures can potentially enhance organizations' control over their data. However, the adoption of data spaces has yet to be fully investigated. Therefore, this study attempts to address this gap by applying a Design Science Research (DSR) approach for developing an artefact to support the adoption of data spaces. We designed and evaluated the framework through expert interviews. As a result, we propose a final data space adoption framework comprising 14 factors originating from the Technological-Organizational-Environment (TOE) framework. The framework contributes to the existing body of knowledge regarding influencing factors and supports the adoption of data spaces integrating data sovereignty within organizations.

Keywords
Data Space, Adoption, TOE Framework, Data Sovereignty.

1 Introduction

The second wave of digitalization has led to significant shifts, notably through the rise of platform markets and digital ecosystems (Kagermann & Winter, 2018). This evolution has brought about data ecosystems, emphasizing data as a key resource (Otto, 2022b). Recognizing data as a strategic asset enhances competitiveness in the digital economy, thereby increasing the research focus on data ecosystems (Gelhaar & Otto, 2020). In particular, the development of dynamic data ecosystems is seen as a means to achieve competitive advantage (De Prieelle et al., 2020). Organizations are beginning to acknowledge the benefits of cross-organizational data sharing, shifting towards horizontal approaches (Gelhaar et al., 2023), yet such practices remain relatively uncommon (Fassnacht et al., 2023; Jussen et al., 2023). Data spaces aim to address these shortcomings by providing solutions to organizational concerns about data sharing (Möller et al., 2024). Data spaces, which are platform-based data ecosystems, have recently attracted academic interest as a potential means of counteracting monopolistic tendencies. (Beverungen et al., 2022; Rotgang et al., 2023). In addition, data spaces are multi-sided platforms, benefitting from multiple participants adding value (Otto & Jarke, 2019). In this context, data sovereignty may enhance the willingness to share data, but further exploration is needed to operationalize data sovereignty (Hellmeier & von Scherenberg, 2023; von Scherenberg et al., 2024).

Besides technical challenges, the legal concept of data ownership in Europe remains undefined (Jarke, 2020). Data spaces are intended to address this ambiguity by establishing data sovereignty within data ecosystems (Hutterer & Krumay, 2022; Jarke et al., 2019), also becoming an

emerging topic for inter-organizational information systems (Opriel et al., 2024). Previous studies have focused on prerequisites such as digital transformation capabilities (Hupperz & Gieß, 2024), governance (Schurig et al., 2024), digital transformation opportunities (Hutterer & Krumay, 2025) and adoption drivers (Hutterer & Krumay, 2024). However, the existing literature lacks a solid foundation for data space adoption (Hirsch-Kreinsen et al., 2022). As cross-organizational data infrastructures, data spaces impact data sharing across boundaries (Möller et al., 2024). Still, limited clarity on how data sovereignty is applied to data spaces (von Scherenberg et al., 2024) challenges the motivation for the adoption of data spaces (Hutterer & Krumay, 2022), which has also not been adequately addressed (Mertens et al., 2024). Given the multi-sided nature (Otto & Jarke, 2019), adoption of data spaces seems to be beneficial from an organizational perspective within the context of a multi-stakeholder community. This study addresses this gap by conceptualizing a data space adoption framework adopting a Design Science Research approach (Beck et al., 2013). The structure of the paper is as follows: first, we provide an overview of the key aspect considered in this study, focusing on data ecosystems and data spaces. This is followed by a description of the research approach employed and the result presentation, which includes the development and evaluation of the data space adoption framework. Finally, we discuss the results, identify limitations and suggest directions for future research.

2. Background Information

Data spaces also offer an adaptable integration model that complements centralized storage of structured data to consider increasing data heterogeneity and the growing demand for flexible data management approaches (Franklin et al., 2005). Although the terms ’data spaces’ and ‘data ecosystems’ are often used interchangeably, strictly speaking, they differ significantly (Curry, 2020). Data spaces are platform-based ecosystems designed to facilitate data sharing among previously isolated systems, functioning as inter-organizational information systems (Hutterer & Krumay, 2022; Möller et al., 2024). This shift from centralized to decentralized data storage is transforming organizational data management concepts (Halevy et al., 2006) overcoming struggles from traditional database management systems (i.e., handling of heterogeneous data) In contrast, data spaces centralize schemas while decentralizing data storage, enabling seamless integration across various structures, locations, or formats (Curry, 2020). Data Space Support Platforms (DSSPs) offer minimal integration requirements and increased autonomy, progressively enhancing their capabilities through a pay-as-you-go approach (Sarma et al., 2009). Data spaces employ approximation techniques for flexible data integration, focusing on data relationships rather than comprehensive semantic integration (Hedeler et al., 2012; Singh, 2013). DSSPs also lay the foundation for multi-sided data platforms that enable federated data sharing among disparate organizations (Otto & Jarke, 2019). The design of data spaces is highly variable, encompassing a range of architectural approaches, from centralized to decentralized (Gieß et al., 2023; Schleimer et al., 2023). Organizational adoption of data spaces is emerging (Mertens & Kuster, 2024), developing artefacts supporting the evolving concept of data sovereignty is necessary (von Scherenberg et al., 2024). For this study, we understand data spaces as follows:

> “Data spaces are flexible and open IT structures that guarantee the absolute sovereignty of the participants involved. As such, they facilitate the trustworthy and transparent use of decentralized data according to pre-defined scopes of use. Data spaces are set up as federal entities and are based on various basic concepts/elements:
>
> a. Data space operation and management services (core services)
>
> b. Technical standards
>
> c. Operational processes
>
> d. Regulatory frameworks (governance model)

A key feature of data spaces is that they create a level playing field for sovereign data sharing. This means that all participants can benefit from the use of data in the same way." (acatech, 2024)

3. Research Approach

Our research approach follows methodological recommendations (Beck et al., 2013), combining the design science paradigm (Baskerville, 2008) with elements of grounded theory—an interpretative method (Strauss & Corbin, 1994)—into a methodology comprising three phases: Literature Collection, Conceptualization, and Evaluation. We focus on the "building" and "evaluating" phases of Design Science Research (Hevner et al., 2004; March & Smith, 1995), which is suitable for research artefact development in specific contexts (Benbasat et al., 1987). This involves developing innovative artefacts to enhance organizational capabilities (Hevner et al., 2004). Grounded Theory (GT), by contrast, allows to analyse of qualitative data and formulate theories from data. Our approach includes identifying needs, designing and evaluating methodologies, and theorizing (Peffers et al., 2020). The adaptability of design science allows integration with various approaches, as seen for blockchain (Beck et al., 2016). We employ methods from GT like theoretical sampling to enhance understanding of artefact utilization. Starting with a literature review on technology adoption and data spaces, we conceptualized a data space adoption framework informed by expert consultations to ensure rigor and practical relevance (Goldkuhl, 2004; Hevner, 2007). To gain more insights on the topic, semi-structured interviews were conducted with experts. Interviews help to reconstruct specialized knowledge vital for practical application (Flick, 2022) and are recommended for understanding the problem and assessing the artefact's quality (vom Brocke & Buddendick, 2006). Our iterative approach remained responsive to emerging insights, allowing continual refinement of the artefact (Gregory, 2011). Data collected from experts led to the refinement of the final data space adoption framework proposed in this paper.

The literature review on data spaces allowed identifying factors influencing data space adoption (Hutterer & Krumay, 2024). We chose to apply the Technology-Organization-Environment (TOE) framework as the theoretical foundation for our research due to its flexibility and its ability to integrate diverse perspectives of data spaces (Hupperz & Gieß, 2024; Hutterer & Krumay, 2023). The TOE framework supports informed technology adoption decisions across multiple technological domains due to its flexibility and integration of diverse perspectives (De Prieelle et al., 2020), as demonstrated by its successful application in areas such as cloud computing, the Internet of Things (IoT), and enterprise systems (Oliveira et al., 2014). In Spring 2024, we consulted experts (A–J) to map the identified factors (Hutterer & Krumay, 2024) with the TOE framework's dimensions, ensuring the framework was technically sound and practically applicable (Morse, 1991; Schultze & Avital, 2011). We rigorously assessed the preliminary data space adoption framework using a criteria-based approach (Betz & Jung, 2021). To refine it, we interviewed additional experts (K–U) until theoretical saturation was reached. We conducted 11 online, semi-structured interviews with data space experts from Central Europe in June and July 2024. The interviews were recorded, transcribed, anonymized, and analysed based on criteria such as completeness, fidelity to real-world phenomena, internal consistency, level of detail, and robustness (Sonnenberg & Vom Brocke, 2012). Table 1 presents a description of the interviewees.

4. Preliminary Data Space Adoption Framework

This section presents the results of the conceptualization phase, which builds on our previously outlined three-stage research approach. Following consultations with experts in the field, a preliminary framework for data space adoption was developed. This framework categorizes the drivers for adoption (Hutterer & Krumay, 2024) into three main dimensions: technological, organizational, and environmental. This categorization is intended to provide a theoretical basis for understanding the process of organizational data space adoption. While the majority of adoption factors were classified into a single dimension by all experts, allowing for straightforward assignment, some factors were initially assigned to multiple dimensions. Notably, the adoption drivers "Controllable complexity", "Technology competence", and "Trust" could not be distinctly classified within one dimension. Consequently, these factors were allocated to both the technological and organizational dimensions, resulting in the additional inclusion of "Manageable complexity" and "Digital literacy" within the organizational dimension. Furthermore, the "Trust" factor was divided into "Technology trust" and "Ecosystem trust". A description was developed to address the characteristics of each factor based on its specific focus. Where possible, independent factors were defined, acknowledging that some may overlap or influence each other, as evidenced by the relationship between “Security” and “Technology trust”. Figure 1 illustrates the preliminary data space adoption framework, which includes 15 factors within the dimensions of technological, organizational, and environmental.

Phase	Expert	Position	Phase	Expert	Position
Conceptualization	Expert A	Industry-University Collaboration Lead	Evaluation	Expert K	Researcher
	Expert B	Project Manager		Expert L	CTO
	Expert C	Co-Founder and Business Lead		Expert M	Project Director Data Space
	Expert D	Associate		Expert N	CEO
	Expert E	Developer at Cloud-Native Architecture		Expert O	Project Manager Industrial Data
	Expert F	Solution Architect		Expert P	Manager IT-Strategy
	Expert G	Project Coordinator		Expert Q	Researcher
	Expert H	Deputy CTO		Expert R	CEO
	Expert I	Project Director Data Space		Expert S	Researcher
	Expert J	Head of Digitalization and Resilient Production		Expert T	Consultant
				Expert U	CEO

Table 1: Participants of the conceptualization and evaluation phase

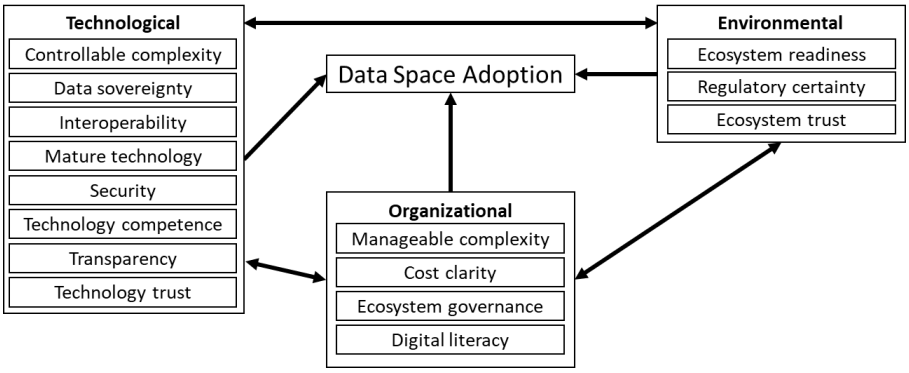

Figure 1. Preliminary Data Space Adoption Framework

Secondly, we conducted the evaluation using the approach of Sonnenberg and vom Brocke (2012), utilizing a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree)

across five criteria. The results demonstrated alignment between the expert interviews and the preliminary data space adoption framework. However, experts indicated that further instantiation might be necessary to validate the preliminary adoption framework. The quantitative evaluation results showed a tendency toward confirmation across all five criteria, with an average rating of 4.05 and a standard deviation of 0.12. The highest ratings were observed for criteria 1, 2, and 3, with values ranging from 4.0 to 4.27 and standard deviations of 0.6 and 0.62, respectively. The fourth criterion was confirmed with a rating of 4.09 and a standard deviation of 0.9. In contrast, the fifth criterion received the lowest rating of 3.91 but exhibited the highest standard deviation of 1.08, suggesting greater variability in responses. Feedback from the interviews was integrated into the final data space adoption framework. Table 2 provides a summary of the evaluation results of the preliminary data space adoption framework.

Criteria	Ave	Std	Expert Feedback
Completeness	4,00	0,6	Experts agreed the framework is well-developed but lacks critical elements. Missing components include data productization (Expert T) and "value-added" aspects (Experts M and K). Integrating a business component to create economic incentives was emphasized by Experts S, R, and U. While Expert L considered it complete, others like Experts O and P noted it doesn't fully cover all dimensions of business value. Expert N praised refinements but agreed further improvements are needed. Overall, incorporating a clearer business perspective and economic incentives would enhance completeness.
Fidelity with real world phenomena	4,00	0,6	Though helpful for decision-making, the framework needs practical enhancements to align better with real-world dynamics. Expert T suggested a simplified checklist to reduce complexity. Experts M and K highlighted the lack of practical guidance and emphasis on added value. Expert S saw potential as a "data space readiness" tool with clear metrics. The abstract nature limits practical utility (Experts L and U). Experts R and O noted the absence of immediacy and clarity, particularly due to the lack of a business perspective. Experts P, N, and Q stressed the need for tangible actions and practical metrics reflecting business realities.
Internal consistency	4,27	0,62	Feedback was mostly positive. Experts T and S found high consistency despite overlaps. Expert K suggested revising factor categorizations to enhance the logical structure. Expert R proposed adding a fourth, cross-cutting dimension for issues like trust and security. Experts O and P recommended refining aspects to minimize overlap and ensure clearer definitions. Overall, while the framework is coherent, adjustments are needed to eliminate overlaps and present elements more clearly.
Level of detail	4,09	0,90	Experts acknowledged the comprehensive approach but agreed refinement could enhance comprehensibility. Expert T rated the detail as high but suggested integrating factors. Experts M and K emphasized the need for clearer definitions of complex constructs like data sovereignty. Expert R called for more detailed explanations to enhance clarity. Expert Q criticized the cursory treatment of topics, indicating room for improvement. Collectively, adding depth and potentially restructuring could improve user understanding.
Robustness	3,91	1,08	Experts had mixed views on robustness due to the field's dynamic nature. Concerns about rapid evolution affecting long-term applicability were noted by Experts T and M. Expert K viewed it as relatively robust because of adaptable terms and non-specific technological factors. Expert S was optimistic about long-term robustness despite potential outdated elements. Experts L and U advised against future alterations to maintain resilience. Overall, while designed to endure technological and market changes, continuous updates are necessary to maintain practicability.
Ave: average; StD: standard deviation; n=11			

Table 2: Evaluation of the preliminary data space adoption framework

5. Results

The final data space adoption framework was developed by incorporating experts' (K–U) feedback from the second evaluation phase. Discussions with experts from both academia and industry led to minor adjustments that enhanced the framework's accuracy and applicability. The final data space adoption framework shows a shift of focus from technological to environmental aspects. In the technological dimension, experts confirmed the relevance of factors such as 'Controllable complexity', 'Data sovereignty', 'Interoperability', 'Mature technology', 'Security', and 'Transparency'. The evaluation particularly focused on factors with discrepancies

in the initial allocation. Notably, there was considerable debate over ‘Controllable complexity‘, ‘Technology competence‘ and ‘Technology trust‘ in relation to ‘Manageable complexity‘, ‘Digital literacy"‘ and ‘Ecosystem trust‘. Experts concluded that ‘Controllable complexity‘ is distinct from ‘Manageable complexity‘, encompassing different elements and closely associated with "Technology competence’. Based on their feedback, ‘Controllable complexity’ was deemed more appropriate within the technological dimension, as it pertains to the inherent complexity of data spaces. Conversely, ‘Technology competence’, reflecting an organization's capacity to align its structures and processes with the data space, was relocated to the organizational dimension and merged with ‘Manageable complexity’. In the organizational dimension, experts confirmed the significance of ‘Cost clarity’ and ‘Digital literacy’. Additionally, ‘Technology trust’ and ‘Ecosystem trust’ were scrutinized. Experts observed that trust is generated by technology and inherent in factors like ‘Data sovereignty’, ‘Interoperability’, ‘Mature technology’, ‘Security’ and ‘Transparency,’ yet it also influences the ecosystem. Consequently, ‘Technology trust’ was removed from the technological dimension and redefined as ‘Trust’ within the environmental dimension. In the environmental dimension, experts confirmed the factors ‘Ecosystem readiness’ and ‘Regulatory certainty’. This shift emphasizes the importance of external factors in the adoption process. Figure 2 illustrates the final data space adoption framework, comprising 14 factors distributed across the three dimensions: technological, organizational, and environmental. This comprehensive framework provides a structured approach for organizations to evaluate and navigate the complexities of adopting data spaces, facilitating informed decision-making in cross-organizational data sharing initiatives.

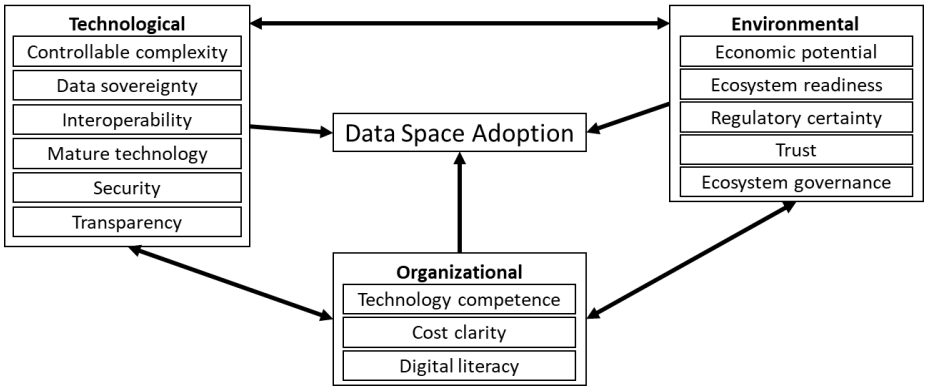

Figure 2: Final Data Space Adoption Framework

Experts identified several improvements within the technological dimension of the preliminary data space adoption framework, which originally consisted of eight factors accounting for half of all factors. One key factor highlighted was ‘Controllable complexity’, addressing the inherent complexity of the data space approach. Expert N noted this as a barrier due to the absence of plug-and-play components, stating that organizations "have to do a lot of software development and customization of existing connectors". Additionally, ‘Data sovereignty’ was emphasized as critically important. It allows data owners to control access permissions, distinguishing data spaces from other technologies. Expert L remarked on its significance for organizations handling sensitive data: "The issue of data sovereignty is of significant importance to organizations that utilize data spaces". However, Expert Q pointed out the technical challenges in implementing data sovereignty effectively. Furthermore, "Interoperability" emerged as a vital aspect, achievable at technical, semantic, organizational, and legal levels. It encompasses both intra-dataspace interoperability—compatibility among participants within a single data space—and cross-dataspace interoperability—accessing data across multiple data spaces. Expert M

underscored its importance: "It's about breaking down silos between domains and data spaces... creating an ecosystem of ecosystems". Moreover, the factor of ‘Mature technology’ remains relevant. While data space technologies offer significant benefits, their maturity level can hinder adoption. Expert S summarized this concern: "An organization is unlikely to adopt a technology that is not yet mature if it still has doubts about it". In addition, ‘Security’ was confirmed by all experts as essential. Data spaces involve certified entities, components, and services, such as IDSA-certified connectors or GAIA-X Digital Clearing Houses certifications. Expert T highlighted the importance of security: "Security has been taken into account in the development of data spaces". Lastly, ‘Transparency’ enables stakeholders to trace data transactions and verify data integrity. However, it's not always viewed positively. Expert T noted: "For me as an organization, the fact that there is this transparency is also more of a negative aspect".

Moving on to the organizational dimension, experts agreed with existing factors but suggested refinements. ‘Manageable complexity’ was reclassified to address an organization's internal data maturity and readiness for cross-organizational data sharing, distinct from the technological ‘Controllable complexity’. It was reframed as ‘Technology competence’. Expert N commented on the challenges: "It takes time to get the organizational structures in place... it's actually a strong change process in the organization". Moreover, the ‘Cost clarity’ factor was affirmed, acknowledging that adopting new technology incurs costs related to data collection, processing, and storage. Expert R described this: "Providing data involves significant costs... I may have to transform it in some way". Furthermore, the broad ‘Digital literacy’ factor was debated. Experts suggested splitting it to focus on internal competencies and separately listing the potential of data monetization. Expert K explained: "The problem is that when I see the term 'digital literacy', I don't immediately think of business models. The data monetization aspect should be a separate factor". Consequently, the ability to monetize data was recognized as a distinct factor, leading to the addition of ‘Economic potential’ in the environmental dimension.

In the environmental dimension, the framework was refined to include more specific ecosystem factors. Experts identified that the preliminary framework lacked emphasis on the potential value of data in data spaces—a crucial adoption factor. The new ‘Economic potential’ factor addresses the need for organizations to benefit from data space adoption through data monetization or increased efficiency. Expert R highlighted this necessity: "Is there added value for an organization to offer data in the data space instead of selling it to a third party?". Additionally, due to network effects, ‘Ecosystem readiness’ was deemed important. Adoption increases when many participants are willing to share data. Expert P stressed: "It is important for organizations that there is an ecosystem ready to work with these techniques and technologies... all parts of the ecosystem need to be in place". Furthermore, ‘Regulatory certainty’ was identified as necessary, ensuring that regulations and standards for data sharing are clear, which builds trust among participants. Expert U mentioned: "Regulatory certainty is very important for us because it builds trust". Moreover, the factor of ‘Trust’ was also emphasized. Data spaces require collaboration toward a common purpose, and mechanisms like GAIA-X Digital Clearing Houses act as trust anchors. Expert M explained: "Trust refers to the confidence that participants have in each other... it's more of a community issue or an environmental management aspect". Lastly, ‘Ecosystem governance’ was included in this dimension, acknowledging that democratized governance is crucial for orchestrating data spaces. The Dataspace Governance Authority (DGA) serves as the governing body, which can exist independently of any single organization. Expert T noted: "There's data space governance and ecosystem governance. For the adoption of a data space, this data space-specific governance is the first thing that has to happen". Expert N added: "If it's about the ecosystem of the data space, then that's part of the environmental dimension... it's crucial to whether I decide to participate".

6. Discussion

The growing interest in data spaces necessitates understanding their adoption (Nagel & Lycklama, 2021). To fill the existing research gap (Beverungen et al., 2022), we developed a data space adoption framework as a foundational base for analysing cross-organizational data sharing, covering technological, organizational, and environmental dimensions. This enhances comprehension of adoption complexities and broadens the focus to include managerial and economic aspects (Hellmeier & von Scherenberg, 2023; Otto, 2022b). Therefore, we are providing a foundation for theoretical development, while also supporting the development of practical applications across various data space initiatives (Mertens & Kuster, 2024). Firstly, the framework integrates a unique 'Economic potential' factor, which has not been identified in recent research. The factor describes the monetization opportunities of data space adoption, generally, the possible added value. Existing literature discusses two generic business model types of data spaces itself (Ammann, 2025). Although most data space initiatives start with public funding, they aim to develop a self-sustainable business model over time by generating revenue from use cases—complementing or replacing public funding—and exploiting economies of scale through the reuse of data products across multiple applications (Mertens et al., 2024). In the context of data spaces, business models can be understood to span the range of participants (data owner, data user, intermediaries) and the data space infrastructure operators. Secondly, we identified data sovereignty as a technological factor, which may lead to organizational commitment and trust for data sharing. Recently, data sovereignty has been regarded as an organisational competence (Moschko et al., 2023). Interestingly, the experts regarded data sovereignty as a technological factor that is operationalised through technological solutions such as usage policies at the infrastructure level (Marino et al., 2023). However, data spaces are regarded as sovereign data infrastructures that establish interoperability through standardization (Möller et al., 2024; Otto, 2022a). Specifically, data sovereignty was operationalised as a technological adoption factor within the context of data spaces (von Scherenberg et al., 2024).

Particularly within the European context, there is a growing emphasis on encouraging organizations to enhance transparency regarding data management through the implementation of sustainable regulatory frameworks (Ubacht et al., 2023). The European Union's Data Act mandates organizations to disseminate previously proprietary data and develop products with user accessibility in mind, a concept known as "data accessibility by design" (European Commission, 2023). We assume, that societal demands for data sovereignty and regulations in Europe—such as the Corporate Sustainability Due Diligence Directive, the Carbon Border Adjustment Mechanism, and the Digital Product Passport—can support the adoption of data spaces. Environmental factors thus shape technological design to balance economic opportunities with societal interests in data sharing and reuse (Labadie & Legner, 2019). Consequently, the data space approach was established to strengthen data sovereignty as a technological function. The adoption of data spaces provides organizations with the ability to empower data sovereignty, which can be regarded as an emerging organizational competence within the context of the data economy. While recent work discusses sovereignty as a capability without providing a theoretical foundation (Nagel & Lycklama, 2021), we posit that data sovereignty is an environmental requirement that facilitates cross-organizational data sharing in the data economy. However, MISSION KI has developed a search engine and cataloging service to enhance data ecosystem connectivity by enabling comprehensive dataset searches (Mission KI, 2025). Daseen streamlines data discovery, analysis, and management by enabling metadata-based searches across public and private data portals and data spaces while automating documentation and compliance processes (Daseen, 2025).

7. Conclusion

This study contributes to the current academic discussion on data space adoption by addressing it from a management perspective. We aimed to enhance the understanding regarding factors driving organizations to adopt a data space and embrace federated data sharing across boundaries. Interestingly, data sovereignty as a technological factor seems to play in important role but requires further operationalization (von Scherenberg et al., 2024). The described data space adoption framework is an initial step toward a comprehensive knowledge base on data space adoption, addressing both technical and managerial challenges. We acknowledge limitations due to potential variations in the expert sample that might have an influence on domain-specific factors. Expanding the expert sample could provide further insights. Future research could extend the framework to include these factors, evaluate the impact, and explore approaches for adoption. Examining correlations between identified factors and data space adoption, as well as among the factors themselves, is necessary. Developing quantifiable metrics and measurement items for each factor is necessary. Since our framework is based on the external perspective of adopting organizations, further research should focus on the internal perspective of data space operators.

References

acatech. (2024). *Fact sheet: Data spaces*. acatech – National Academy of Science and Engineering. https://mobility-dataspace.eu/fileadmin/05_presse_medien/230904_factsheet_datenraeume.pdf

Ammann, J. (2025). Data Sharing is Caring -A Multiple Case Study on Business Model Types in Decentralized Data Ecosystems. *58th Hawaii International Conference on System Sciences (HICSS)*. Hawaii, USA.

Baskerville, R. (2008). What design science is not. *European Journal of Information Systems*, *17*(5), 441–443. https://doi.org/10.1057/ejis.2008.45

Beck, R., Czepluch, J. S., Lollike, N., & Malone, S. (2016). Blockchain–the gateway to trust-free cryptographic transactions. *Twenty-Fourth European Conference on Information Systems (ECIS), İstanbul, Turkey, 2016*, 1–14.

Beck, R., Weber, S., & Gregory, R. W. (2013). Theory-generating design science research. *Information Systems Frontiers*, *15*(4), 637–651. https://doi.org/10.1007/s10796-012-9342-4

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The Case Research Strategy in Studies of Information Systems. *MIS Quarterly*, *11*(3), 369. https://doi.org/10.2307/248684

Betz, C., & Jung, R. (2021). Value Creation in Business Ecosystems-A Design Theory for a Reference Model. *AMCIS 2021*. Americas Conference on Information Systems (AMCIS). https://aisel.aisnet.org/amcis2021/virtual_communities/virtual_communities/4

Beverungen, D., Hess, T., Köster, A., & Lehrer, C. (2022). From private digital platforms to public data spaces: Implications for the digital transformation. *Electronic Markets*, *32*(2), 493–501. https://doi.org/10.1007/s12525-022-00553-z

Curry, E. (2020). Dataspaces: Fundamentals, Principles, and Techniques. In E. Curry, *Real-time Linked Dataspaces* (pp. 45–62). Springer International Publishing. https://doi.org/10.1007/978-3-030-29665-0_3

Daseen. (2025). *Large Dataset Search Engine*. https://daseen.de/

De Prieelle, F., De Reuver, M., & Rezaei, J. (2020). The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry. *IEEE Transactions on Engineering Management*, 1–11.

European Commission. (2023). *REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on harmonised rules on fair access to and use of data (Data Act)*.

European Commission. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022SC0034

Fassnacht, M., Benz, C., Heinz, D., Leimstoll, J., & Satzger, G. (2023). Barriers to Data Sharing among Private Sector Organizations. *Proceedings of the 56th Hawaii International Conference on System Sciences*. HICCS.

Flick, U. (2022). *An introduction to qualitative research*.

Franklin, M., Halevy, A., & Maier, D. (2005). From databases to dataspaces: A new abstraction for information management. *ACM SIGMOD Record*, *34*(4), Article 4. https://doi.org/10.1145/1107499.1107502

Gelhaar, J., Müller, P., Bergmann, N., & Dogan, R. (2023). Motives and incentives for data sharing in industrial data ecosystems: An explorative single case study. *Proceedings of the 56th Hawaii International Conference on System Sciences (HICSS)*. Hawaii, USA. https://hdl.handle.net/10125/103085

Gelhaar, J., & Otto, B. (2020). Challenges in the Emergence of Data Ecosystems. *Pacific Asia Conference on Information Systems (PACIS)*, 175.

Gieß, A., Möller, F., Schoormann, T., & Otto, B. (2023). *Design options for data spaces*. Thirty-first European Conference on Information Systems (ECIS 2023).

Goldkuhl, G. (2004). Design theories in information systems-a need for multi-grounding. *Journal of Information Technology Theory and Application (JITTA)*, *6*(2), 7.

Gregory, R. W. (2011). Design Science Research and the Grounded Theory Method: Characteristics, Differences, and Complementary Uses. In A. Heinzl, P. Buxmann, O. Wendt, & T. Weitzel (Eds.), *Theory-Guided Modeling and Empiricism in Information Systems Research* (pp. 111–127). Physica-Verlag HD.

Halevy, A., Rajaraman, A., & Ordille, J. (2006). Data integration: The teenage years. *Proceedings of the 32nd International Conference on Very Large Data Bases*, 9–16.

Hedeler, C., Belhajjame, K., Mao, L., Guo, C., Arundale, I., Lóscio, B. F., Paton, N. W., Fernandes, A. A., & Embury, S. M. (2012). Dstoolkit: An architecture for flexible dataspace management. In *Transactions on Large-Scale Data-and Knowledge-Centered Systems V* (pp. 126–157). Springer.

Hellmeier, M., & von Scherenberg, F. (2023). *A delimitation of data sovereignty from digital and technological sovereignty*. Thirty-first European Conference on Information Systems (ECIS 2023).

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems*, *19*(2), 4.

Hevner, March, Park, & Ram. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75. https://doi.org/10.2307/25148625

Hirsch-Kreinsen, H., Kubach, U., Stark, R., von Wichert, G., Litsche, S., Sedlmeier, J., & Steglich, S. (2022). *Themenfelder Industrie 4.0: Forschungs-und Entwicklungsbedarfe zur erfolgreichen Umsetzung von Industrie 4.0* (No. 2). Forschungsbeirat der Plattform Industrie 4.0/acatech – Deutsche Akademie der Technikwissenschaften.

Hupperz, M., & Gieß, A. (2024). *The Interplay of Data-Driven Organizations and Data Spaces: Unlocking Capabilities for Transforming Organizations in the Era of Data Spaces*. Hawaii International Conference on System Sciences (HICSS).

Hutterer, A., & Krumay, B. (2022). Integrating Heterogeneous Data in Dataspaces—A Systematic Mapping Study. *PACIS 2022 Proceedings*. Pacific Asia Conference on Information Systems (PACIS). https://aisel.aisnet.org/pacis2022/222

Hutterer, A., & Krumay, B. (2023). Adoption of data spaces as multi-sided platforms: Towards a preliminary adoption framework. *ICEB 2023 Proceedings (Chiayi, Taiwan)*. International Conference on Electronic Business (ICEB). https://aisel.aisnet.org/iceb2023/1

Hutterer, A., & Krumay, B. (2024). *The adoption of data spaces: Drivers toward federated data sharing*. Hawaii International Conference on System Sciences (HICSS).

Hutterer, A., & Krumay, B. (2025). *Exploring Digital Transformation Opportunities of Data Spaces*. Hawaii International Conference on System Sciences (HICSS).

Jarke, M. (2020). Data sovereignty and the internet of production. *International Conference on Advanced Information Systems Engineering*, 549–558.

Jarke, M., Otto, B., & Ram, S. (2019). Data sovereignty and data space ecosystems. In *Business & Information Systems Engineering* (Vol. 61, Issue 5, pp. 549–550). Springer.

Jussen, I., Schweihoff, J., & Möller, F. (2023). Tensions in Inter-Organizational Data Sharing: Findings from Literature and Practice. *2023 IEEE 25th Conference on Business Informatics (CBI)*, 1–10. https://doi.org/10.1109/CBI58679.2023.10187530

Kagermann, H., & Winter, J. (2018). The second wave of digitalisation: Germany's chance. In S. Mair, D. Messner, & L. Meyer (Eds.), *Germany and the World 2030: What will change, how we must act. Berlin*. Econ Publishers.

Labadie, C., & Legner, C. (2019). Understanding data protection regulations from a data management perspective: A capability-based approach to EU-GDPR. *Proceedings of the 14th International Conference on Wirtschaftsinformatik (2019)*.

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, *15*(4), 251–266. https://doi.org/10.1016/0167-9236(94)00041-2

Marino, J., Camiciotti, L., Cheinasso, F., Olivero, A., & Risso, F. (2023). Enabling Compute and Data Sovereignty with Infrastructure-Level Data Spaces. *Proceedings of the 3rd Eclipse Security, AI, Architecture and Modelling Conference on Cloud to Edge Continuum*, 77–85. https://doi.org/10.1145/3624486.3624509

Mertens, C., & Kuster, A. (2024). *The Data Spaces Radar*. International Data Spaces Association. https://internationaldataspaces.org/wp-content/uploads/dlm_uploads/The-Data-Spaces-Radar-Version-4.pdf

Mertens, C., Verbrugge, S., Herregodts, A.-L., & Kraemer, P. (2024). *Data Spaces Business Models* (Version Version 1.0). Zenodo. https://doi.org/10.5281/ZENODO.14101303

Mission KI. (2025). *Enhancing Data Ecosystem Connectivity: MISSION KI Develops Innovative Dataset Search Engine*. https://mission-ki.de/en/enhancing-data-ecosystem-connectivity-mission-ki-develops-innovative-dataset-search-engine

Möller, F., Jussen, I., Springer, V., Gieß, A., Schweihoff, J. C., Gelhaar, J., Guggenberger, T., & Otto, B. (2024). Industrial data ecosystems and data spaces. *Electronic Markets*, *34*(1), 41. https://doi.org/10.1007/s12525-024-00724-0

Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research*, *40*(2), 120–123.

Moschko, L., Blazevic, V., & Piller, F. T. (2023). Data Sovereignty as Organisational Competence for Digital Innovation in Industrial Incumbents. *Academy of Management Proceedings*, *2023*(1), 19081. https://doi.org/10.5465/AMPROC.2023.19081abstract

Nagel, L., & Lycklama, D. (2021). *Design Principles for Data Spaces—Position Paper* (Version 1.0). Zenodo. https://doi.org/10.5281/ZENODO.5105744

Oliveira, T., Thomas, M., & Espadanal, M. (2014). Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors. *Information & Management*, *51*(5), 497–510. https://doi.org/10.1016/j.im.2014.03.006

Opriel, S., Möller, F., Strobel, G., & Otto, B. (2024). Data Sovereignty in Inter-organizational Information Systems: Findings from Demand and Capacity Management in the Automotive Industry. *Business & Information Systems Engineering*.

Otto, B. (2022a). A federated infrastructure for European data spaces. *Communications of the ACM*, *65*(4), 44–45. https://doi.org/10.1145/3512341

Otto, B. (2022b). The evolution of data spaces. In *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage* (pp. 3–15). Springer International Publishing Cham.

Otto, B., & Jarke, M. (2019). Designing a multi-sided data platform: Findings from the International Data Spaces case. *Electronic Markets*, *29*(4), Article 4.

Peffers, K., Tuunanen, T., Gengler, C. E., Rossi, M., Hui, W., Virtanen, V., & Bragge, J. (2020). *Design Science Research Process: A Model for Producing and Presenting Information Systems Research*. https://doi.org/10.48550/ARXIV.2006.02763

Rotgang, A., Duparc, E., Steffen, B., Hövel, S., Mayer, J., & Stodolka, W. (2023). Federal Platform Ecosystems to Counter Monopolists: A Value-Based Vision for the Logistics Industry. *Proceedings of the 56th Hawaii International Conference on System Sciences*.

Sarma, A. D., Dong, X. L., & Halevy, A. Y. (2009). Data modeling in dataspace support platforms. In *Conceptual Modeling: Foundations and Applications* (pp. 122–138). Springer.

Schleimer, A. M., Jahnke, N., & Otto, B. (2023). *Architecture Design Options for Federated Data Spaces*. Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2023.447

Schultze, U., & Avital, M. (2011). Designing interviews to generate rich data for information systems research. *Information and Organization*, *21*(1), 1–16. https://doi.org/10.1016/j.infoandorg.2010.11.001

Schurig, T., Kari, A., & Fürstenau, D. (2024). The Symphony of Orchestrated Participatory Data Space Governance: A Systematic Review. *ECIS 2024 Proceedings. 9*. Thirty-Second European Conference on Information Systems (ECIS 2024).

Singh, M. (2013). A framework for data modeling and querying dataspace systems. *Seventh International Conference on Data Mining and Warehousing (ICDMW)*.

Sonnenberg, C., & Vom Brocke, J. (2012). Evaluations in the Science of the Artificial – Reconsidering the Build-Evaluate Pattern in Design Science Research. In K. Peffers, M. Rothenberger, & B. Kuechler (Eds.), *Design Science Research in Information Systems. Advances in Theory and Practice* (Vol. 7286, pp. 381–397). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-29863-9_28

Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In *Handbook of qualitative research* (pp. 273–285). Sage Publications, Inc.

Ubacht, J., Schmid, J., Rietveld, E., Lennartz, J., Rukanova, B., & Tan, Y.-H. (2023). *Data Sharing Arrangements for Monitoring in the EU Circular Economy: The Case of CBAM and Steel Import for the EU Automotive Sector*. EGOV-CeDeM-ePart conference. https://ceur-ws.org/Vol-3449/paper17.pdf

vom Brocke, J., & Buddendick, C. (2006). Reusable conceptual models–requirements based on the design science research paradigm. *Proceedings of the First International Conference on Design Science Research in Information Systems and Technology (DESRIST)*, 576–604.

von Scherenberg, F., Hellmeier, M., & Otto, B. (2024). Data Sovereignty in Information Systems. *Electronic Markets*, *34*(1), 15. https://doi.org/10.1007/s12525-024-00693-4

2. A METHODOLOGICAL PROCESS MODEL FOR RECOGNIZING DIGITAL TRANSFORMATION OPPORTUNITIES

Manuel Muehlburger
Johannes Kepler University Linz
Manuel.muehlburger@jku.at

Thomas Deutsch
Johannes Kepler University Linz
Deutsch-thomas@gmx.at

Barbara Krumay
Johannes Kepler University Linz
Barbara.Krumay@jku.at

Abstract

Organizations' ability to identify the potential of digital technologies represents a crucial challenge for successfully managing their digital transformation (DT). To tackle this challenge processes are needed that integrate the technological, domain specific and economic expertise within an organization for identifying digital transformation opportunities. Currently, however, approaches to support these recognition processes are scarcely described in information systems engineering. To close this gap an ensemble artefact - namely a methodological process model that enables the recognition of DT opportunities in collaborative settings - has been developed by applying action design research (ADR), involving researchers and practitioners. The final artefact, developed in four iterations, combines a taxonomy and a procedural model. The taxonomy conceptualizes DT as a combination of context-specific instances of three meta-elements (representation, technology and effect). In addition, the procedural model of activities and states guides the recognition process. The artefact's evaluation regarding usability, productivity, validity as well as overall acceptance indicates that this approach is a viable solution for organizations aiming to bridge the silos between domain, business and technology expertise in the context of DT opportunity recognition.

Keywords

Digital Transformation, Opportunity Recognition, Process Model.

1. Introduction

DT has emerged as a critical factor for organizations' ability to be successful, compelling them to strategically recognize opportunities for their business when adopting digital technologies (Bharadwaj et al., 2013; Fitzgerald et al., 2014; Vial, 2019). Early DT focused on adopting new possibilities, but organizations learned that technology adoption fails when organizational capabilities don't fit (Blanka et al., 2022; Vial, 2019). Therefore, approaches that consider the current situation of the organizations including the business model (Berman, 2012) and the entrepreneurial competencies within the organizations (Vargas-Halabí et al., 2017) have been developed. Still, DT in a company depends on many different aspects, among them the ability to recognize opportunities from a more holistic point of view, beyond simply adopting the next digital technology. Current frameworks at hand integrate such a holistic approaches (Muehlburger & Krumay, 2023; Mühlburger et al., 2020), by focusing on the organizational,

economic and technological aspects. However, the frameworks lack operationalization, to become applicable in organizations. To close this gap, this study aims at developing a methodological process model that reflects an engineering approach to enable the recognition of DT opportunities in collaborative settings. Based on the existing DT opportunity recognition (DTOR) framework (Muehlburger et al., 2020), a methodological procedure model is developed based on action design research (ADR) (Mullarkey & Hevner, 2019; Sein et al., 2011), involving a case company providing software services to the tourism sector. By going through four iterations, the artefact has been evaluated and improved. The final artefact combines a taxonomy and a procedural model. The taxonomy conceptualizes DT as a combination of context-specific instances of three meta-elements (representation, technology and effect). In addition, the procedural model of activities and states guides the recognition process.

The remainder of this paper is structured as follows. First, we provide a brief overview on the current state of the field regarding DT and DT opportunity recognition. Next, we show how the method was applied and describe the results of our iterative development process. Finally, we discuss the implications of our findings, limitations of our approach, and directions for further research.

1.1. Digital Transformation

Research investigating digitalization and DT has gained significant attention recently (Vial, 2019). DT has profoundly impacted organizations due to the rapid emergence of new market entrants and substantial changes in business models (Verhoef et al., 2021). The continuously increasing adoption of digital technologies by organizations contributes to market turbulence and challenges traditional determinants of organizational performance (Nan & Tanriverdi, 2017). Organizational competitiveness now frequently depends on how quickly and effectively they adopt digital technologies to drive transformation (Verhoef et al., 2021). However, the term "DT" encompasses various concepts (Vial, 2019) and is used ambiguously in research an practice (Riedl et al., 2024). DT has been defined as "a process that aims to improve an entity by triggering significant changes to its properties through combinations of information, computing, communication, and connectivity technologies" (Vial, 2019). DT represents a process rather than a state, built on digital technologies related to disruptions at various organizational levels (Vial, 2019). By implementing digital technologies, organizations may transform their value creation approaches, necessitating structural changes and overcoming barriers, resulting in both positive and negative outcomes (Vial, 2019). To distinguish DT from other forms of technology-induced organizational change, research has established that DT requires specific capabilities (Blanka et al., 2022) and produces certain transformative outcomes (e.g., (Gong & Ribiere, 2021; Wessel et al., 2021)). This understanding has led to research exploring entrepreneurial aspects of DT and the opportunities arising from it (Kreuzer et al., 2022). An ongoing academic discussion addresses the conceptual nature of organizational DT (Chen & King, 2022; Riedl et al., 2017). Most recently, representation (what changes?), technology, and effect (outcome of DT) have been identified as three context-independent conceptual elements (Muehlburger & Krumay, 2023).

1.2 Digital Transformation Opportunity Recognition

As organizations face increasing pressure to maintain agility and innovativeness to remain competitive, they must continuously evaluate new technologies and assess their organizational fit. In entrepreneurship research, opportunity recognition has been extensively investigated (Baron, 2006; George et al., 2016). Research has examined individuals' ability to recognize business opportunities based on cognitive frameworks and prior experience (Baron, 2006). At

the organizational level, however, opportunity recognition has been attributed to organizational learning processes (Lumpkin & Liechtenstein, 2008).

DT specific opportunity recognition, by comparison, has received considerably less attention. Drawing on theoretical foundations from entrepreneurial literature, recent work has illustrated the effects of digital technology on opportunity recognition, demonstrating that digital technology impacts resources, actors, and markets (Kreuzer et al., 2022). This suggests that digital technology and DT directly influence how opportunities are recognized (Kreuzer et al., 2022), yet significant gaps remain in developing applicable methods to identify opportunities arising from DT. An approach addressing this gap has been proposed based on a design-oriented perspective of DT and digital innovation (Wiesböck & Hess, 2019), resulting in a descriptive framework for recognizing technology-induced opportunities (Muehlburger et al., 2020). The DTOR framework defines three context-independent elements that enable a holistic approach to opportunity recognition (Muehlburger et al., 2020). Organizational representations are models of an organization or its aspects, including business ecosystems, business models, customer journeys, products, services, processes, and data. Technology lenses represent internal representations of technological constructs along function, behavior, and structure dimensions, derived from the function-behavior-structure framework (Gero & Kannengiesser, 2004). Function refers to automation, connectivity, and data availability capabilities; behavior encompasses smart manufacturing and digital platform paradigms; and structure includes physical technologies, software technologies, working techniques, and composite technologies. Digital transformation opportunities (DTOs) represent potential technology-induced changes to organizational value creation paths, categorized as either supportive opportunities (enhancing quality, efficiency, insight, flexibility, accessibility) or transformative opportunities (altering business models, services, products, or markets). The framework conceptualizes DT opportunity recognition as the systematic pairing of organizational representations with technology lenses to identify relevant DTOs. However, this framework presents the process from a high-level perspective that is not directly applicable in information systems engineering practice. Our study aims to address this gap by operationalizing the DTOR framework.

2 Methodological Approach

In this study, we apply Action Design Research (ADR) because practitioner involvement directly contributes to achieving our research goals (Sein et al., 2011). While the widely-used Design Science Research (DSR) approach (Hevner, 2007; Peffers et al., 2007) typically separates practitioners from the development and shaping phase of an artifact, ADR deliberately bridges the gap between researchers and practitioners who experience the problem firsthand (Sein et al., 2011). ADR integrates concepts from both DSR and Action Research, aiming to "generate prescriptive design knowledge through building and evaluating IT artifacts in an organisation setting" (Sein et al., 2011). This methodology addresses organizational problems through intervention and evaluation while simultaneously constructing and refining IT artifacts based on the specific challenges encountered.

ADR builds upon existing DSR frameworks and synthesizes six commonly applied steps (Peffers et al., 2007) into four stages with corresponding principles (P1–P7): Stage 1: Problem Formulation (P1: Practice-Inspired Research, P2: Theory-Ingrained Artifact); Stage 2: BIE - Building, Intervention and Evaluation (P3: Reciprocal Shaping, P4: Mutually Influential Roles, P5: Authentic and Concurrent Evaluation); Stage 3: Reflection and Learning (P6: Guided Emergence); and Stage 4: Formalization of Learning (P7: Generalized Outcomes) (Sein et al., 2011). Mullarkey and Hevner (2019) further developed the ADR approach by proposing a cyclical process with five steps: Problem Formulation/Planning (P), Artifact creation (A),

Evaluation (E), Reflection (R), and Learning (L). This cycle incorporates the four stages previously identified (Sein et al., 2011) while integrating the aforementioned principles. While steps P, E, R, and L closely align with the original stages, Artifact creation (A) has been added as a distinct step with an additional principle (P8: Abstraction). Building on this foundation, Mullarkey and Hevner (2019) developed a comprehensive model connecting different ADR cycles across four stages (Diagnosis, Design, Implementation, Evolution), each associated with specific research entry points (see Fig. 1). For this study, we employ the ADR process model developed by Mullarkey and Hevner (2019) to guide our artifact development.

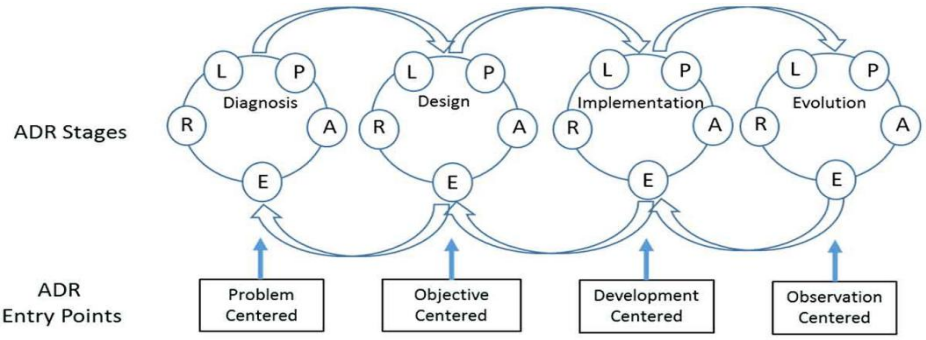

Fig 1: ADR process model (Mullarkey & Hevner, 2019).

2.1 Artefact Development

To design and develop the artifact, we employed various methods for collecting data from both academic sources and practitioners who were directly involved in the design process. Practitioner involvement was essential because the specialized knowledge and perspectives of these stakeholders were required to properly inform the study. The research was conducted in collaboration with a case company from the tourism sector—an industry facing significant challenges due to dynamic and rapidly changing market conditions driven by emerging technologies. The case company is a small software agency providing B2B services primarily in the central European tourism market. We structured our development process using the five steps of the ADR cycle (Mullarkey & Hevner, 2019) as guiding principles for our iterative approach. We adopted an objective-centered approach (beginning with the Design stage) because the problem had already been conceptualized in prior studies (Muehlburger et al., 2020). Our process began with problem formulation, followed by developing a preliminary artifact based on literature, unstructured observations, and interviews within a single case company in the tourism software sector, limiting our evaluation to one organizational context.

The final artifact resulted from four ADR cycles in which it was repeatedly evaluated in workshop settings. During these workshops, participants were presented with the artifact in its current form and asked to apply it while discussing their experiences. We observed their behaviour throughout this process to gather additional insights, combining structured data collection with unstructured observations of participants' interactions and difficulties with the model. After each workshop, participants completed a post-hoc questionnaire designed to assess specific evaluation metrics: usability, productivity, validity, and overall acceptance. These metrics were selected from a broader set of potential criteria through preliminary interviews, prioritizing those most relevant to practitioners. The researcher actively moderated each workshop, with the moderation approach evolving across iterations—from highly directive in early workshops when participants struggled with conceptual understanding, to a more hands-off approach in later iterations as the improved model required less intervention, allowing assessment of whether practitioners could independently apply the process. The workshop participants included three employees from project management (workshop 1), three from innovation management (workshop 2), and three from customer service (workshop 3). The fourth

workshop involved the same three participants from workshop one, enabling direct comparison of their experiences with both the initial and final versions of the artifact and providing valuable insights into how specific improvements addressed their original challenges. Time constraints sometimes limited exploration of all potential opportunities or lens combinations, reflecting real-world implementation challenges. Through these four iterations we employed various approaches to develop and evaluate the artifact, incorporating feedback at each stage. Figure 2 illustrates key changes across iterations, though space constraints prevent showing the comprehensive evolution where each modification directly addressed specific participant challenges observed in workshops.

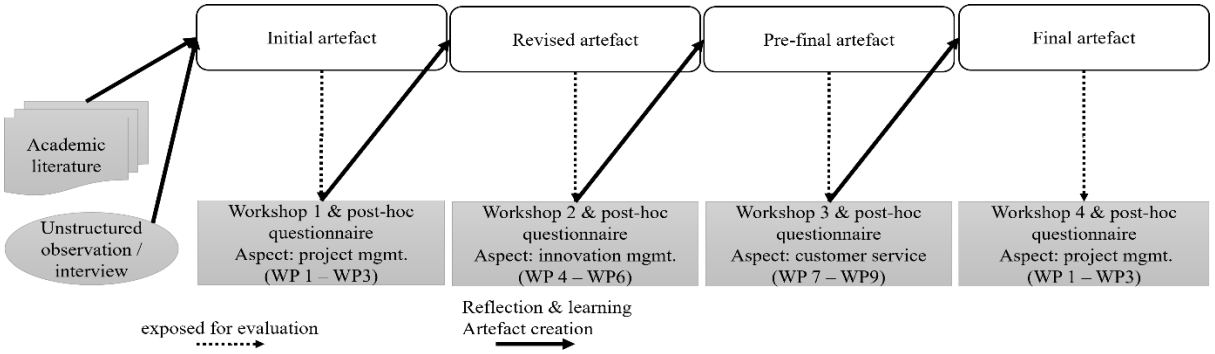

Fig 2: Documentation of development steps.

3 Results

The results of the ADR process are presented chronologically through four iterations, demonstrating the evolutionary development of the final artifact and highlighting the changes made at each stage based on stakeholder feedback and evaluation.

3.1 Initial Artefact

In the first iteration, we focused on understanding how an artifact might address the identified problem based on existing literature (Muehlburger et al., 2020). Our primary objective was to sketch a process that would support organizations in identifying DT opportunities. Additionally, we aimed to provide the necessary tools and descriptions to establish a shared understanding among stakeholders. This iteration was pivotal as it represented the transition from conceptual framework to applicable design—a common progression in information systems engineering. The iteration had a normative character, providing an initial design template for the process. Following design principles (Mullarkey & Hevner, 2019), this phase played a critical role in development by bridging the conceptual framework (Muehlburger et al., 2020) with an implementable, testable design procedure. Simultaneously, it established guidelines that systems engineers could follow in practice.

Beyond literature review, the initial artifact development was informed by an unstructured interview with an employee responsible for identifying innovative technologies at the case company. The interviewee described the company's current discovery process as primarily "gap-focused" and experience-based, without any structured methodology—essentially relying on coincidental discoveries. This highlighted the need for a balance between methodological rigor and accessibility, as many participants would later struggle with selecting appropriate technological lenses, particularly those with less technical backgrounds. Both the interview and unstructured observations of company practices revealed that opportunity recognition efforts were almost exclusively oriented toward efficiency improvements and cost reduction. Opportunities related to quality enhancements or service improvements received minimal attention. The interviewee also noted resistance within the organization toward frequent changes,

particularly regarding the implementation of new technologies. For successful adoption, the benefits of new technologies needed to be clearly articulated; otherwise, acceptance and usage remained low.

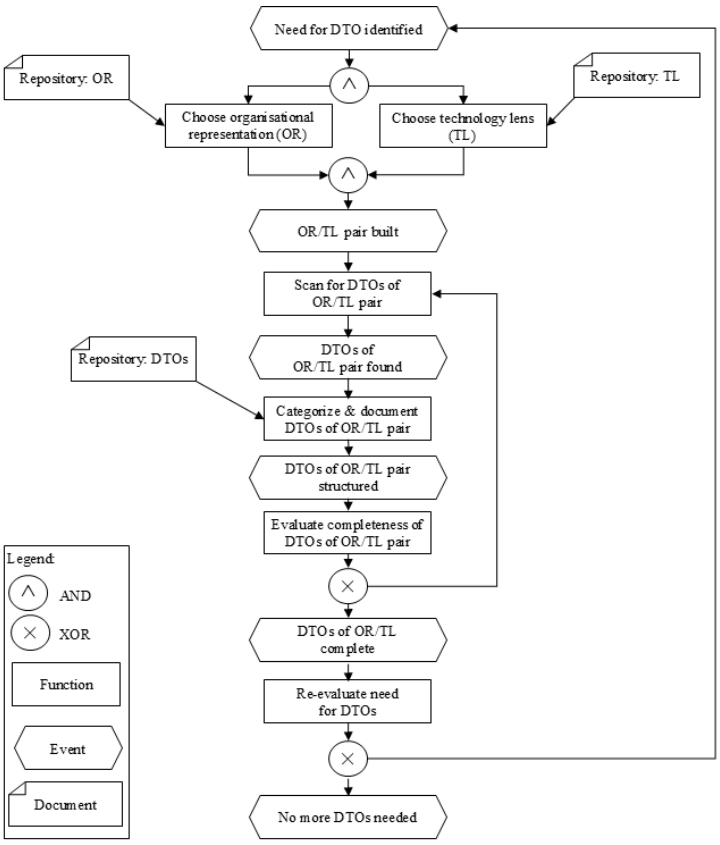

Fig 3 Initial Artefact (Procedural Model)

The initial artifact addresses the identified shortcomings while incorporating insights from literature, interviews, and observations. We based our design on the extended event-driven process chain (eEPC) modeling approach, which is recognized for its accessibility to users without specialized process modeling training (Felder & Brent, 2005). In our model, the process begins when a need for DT opportunities (DTOs) is identified. This need may emerge internally within the company (e.g., to improve efficiency in a specific process) or externally from market conditions (e.g., competitors gaining competitive advantage). The process continues with the selection of both an organizational representation (OR) and a technology lens (TL) (Muehlburger et al., 2020) to create an OR/TL pair. This pairing reflects the fundamental assumption that DTOs must address both technological and organizational dimensions (Muehlburger et al., 2020). For organizational representations, we utilized repositories developed in a previous study (Muehlburger & Krumay, 2023), comprising seven categories: business ecosystem, business model, customer journey, product, services, process, and data. Similarly, the technology lens drew from repositories following the function-behavior-structure framework ((Gero & Kannengiesser, 2004), categorized by function (automation, connectivity, data availability), behavior (function, smart manufacturing paradigm, digital platform paradigm), and structure (physical technology, software technology, working techniques, composite technology). The scanning for DTOs is conducted using the defined OR/TL pair, assuming that at least one opportunity exists. Once identified, DTOs are categorized into two groups: supportive opportunities (quality, efficiency, insight, flexibility, accessibility) and transformative opportunities (business model, service, product, market). The process then evaluates the completeness of the identified DTOs—whether a sufficient number has been found for the

current OR/TL pair. If deemed incomplete, the process cycles back to scanning for additional DTOs with the same pair. If complete, the overall need for additional DTOs is reassessed, potentially initiating a new cycle with different OR/TL pairs or concluding the process if no further DTOs are needed. Figure 3 illustrates this initial procedural model.

In summary, the initial artifact (procedural model) expanded beyond the case company's current focus on efficiency-enhancing technologies to encompass a broader range of opportunities arising from new technologies. We deliberately use the term "procedural model" to emphasize the step-by-step nature of our approach. The repositories for pairing organizational representations (OR) and technology lenses (TL), as well as those for categorizing DT opportunities (DTOs), are equally critical components of the overall artifact.

To evaluate the initial artifact, we conducted a workshop with three project management employees from the case company (workshop 1). Participants were provided with predefined settings, including specified DTO needs, repositories (OR and TL), and a categorization scheme for structuring DTOs. Workshop observations revealed that while participants generally navigated the process without difficulty, they struggled with specific aspects. Most notably, they had trouble interpreting the "AND" relationship. Participants with non-technical backgrounds particularly struggled with understanding and applying the technological lens concept, finding it highly abstract and difficult to translate into practical application—a challenge that persisted throughout all iterations despite attempts to provide more concrete examples. Participants found the repositories overwhelming and requested additional contextual information to facilitate selection. The technology lens repository, in particular, provided high-level definitions that participants found too abstract; they requested concrete examples of technologies (e.g., explaining "composite technologies" with examples like 3D-printing). The categorization of DTOs also prompted significant discussion, as participants could not readily discern its practical benefit. Although participants found the process generally straightforward, they struggled to assess when the "completeness" criterion had been satisfied. Additionally, observations indicated that the procedural model implicitly contained distinct phases that were not explicitly represented in the visual model. The post-hoc questionnaire revealed that participants found the initial artifact relatively easy to use and believed it improved productivity. However, they identified several challenges: difficulty understanding the pairing process, lack of clear purpose, and need for more concrete examples of technologies. While validity and acceptance ratings were high, participants suggested improvements to clarify when the process ends ("When is the need fulfilled") and the model's overall purpose ("Why should it be used?"). All participants emphasized the need for comprehensive instructions before beginning the process, particularly when working with the repositories. Finally, they expressed concern about the high abstraction level, noting that effective use required the ability to "think on an abstract level." Based on this valuable feedback, we reflected on necessary changes and proceeded to develop subsequent iterations of the artifact.

3.2 Revised Artefact

The next two iterations followed the five steps of the ADR cycle (Sein et al., 2011), beginning with problem formulation based on feedback from previous evaluations. To address the identified issues, we made several key modifications to the artifact. First, to enhance clarity regarding DTO needs, we added a specific process step for formally articulating these needs. Second, we addressed concerns about the OR/TL pairing by consolidating it into a single, more intuitive process step. Third, to resolve difficulties with completeness assessment, we introduced additional process steps and feedback loops that clearly differentiate between identifying all DTOs for a specific pair versus exploring all potential pairs. Fourth, we incorporated the concept of "satisfaction" ("Re-evaluate if need for DTOs is satisfied") to provide clearer guidance on

determining when the process should conclude. Fifth, we visually distinguished the different phases (Initiation, Application, Evaluation) using color coding to make the previously implicit process stages more explicit (see Fig. 3). Finally, we enhanced the technology lens repository by adding concrete examples of actual technologies to address participants' requests for more tangible illustrations.

We evaluated this updated artifact through a second workshop with three innovation management employees from the case company. The workshop observations and post-hoc questionnaire results indicated that while the second iteration had successfully addressed some previously identified issues (particularly by adding concrete technology examples to the TL repository), several challenges remained unresolved, resulting in decreased usability ratings.
Participants continued to struggle with understanding the practical benefits of DTO categorization. They provided specific recommendations, emphasizing that the color-coded phase marking was unnecessary and requested a handbook or clear guidelines for each process step. They highlighted the importance of focusing on the target group and suggested using the term "activity" instead of "function." Additionally, they proposed clearer explanations of process elements such as events and functions, and suggested either clear color-coding or none. Lastly, they questioned the use of the term "repository," finding it too technical.

3.3 Pre-final and Final Artefact

Reflecting on and learning from workshop 2 results and post-hoc questionnaire feedback, we conducted another ADR cycle (Sein et al., 2011). Our primary focus for this iteration was improving the visualization of the procedural model and refining terminology to address participant concerns. We made significant changes to the visual representation, introducing numbered steps, consistent color-coding between model elements and referenced materials, and adding intuitive icons to quickly distinguish between activities and statuses—visual enhancements that proved crucial for improving usability according to participant feedback. A major enhancement was the transformation of the existing repositories into a comprehensive taxonomy that functioned as the requested handbook. This component of the artifact provides clearly articulated guidelines for each step in the procedural model. We implemented several user-friendly features in this taxonomy: (1) numbered steps corresponding to the procedural model for easy cross-reference; (2) explanations of each step's relevance to the overall process; (3) practical examples illustrating how to execute each step; (4) clear delineation of inputs and outputs for each activity; (5) reorganization of repositories as accessible lists (avoiding the technical term "repository"); and (6) detailed explanations for each list entry. Additionally, we included purpose statements to help users understand the rationale behind each component.
For the third evaluation, we presented the revised artifact (procedural model shown in Fig. 5 and accompanying taxonomy) to three customer service employees during workshop 3. Given the decreased usability ratings from the previous workshop, this evaluation specifically focused on assessing usability improvements. We were particularly interested in evaluating how effectively the taxonomy (handbook) complemented the procedural model. To test this integration, the workshop facilitator instructed participants to consult the handbook first whenever questions arose during the process application. The workshop 3 evaluation revealed that most remaining issues centered on terminology within the taxonomy (handbook) rather than structural concerns. Participants found the procedural model itself easy to follow, with clear wording and terminology. Based on reflection and learning from this feedback, we conducted a final ADR cycle (Sein et al., 2011) focused specifically on refining the taxonomy's terminology while maintaining the procedural model unchanged. For the final evaluation (workshop 4), we presented both components of the artifact—the procedural model and the refined taxonomy—to the same participants who had attended workshop 1. This approach allowed us to gather insights

from participants who had experienced the artifact's complete evolution from initial conception to final form. The comprehensive feedback confirmed the general effectiveness of both components, validating our iterative development approach. While most issues identified throughout the evaluation process were addressed, participants across all iterations consistently questioned the value of the categorization step. Despite attempts to clarify its purpose, this remained an unresolved concern even in the final evaluation, suggesting a fundamental tension between theoretical completeness and practical utility in this aspect of the model. A particularly significant finding from the final evaluation was that participants across different organizational functions expressed confidence in their ability to apply the model independently. By workshop 3, participants without prior experience with process models stated they "would feel comfortable applying the artifact in a new workshop group unfamiliar with the process as a moderator." Multiple groups had already scheduled additional workshops to apply the model to other business cases without researcher intervention, indicating genuine organizational adoption beyond the evaluation context. This unprompted implementation represents strong validation of the artifact's practical utility and accessibility. To evaluate the effectiveness of model improvements across iterations, we tracked the number of DTOs identified during each workshop. The results showed a clear progression: workshop 1 yielded only 3 DTOs, workshop 2 produced 13 DTOs, workshop 3 generated 29 DTOs, and workshop 4 identified 22 DTOs. This quantitative improvement demonstrated the artifact's increasing effectiveness in facilitating opportunity recognition as design improvements were implemented.

4 Discussion, Limitations and Future Research

Organizations today face increasing pressure for DT to maintain competitive advantage, creating an urgent need for structured approaches to recognize digital opportunities. This study addresses this gap by developing a methodological process model comprising two integrated components: a procedural model and a supporting taxonomy. Through four iterations following the Action Design Research methodology (Sein et al., 2011), we created an artifact that builds upon prior research and incorporates practical insights from industry stakeholders. The resulting model enables companies to balance technological and organizational dimensions when identifying opportunities aligned with their needs. The primary contribution is connecting business, technology, and domain-specific knowledge.

Our model consists of two complementary components: the procedural model and the taxonomy (handbook). These elements evolved significantly throughout our iterative research process, aligning with established engineering approaches found in other frameworks. The taxonomy proved valuable for ensuring clarity and accessibility across stakeholders with varied backgrounds and expertise levels. Our model addresses key challenges in DT opportunity recognition. First, it expands the focus beyond efficiency improvements to encompass transformative opportunities. Second, it provides a structured methodology for what was previously an ad hoc process. Third, it creates a common language and conceptual framework that facilitates communication across different organizational functions.

The evaluation results from our workshops demonstrate that the model balances structure with flexibility, providing guidance while allowing customization to specific contexts. Participants consistently highlighted the value of a systematic approach that accommodates their environment and objectives. The procedural model guides users through the recognition process, while the taxonomy provides the necessary depth of information to support decision-making at each step. A distinctive strength of the model is its practical adaptability to organizational constraints. Participants valued the ability to pause and resume the process across multiple sessions without losing coherence—a critical feature for implementation in time-constrained environments. This

flexibility enabled practitioners to integrate the model into existing workflows, enhancing implementation potential.

From a theoretical perspective, this study contributes to DT literature by operationalizing conceptual frameworks. By translating abstract elements into concrete procedures, we bridge the gap between theory and practice. This demonstrates how conceptual frameworks can be transformed into tools through collaborative design processes involving researchers and practitioners. Practically, our model offers organizations a tool for systematically identifying DT opportunities beyond the obvious. By considering multiple organizational representations and technology lenses, it encourages exploration of overlooked possibilities and potential synergies. This approach is valuable given the complexity and high stakes of DT initiatives.

4.1 Limitations and Future Research:

While our study provides valuable insights, several limitations must be acknowledged, our model was developed and evaluated within a single case company in the tourism sector, which may limit its generalizability to other industries or organizational contexts. Second, our evaluation focused primarily on the model's usability and acceptance rather than its long-term effectiveness in identifying valuable DT opportunities. Third, the model in its current form requires facilitation by individuals with some familiarity with process modelling and digital technologies. Further development could explore simplified versions for organizations with limited technical expertise or resources. Fourth research on how this approach might be integrated with other DT frameworks and tools, particularly those focused on implementation rather than opportunity recognition. Additionally, exploring how emerging technologies such as artificial intelligence might be incorporated into the opportunity recognition process represents a promising direction for extending this work.

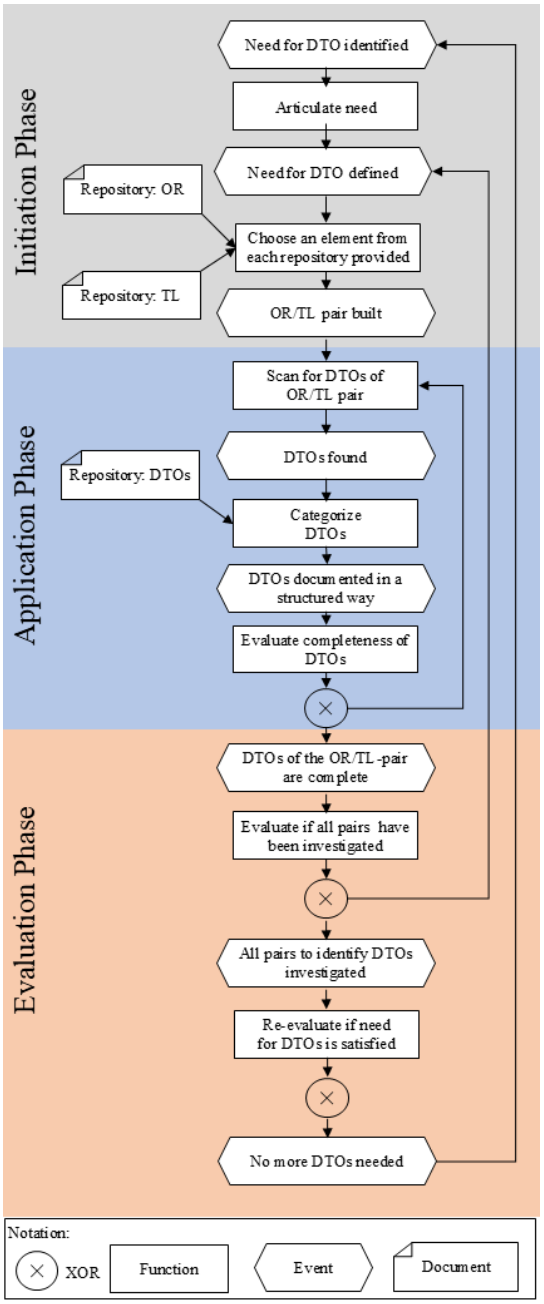

Fig 4: Revised Artefact (Procedural Model), 2nd Iteration

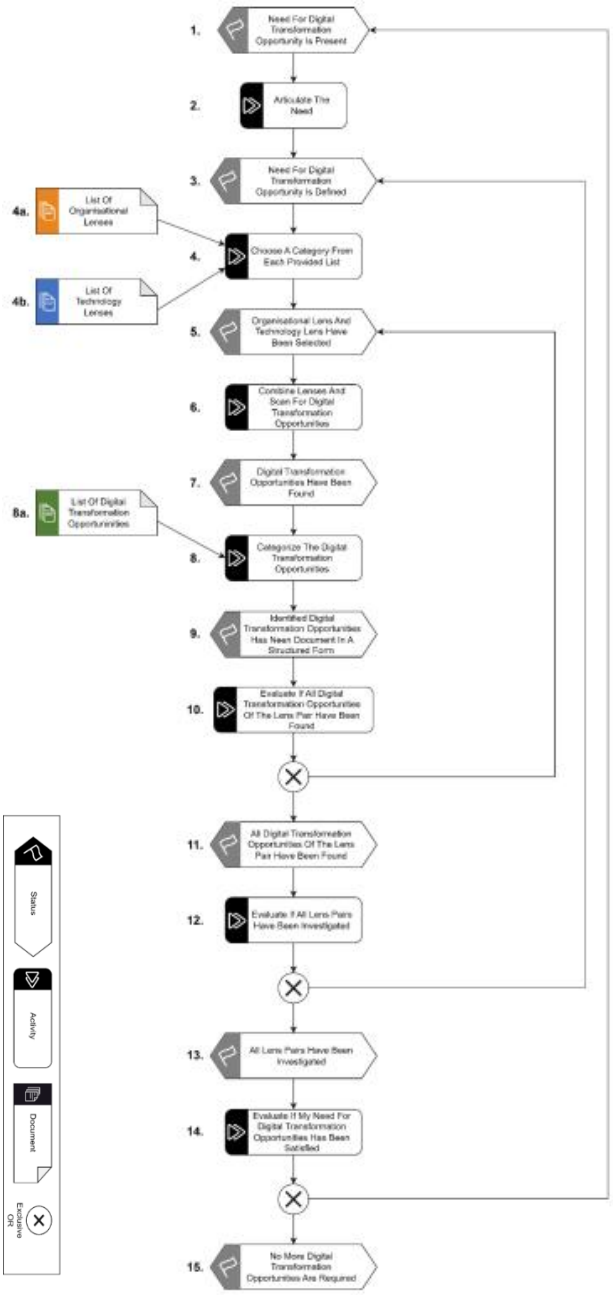

Fig. 5. Final Artefact (Procedural Model)

References

Baron, R. A. (2006). Opportunity Recognition as Pattern Recognition: How Entrepreneurs "Connect the Dots" to Identify New Business Opportunities. *Academy of Management Perspectives*, *20*(1), 104–119.

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. (2013). Digital Business Strategy: Toward a Next Generation of Insights. *MIS Quarterly*, *37*(2), 471–482.

Blanka, C., Krumay, B., & Rueckel, D. (2022). The interplay of digital transformation and employee competency: A design science approach. *Technological Forecasting and Social Change*, *178*, 121575.

Chen, S., & King, J. (2022). Policy and Imprecise Concepts: The Case of Digital Transformation. *Journal of the Association for Information Systems*, *22*(2), 401–407.

Felder, R. M., & Brent, R. (2005). Understanding Student Differences. *Journal of Engineering Education*, *94*(1), 57–72.

Fitzgerald, M., Kruschwitz, N., Bonnet, D., & Welch, M. (2014). Embracing digital technology: A new strategic imperative. *MIT sloan management review*, *55*(2), 1–16.

George, N. M., Parida, V., Lahti, T., & Wincent, J. (2016). A systematic literature review of entrepreneurial opportunity recognition: Insights on influencing factors. *International Entrepreneurship and Management Journal*, *12*(2), 309–350.

Gero, J. S., & Kannengiesser, U. (2004). The situated function–behaviour–structure framework. *Design Studies*, *25*(4), 373–391.

Gong, C., & Ribiere, V. (2021). Developing a unified definition of digital transformation. *Technovation*, *102*, 102217.

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems*, *19*(2), 4.

Kreuzer, T., Lindenthal, A.-K., Oberländer, A. M., & Röglinger, M. (2022). The Effects of Digital Technology on Opportunity Recognition. *Business & Information Systems Engineering*, *64*(1), 47–67.

Lumpkin, G. T., & Liechtenstein, B. (2008). The role of organizational learning in the opportunity recognition process. In *Entrepreneurial Learning*. Routledge.

Muehlburger, M., Kannengiesser, U., Krumay, B., & Stary, C. (2020). A Framework for Recognizing Digital Transformation Opportunities. *ECIS 2020 Proceedings*. European Conference on Information Systems, A Virtual Conference.

Muehlburger, M., & Krumay, B. (2023). Towards a context-sensitive conceptualisation of digital transformation. *Journal of Information Technology*, 02683962231219514.

Mullarkey, M. T., & Hevner, A. R. (2019). An elaborated action design research process model. *European Journal of Information Systems*, *28*(1), 6–20.

Nan, N., & Tanriverdi, H. (2017). Unifying the Role of IT in Hyperturbulence and Competitive Advantage Via a Multilevel Perspective of IS Strategy. *MIS Quarterly*, *41*(3), 937–958.

Peffers, 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.

Riedl, R., Benlian, A., Hess, T., Stelzer, D., & Sikora, H. (2017). On the Relationship Between Information Management and Digitalization | SpringerLink. *Business & Information Systems Engineering*, *59*, 475–482.

Riedl, R., Stieninger, M., Muehlburger, M., Koch, S., & Hess, T. (2024). What is digital transformation? A survey on the perceptions of decision-makers in business. *Information Systems and E-Business Management*, *22*(1), 61–95.

Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action Design Research. *MIS Quarterly*, *35*(1), 37–56. https://doi.org/10.2307/23043488

Vargas-Halabí, T., Mora-Esquive, R., & Siles, B. (2017). Intrapreneurial competencies: Development and validation of a measurement scale. *European Journal of Management and Business Economics*, *26*(1), 86–111.

Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Qi Dong, J., Fabian, N., & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. *Journal of Business Research*, *122*, 889–901.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*, *28*(2), 118–144.

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.

Wiesböck, F., & Hess, T. (2019). Digital innovations: Embedding in organizations. *Electronic Markets*.

3. ADOPTING DEVSECOPS: A FRAMEWORK FOR IT GOVERNANCE AND CULTURE CHANGE BASED ON A PLAN-DO-CHECK-ACT (PDCA) APPROACH

Bilgin Metin
Bogazici University
bilgin.metin@bogazici.edu.tr

Martin Wynn
University of Gloucestershire
mwynn@glos.ac.uk

Ali Can Keserel
Bogazici University
akeserel@gmail.com

Timur Demir
Bogazici University
tdtimurdemir@gmail.com

Barışcan Güngör
Bogazici University
gbariscan@gmail.com

Abstract

As digital transformation accelerates, organizations increasingly turn to agile software development and deployment practices like DevOps. However, incorporating security into these processes through DevSecOps presents significant challenges, particularly in cultural adaptation and alignment with IT governance. This study explores the challenges of adopting DevSecOps from two crucial perspectives: organizational culture and IT governance. Through a thorough literature review and the development of a conceptual framework, we identify human-related barriers such as resistance to change, lack of awareness, and communication gaps, along with governance-related constraints such as inadequate policies, misalignment of risks, and compliance issues. To tackle these challenges, we propose a Plan-Do-Check-Act (PDCA) implementation model that provides a practical approach for transforming organizational culture and improving IT governance. This approach aims to bridge the gap between development, security, and operations while aligning with strategic business objectives. Future research in this field could include empirically validating the model through case studies.

Keywords

DevOps, DevSecOps, Cybersecurity,

1 Introduction

Cybersecurity is more crucial than ever as digitalization is becoming increasingly evident within business and across wider society. Cyberattacks and security concerns emphasize the importance of security in software development and operations (Nagasundari *et al.*, 2025). In the context of the business IT function, Ahmed and Francis (2019) indicate that using the method of DevOps not only enables faster, better collaboration and communication for team building, testing, and releasing software but also brings great agility, flexibility, and quality, as well as cost efficiencies. DevOps is described as merging software development and operations to align their priorities, enabling them to work together to achieve a successful project (Cois, 2014). DevSecOps integrates a security layer into DevOps, recognizing the increasing importance of cybersecurity. The aim is to embed modern security practices in IT development and operations by involving security from the beginning (Myrbakken & Colomo-Palacios, 2017).

DevOps is divided into four processes: culture change, automation, measurement, and sharing (CAMS) (Tomas & Huang, 2019). All these are important, but culture change is crucial to establishing DevOps and DevSecOps in an organization. DevSecOps needs three different areas

to work together to build a secure software project, for example (Sánchez-Gordón & Colomo-Palacios, 2020). Organizations must first create an appropriate security culture within which technical procedures can be more effectively communicated and adhered to.

A survey conducted by Synk (2020) emphasizes cultural issues with DevSecOps. It indicates that security is perceived as an activity that slows down the business and overall software delivery, and 29% of all organizations positioned at the highest level of security integration still feel that security teams and delivery teams encounter a lot of friction, even when collaborating together. Another factor that affects both culture change, and DevSecOps adoption overall, is IT governance, which impacts a company's internal processes and culture. To successfully adopt DevSecOps, alignment with business and IT goals is necessary. Thus, this research aims to study how culture and IT governance influence the adoption of DevSecOps. Based on these insights, a PDCA-based framework of actionable measures is developed to address the challenges of adoption and encourage successful implementation. More specifically, this study answers the following research questions (RQs):

RQ1. What are the main challenges in a business's adoption of DevSecOps?
RQ2. What are the effects of culture on a business's adoption of DevSecOps?
RQ3. What are the effects of IT governance on a business's adoption of DevSecOps?
RQ4. What measures or framework can be adopted as regards culture and IT governance to overcome DevSecOps adoption challenges?

This study undertakes a narrative synthesis to explore the organizational and governance-related challenges in DevSecOps adoption. Popay et al. (2006) note that a "narrative synthesis relies on the use of words and text to summarise and explain the findings of multiple studies, rather than relying solely on numerical data". One particular strength of a narrative synthesis is its flexibility and adaptability, which allows the researcher to accommodate a range of approaches and techniques in locating relevant sources and to assess material from a wide range of different types of literature. Petticrew and Roberts (2006) observe that a "narrative synthesis is particularly valuable when studies vary in design, population, or intervention, as it enables researchers to construct a coherent story from fragmented evidence". In this instance, the selection of literature was guided by the authors' experience and focused on key academic databases, such as Web of Science (WoS), Scopus, and IEEE Xplore, to gather foundational theories and empirical findings. This study synthesizes significant contributions from the academic literature and incorporates selected grey literature, including industry reports from sources like Synk, as well as technical guides from MITRE, GSA, and Interpol. This approach ensures both practical relevance and contemporary context, helping to develop a conceptual understanding and propose a practical framework. To ensure our analysis remains relevant, especially with the rapid evolution of DevSecOps practices, the literature search focused on publications from the last 5-7 years (approximately 2018-2025). In the remainder of this paper, Section 2 presents a literature review that establishes the theoretical foundation of DevSecOps, covering both its cultural and governance aspects. Section 3 introduces a conceptual framework that highlights the key challenges associated with adopting DevSecOps from the perspectives of culture and IT governance. Section 4 proposes a structured Plan-Do-Check-Act (PDCA) approach to systematically address these challenges and promote effective adoption. Section 5 provides a critical discussion of the findings in relation to the RQs and relevant literature. Finally, Section 6 concludes the paper by summarizing the main insights and offering suggestions for future research directions

2 Literature Review

This study includes both technical and business-related concepts from distinct research areas, which are considered systematically in the sections below, in which definitions and theoretical background of each term are given.

2.1 DevOps and Security: DevSecOps

Zhu *et al.* (2016) define DevOps in the context of software development as a group of functions used primarily to make the time between injecting an edit to a system and injecting a change to normal production shorter without compromising overall quality. Lwakatare *et al.* (2019) indicate that, as a concept, the main purpose of DevOps is to handle the problems with software development, release, and operational processes, which are brought about by organizational distinctions between processes. They further claim that, within the IT function, the development and operations departments may have some conflict of objectives and practices as regards agility and stability. This may be evidenced in particular when these departments have to deal with issues such as inefficient data flow and poor test environments (Lwakatare *et al.,* 2019).

Luz *et al.* (2019) developed a model for DevOps adoption, considering the view of the people in charge of the adoption process of DevOps in their workplace, and they also applied the model to a real-world DevOps adoption scenario. The research concludes that the adoption of DevOps is related to a number of main factors, including agility, automation, collaborative culture, continuous measurement, quality assurance, resilience, sharing and transparency, and a collaborative culture, which plays the most important role among them (Luz *et al*., 2019).

In recent years, the global shift to online communication tools and remote working has increased dramatically. Besides technological advances, COVID-19 played a crucial role in rapidly adopting and deploying remote systems and applications for organizations and individuals. Organizations have become increasingly vulnerable to cybercriminals focusing on online systems and connectivity due to this sudden global shift (Interpol Cybercrime Report, 2020). At the same time, many organizations want to speed their adoption of IT-related change by minimizing IT operational overheads and saving time by automating development processes and increasing their deployment of newly created software and features. However, security aspects are a concern - of each line of code and for infrastructure changes. DevOps and Security have thus become strictly interconnected practices, and as a result, a new term, DevSecOps, emerged (Mohan & Othmane, 2016).

DevSecOps can be conceptualized as the set of automated procedures and security practices applied to the development, integration, and deployment pipelines for IT products in organizations by fundamentally strengthening awareness and knowledge and by using appropriate comprehensive tools (Tomas *et al*., 2019). Achieving better quality of secure software requires integrating security principles and practices into the DevOps processes and the software development lifecycle (SDLC) from inception to delivery, including automated and non-automated activities. Automated activities are related to the common software practices in DevOps, such as automation of monitoring, testing, code reviews, security updates, and patches. Besides, non-automated activities, especially in terms of security, include security requirements analysis, design review, configurations and policies, compliance analysis, threat modeling, and risk analysis (Nagasundari *et al.,* 2025). In their study of security practices in DevOps, Rahman *et al.* (2016) found that security teams were actively collaborating with development and operations teams to provide better DevSecOps processes in terms of automated and non-automated activities.

2.2 Challenges in DevSecOps Adoption

Kumar and Goyal (2020) identify the different challenges in DevSecOps adoption from people, process, and technology perspectives, encompassing culture-related challenges such as resistance to change in the organization's culture and people mindset and teams working in isolation without sharing information between them. They further indicated that the selection of Open Source Software (OSS) policy should encourage a collaborative culture to make connections between staff from development, security, and operations to decide on the best open-source tools (Kumar & Goyal, 2020). In their multivocal literature review, Myrbakken and Colomo Palacios (2017) indicated that most problems with DevSecOps adoption are related to speed and agility, organizational changes, tools, and practices. Moreover, they stated that organizational issues can be related to organizational culture, and new procedures for security prevention, detection, and response will be needed (Myrbakken & Colomo Palacios, 2017).

Zaydi and Nassereddine (2020) analyzed the practices that DevOps practitioners have applied to date and summarized the key challenges of DevSecOps adoption. In conclusion, the authors observed that many organizations have limited resources, but a range of challenges need to be addressed to achieve true agility in a DevSecOps approach. These include the need to enhance coordination among several teams for effective facilitation and the insertion of IT security practices into the full software/product lifecycle, including continuous planning, feedback loops, and automation for the whole pipeline to provide end-to-end coverage.

2.3 DevSecOps and Culture

Transitioning to DevSecOps requires a significant cultural shift. Resistance to change and the need for collaboration among development, security, and operations teams can hinder adoption (Lumpatki *et al.,* 2024; Rajapakse *et al.,* 2022). DevSecOps requires three different groups to work together, which traditionally have worked separately in the IT arena: the development team, the operations team, and the security overseers. Organizational changes and improvements are required before technological developments to ensure that these three groups work together. It is necessary to maintain a culture of innovation that consists of openness, transparency, ownership, and accountability to enable these three teams to work together efficiently and reduce organizational-level problems (Mao *et al.,* 2020). The authors examined grey literature and identified internal and external factors impacting DevSecOps implementation and emphasized that there are few studies conducted that identify challenges in DevSecOps implementation. However, cultural resistance, implementation cost, and solidified organizational structures, are identified as the three main challenges in implementing DevSecOps (Mao *et al.,* 2020).

Myrbakken and Colomo-Palacios (2017) emphasize the importance of culture as a principle in DevSecOps in their multivocal literature review. They define culture in DevSecOps differently from DevOps by adding collaboration with security from the start and giving importance to security throughout. Sánchez-Gordón and Colomo-Palacios (2020) emphasized the importance of people when security is integrated into DevOps. The study provides a literature review focused on the cultural aspects and human factors of DevSecOps. Sánchez-Gordón and Colomo-Palacios (2020) identified 13 attributes related to DevSecOps culture: collaboration, sharing knowledge, feedback, continuous improvement mindset, communication, responsibility, trust, experimentation, leadership, commitment and agreement, blamelessness, hiring new personnel, transparency. Their research points to the scarcity of studies in DevSecOps and culture.

Tomas *et al*. (2019) emphasized that unifying software engineering and security teams enables developers to write more secure code and helps the whole team to design systems with security

in mind. The study focused on all the topics of CAMS (culture change, automation, measurement, sharing) and the status of IT companies in terms of security in CAMS terms. They interviewed six engineers working in Norwegian IT companies. The main problems identified in culture were management's prioritization of security, lack of responsibility for security, conflicts between security and development, risk and cost battle, and developers' autonomy (Tomas *et al.,* 2019). Their paper also showed that there is an additional gap between the security and development teams. Tomas *et al*. (2019) identified culture as the foundation of DevSecOps, and their study shows that it is essential for enabling developers to write secure code. Zaydi and Nassereddine (2020) examined applying DevSecOps culture to IT service management (ITSM) in their study. They gathered the most used DevSecOps practices and investigated their usage in ITSM.

Da Veiga and Eloff (2010) developed a framework for assessing information security culture. Their study focuses on different dimensions such as leadership and governance, and security policies. Sas *et al.* (2020) conducted research on measuring the security culture in organizations, and they reviewed existing tools. Orehek and Petrič (2020) also reviewed the scales measuring information security culture.

2.4 IT Governance and DevSecOps

IT governance is an important concept that emphasizes the alignment of IT with business goals (De Haes & Van Grembergen, 2004). Tomas *et al.* (2019) identified reluctance to prioritize security as one of the two factors related to security culture in their study. As their interviews show, management's failure to prioritize security is one of the crucial factors that shape security culture. Alignment and coordination of the organization's goals with IT goals is necessary for DevSecOps to be successfully implemented. In the implementation, Shin *et al.* (2025) proposed a zero-trust framework to enhance security and evaluated its applicability

The relationship between IT governance and DevSecOps has hitherto received little attention in the literature. Two studies are directly related to DevOps and IT governance topics. Fox (2020) examined the relationship between IT governance and DevOps from the Department of Defense's point of view. Wiedemann (2018) analyzed the creation of IT governance systems for building DevOps teams and emphasized that IT governance mechanisms can help establish DevOps in a company.

3 Conceptual Framework

Building upon the literature analysis presented above, a conceptual framework is put forward here to encompass the relationship between the concepts of DevSecOps, culture, and IT governance. The challenges confronting organizations when adopting DevSecOps are represented in Figure 1 from both cultural and IT governance aspects.

The cultural aspects consider both the employee and management perspectives, while IT governance aspects cover structural and operational challenges. A lack of security awareness may result from a lack of knowledge of possible cyber-attacks (Mutlutürk *et al*., 2024). According to Bucena and Kirikova (2017), the challenges regarding the lack of awareness include missing concept maturity, buzzword perception, lack of proper training, and miscommunication. Employees' non-compliant behaviors can create new security vulnerabilities. Inadvertent human error is a contributing factor in 48% of security breaches in businesses in the United Kingdom in 2015 (Statista, 2015). Hence, it is important that employees should take cybersecurity awareness training.

Lack of motivation can be considered a significant factor in employee commitment and performance, according to Metalidou *et al.* (2014). Therefore, in terms of cultural aspects, lack of motivation can result in a decrease in alignment between an employee and the strategic purpose and, in turn, a tendency to resist change during the adoption process of new paradigms such as DevSecOps. Accordingly, thriving in an environment where employees are motivated to continually commit to DevSecOps' purposes strengthens the adoption activities and results.
Lack of support refers to senior management attitude and the existing culture in the organization when DevSecOps adoption is not given adequate prioritization. Tomas *et al.* (2019) identified reluctance to prioritize security as one of the main themes that affect security culture. The DevSecOps process will be hard to implement when senior management and other stakeholders within the organization underestimate security. According to Sánchez-Gordón and Colomo-Palacios (2020) leadership is crucial to promote and spread the DevSecOps culture within the organization. Therefore, top management should support the adoption process and prioritize security.

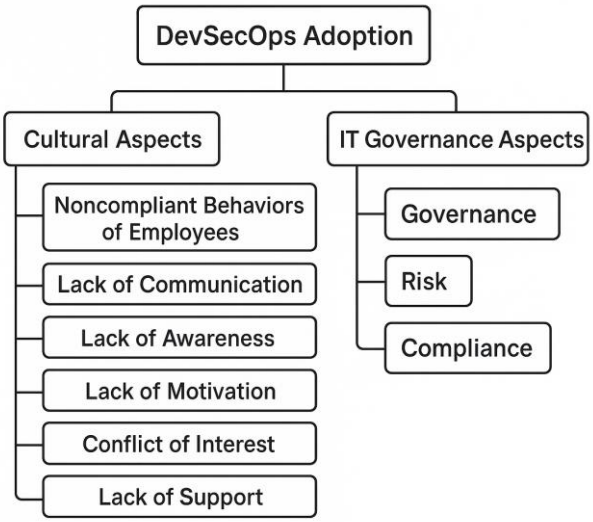

Fig 1: DevSecOps Adoption Challenges: Conceptual Framework

Mansfield-Devine (2018) states that the DevSecOps team needs to communicate with the DevOps team and ask how they can help each other, in what way they do not slow down each other's processes, and how they can initiate new activities together. It is also crucial for the security team to be involved as early as possible to make sure that the communication takes place at the right time and the delivery is done continuously (Sanchez-Gordon & Colomo-Palacios, 2020). Tomas *et al.* (2019) state that there may be a conflict between security and development teams as the security teams may criticize the development in terms of security issues. This problem may cause negative feelings, which may affect the working relationship negatively (Sanchez-Gordon & Colomo-Palacios, 2020).

According to Wiedemann (2018), agile IT organizations reduce the gap between business and IT by making IT a partner instead of a service provider. DevSecOps and Agile have similar goals as eliminating silos and increasing collaboration and teamwork (GSA Tech Guides, n.d.). Therefore, DevSecOps can help reduce the gap between business and IT. Since one of the goals of IT Governance is ensuring that IT investments generate business value (MITRE, n.d.), using DevSecOps in an organization enables this by reducing the gap between business and IT while providing a more secure way to deploy software.

The GRC (governance, risk, compliance) approach will be used when looking into the IT governance aspects (Racz *et al.,* 2010). Governance is mainly about alignment with business and IT goals; the risk is related to the risk appetite of the organization, and compliance is related to laws and regulations concerning the organization.

Potential benefits of adopting a GRC approach include the following:

- *IT Governance and DevSecOps:* DevSecOps can eliminate silos and increase collaboration and teamwork. Hence it can reduce the gap between IT and business and help to align business and IT with the same goals.
- *IT Risk Management and DevSecOps:* DevSecOps extends DevOps by including security processes, and this approach minimizes security risk and breaches from the start of the journey. Including security from the beginning helps to minimize the risks. Therefore, DevSecOps can help organizations' risk management processes.
- *IT Compliance and DevSecOps:* DevSecOps aims to reduce security breaches by integrating security from the beginning; this can help to reduce regulation-related problems with external parties, such as data breaches and GDPR-related problems.

DevSecOps can provide good results for these three GRC aspects, but other problems may arise during its implementation. First, the autonomy given to teams may create additional problems. Organizations can create internal policies to eliminate the possible problems that may occur because of the additional freedom for teams. The autonomy may also create new risks, so more audits and controls may be required. Although DevSecOps enables teams to work more collaboratively, autonomous teams may focus on their issues and lose focus on the business side. Therefore, management should check that the business goals are followed properly. In the end, DevSecOps still provides beneficial results in terms of IT Governance, but in order not to experience adoption problems, it is necessary to take appropriate steps in each aspect of IT GRC and not ignore the governance side.

4 Overcoming Challenges in DevSecOps Adoption: A Plan-Do-Check-Act Approach

In DevSecOps adoption, identifying possible threats is crucial for facing and overcoming problems in the process (Bucena & Kirikova, 2017). Well-defined alignment between teams and strategic direction should be carefully organized and managed to build a culture of ownership with employees and make them consciously accept adopting the new paradigm. When an employee is not clear on the processes, it creates ambiguity and, as a result, diminishes commitment and ownership drastically (Layton, 2005). Therefore, instructions and policies must be translated into specific, measurable, assignable, realistic, and time-bound goals to increase transparency, clarity, and level of satisfaction (SMART criteria, 2021).

Automation can minimize the risk of employee manipulation and human error. Determining and creating automated workflows (a.k.a. DevOps pipelines) is one of the most efficient ways of providing a concrete system to prevent further divergence that might lead to a break or manipulation within the flow. Thus, standards and policies can be applied automatically without any intervening drawbacks. Continuous feedback loops should happen within the teams and across the teams to fulfill the need to balance expectations and create valuable results. Continuous Delivery is essential to release early and faster with incremental outcomes to end users, which simply leads to an increase in customer satisfaction in turn. Accordingly, continuous feedback becomes fundamental when it comes to DevSecOps for the teams to develop applications agilely.

Building a culture of transparency from the risk, safety, and security perspective not only enables seeing the problems as they occur but also increases agility with the ability to adopt continuous learning and improvement activities. Transparency around principles, processes, and methods helps employees to clarify ambiguity, be on the same track, and share a single source of truth among teams. In light of these, management can make good and quick decisions while monitoring and measuring the DevSecOps adoption process within an organization that has a transparent mindset.

Getting top management's support is crucial to changing the organization's culture and adopting DevSecOps processes within the organization. IT Governance is an important factor in aligning business and IT on the same page. Also, IT GRC's other aspects, such as risk management and compliance, should be examined to implement changes to have a successful DevSecOps adoption. DevSecOps needs to have faster and more autonomous processes. This creates new risks, and new risk audit mechanisms should be created. This autonomous structure also needs new policies, internal compliance should be revised to create new policies for a smoother DevSecOps adoption. New risk audit mechanisms and policies will increase the chance of successful DevSecOps adoption. Changes in IT GRC will be a complementary force as they will also affect the culture more formally.

Both cultural changes and IT Governance changes are necessary to have a successful DevSecOps adoption. Both changes will affect each other. Changes in the IT Governance side will provide a basis for cultural changes. Hence, necessary steps should be taken on both sides to overcome challenges in DevSecOps adoption.

5 Findings and Discussion: Addressing the Research Questions

This section discusses the study's findings, considering the RQs and relevant literature. The aim is to critically interpret the results and highlight the implications of cultural and IT governance dimensions in DevSecOps adoption. While commonly identified human-related challenges such as "resistance to change", "lack of awareness", and "communication gaps" can appear generic to many technological transformations, their manifestation, overall impact, and the solutions they necessitate become particularly nuanced and acute within the specific context of DevSecOps. This distinction emanates primarily from DevSecOps' core principles, which include integrating security at speed, fostering deep collaboration across traditionally siloed Development, Security, and Operations (Dev-Sec-Ops) teams, and cultivating a culture of shared responsibility. These principles transform such general barriers into specific impediments to achieving the agile and secure software delivery central to DevSecOps. Understanding these amplified challenges, as categorized in this study's conceptual framework (Figure 1), is therefore crucial for the subsequent discussion.

5.1 RQ1 – What are the main challenges in adopting DevSecOps?

The adoption of DevSecOps in organizations is not just a technical challenge; it is a transformational process that affects structural, procedural, and particularly cultural aspects. This study has highlighted several organizational challenges, including fragmented communication between development, security, and operations teams, a reluctance to share responsibilities, and a general resistance to cultural change. These findings are consistent with those reported by Kumar and Goyal (2020), who point out that cultural inertia and siloed team structures are significant barriers. Additionally, issues such as a lack of security awareness and a misalignment between security and delivery priorities were especially notable. These concerns echo the results

of the Synk (2020) survey and underscore the systemic friction between agility and security enforcement.

	Plan	**Do**	**Check**	**Act**
Culture Change	Identify the potential cultural resistance factors to DevSecOps.	Create a program to clear identified resistances against DevSecOps.	Observe new situations about resistance factors.	Find a way to eliminate the remaining resistance factors.
	Create a security awareness training program.	Implement the education program.	Measure the new security awareness level of employees (with some form of test).	Identify problems regarding security awareness and create new training materials based on weak spots.
	In-house training on DevSecOps to increase employee motivation.	Organize the planned training programs.	Measure the new employee motivation towards DevSecOps.	Identify missing aspects and perspectives. Set new training or info sessions as necessary.
	Seek and enrol top management support for DevSecOps adoption	Explain how DevSecOps work and clarify implementation benefits.	Check the new level of top management support towards DevSecOps.	If full support from top management could not get, try to find ways to get it.
	Identify possible conflict of interest points.	Find the source of these conflict-of-interest points. Address these issues through new procedures, info sessions, training, etc.	Gather feedback from different teams such as security, DevOps and Software Development and identify conflicts.	Resolve possible conflicts by communicating frequently with the security, DevOps and Software Development teams.
	Identify the communication problems between security and development teams	Prepare the means of communication between teams and schedule meetings if necessary	Get feedback from the teams and make sure the communication occurs smoothly	Identify the communication failures and resolve the problems with the corresponding teams
IT Govern-ance	List the possible policies for effective DevSecOps processes.	Implement new policies.	Check the new policies to help employees to stay on track.	If existing policies are not enough, create new policies.
	Identify possible risks within the DevSecOps processes.	Implement new risk management procedures and audit mechanisms.	Check the new risk audit mechanisms are working properly.	Apply working (risk audit) mechanisms and add new features if necessary.
	Reinforce the alignment of business and IT goals.	Improve communication between business and IT to facilitate a smoother DevSecOps adoption process.	Check if business and IT are “on the same page”.	Identify other steps to strengthen communication.

Table I: A proposed PDCA framework and action list for DevSecOps Adoption (addressing IT governance and culture change issues)

5.2 RQ2 – What are the effects of culture on a business’s adoption of DevSecOps?

Culture plays a crucial role in enabling or hindering successful DevSecOps practices. Research by Sánchez-Gordón and Colomo-Palacios (2020) highlights that cultural factors such as transparency, accountability, communication, and shared responsibility are essential for aligning teams and integrating security early in the software development lifecycle. This study supports these findings by demonstrating how employee motivation, leadership commitment, and awareness levels affect DevSecOps initiatives. A significant issue identified is the lack of support from top management and the unclear ownership of security responsibilities, which can undermine secure development practices. Additionally, these results align with the work of Tomas and Huang (2019), who found that cultural misalignment often creates conflicts between development and security teams, ultimately hindering secure coding efforts.

5.3 RQ3 – What are the effects of IT Governance on a business's adoption of DevSecOps?

The role of IT governance in adopting DevSecOps has been relatively underexplored in existing literature. However, this study offers new insights by examining governance-related challenges through the lens of GRC (Governance, Risk, Compliance). The findings indicate that unclear policies, inconsistent risk assessment practices, and weak alignment between business and IT objectives are obstacles to secure and agile software delivery. By implementing effective governance mechanisms, organizations can institutionalize DevSecOps principles, ensuring accountability and compliance across teams. As highlighted by Wiedemann (2018), governance frameworks such as COBIT can provide the structural foundation necessary for DevSecOps to succeed in regulated and process-oriented environments. The study also confirms that IT governance mechanisms indirectly influence organizational culture, thereby enabling or hindering the success of DevSecOps.

5.4 RQ4 – What measures or framework can be adopted as regards culture and IT governance to overcome DevSecOps adoption challenges?

In response to the challenges identified in the current literature, this research proposes a Plan-Do-Check-Act (PDCA) cycle specifically designed for DevSecOps adoption. The PDCA framework has been used in related research to underpin an operational framework for cybersecurity in small-to-medium sized enterprises (Metin *et al.,* 2024). Here, it is adapted to outline actionable steps to promote cultural change, including structured awareness training, motivational workshops, and leadership engagement. At the same time, it addresses IT governance issues by establishing internal policies, risk audit mechanisms, and compliance structures. The challenges presented in Figure 1 are systematically addressed using the PDCA (Plan-Do-Check-Act) structure outlined in Table I. This table details practical steps for planning, executing, monitoring, and improving both cultural and governance-related aspects. Consequently, Figure 1 serves as a diagnostic framework, while Table 1 operationalizes this framework through the PDCA methodology. The integration of these two components supports the characterization of the overall approach as a PDCA-based conceptual framework.

6 Conclusion

This study examines the challenges organizations face when adopting DevSecOps, particularly regarding organizational culture and IT governance. The findings emphasize that successful adoption of DevSecOps is not merely a technical issue; it also requires a socio-organizational transformation. This transformation involves changes in human behavior, managerial support, policy structures, and cultural values.

To tackle these complex challenges, this research makes a number of contributions: firstly, it presents a conceptual framework (Figure 1) that clearly identifies and categorizes essential human-related cultural barriers and IT governance constraints. Secondly, it proposes an actionable PDCA implementation framework (Table 1), providing a structured approach with an action list for cultural adaptation and IT governance improvements. This integrated, PDCA-based framework highlights interdependencies between these elements and is designed to bridge the gap between security practices and strategic business goals, fostering a more secure and agile software development environment.

From a practitioner's perspective, while this PDCA-based framework provides a foundational structure, its real-world application requires careful contextualization. Organizations should adapt the specific actions within the PDCA cycle (as detailed in Table 1) to align with their available resources, operational scale, and any relevant industry-specific regulations or unique business challenges. The iterative nature of the PDCA model supports this tailored

implementation, allowing companies to prioritize interventions and progressively evolve their DevSecOps practices.

Future research could focus on applying and empirically validating the proposed framework through real-world case studies, especially in regulated industries such as finance and healthcare, where IT governance is crucial. Additionally, longitudinal studies could evaluate the cultural transformation process over time, while metrics-based assessments could refine the PDCA cycle for different organizational contexts.

References

Ahmed, Z., & Francis, S. (2019). Integrating security with DevSecOps: Techniques and challenges. *Proceedings of the ICD 2019*, 178–182. https://doi.org/10.1109/ICD47981.2019.9105789

Bucena, I., & Kirikova, M. (2017). Simplifying the DevOps adoption process. In *BIR Workshops* (pp. 1–15).

Cois, C. A., Yankel, J., & Connell, A. (2014, October). Modern DevOps: Optimizing software development through effective system interactions. In *2014 IEEE International Professional Communication Conference (IPCC)* (pp. 1–7). IEEE. https://doi.org/10.1109/IPCC.2014.7020363

Da Veiga, A., & Eloff, J. H. (2010). A framework and assessment instrument for information security culture. *Computers & Security, 29*(2), 196–207. https://doi.org/10.1016/j.cose.2009.09.002

De Haes, S., & Van Grembergen, W. (2004). IT governance and its mechanisms. *Information Systems Control Journal, 1*, 27–33.

Fox, A. (2020). DevOps and IT Governance in the Department of Defense. *MITRE Technical Report*.

GSA General Services Administration. (n.d.). *Understanding the differences between Agile & DevSecOps – From a business perspective*. GSA Technology Transformation Services. Retrieved April 7, 2025, from https://tech.gsa.gov/guides/understanding_differences_agile_devsecops/

INTERPOL. (2020). *Cybercrime: COVID-19 impact*. https://www.interpol.int/en/content/download/15526/file/COVID-19%20Cybercrime%20Analysis%20Report-%20August%202020.pdf

Kumar, R., & Goyal, R. (2020). Modeling continuous security: A conceptual model for automated DevSecOps using open-source software over cloud (ADOC). *Computers & Security, 97*, 101967. https://doi.org/10.1016/j.cose.2020.101967

Layton, T. P. (2005). *Information security awareness: The psychology behind the technology*. AuthorHouse.

Lumpatki, S. S., Patwardhan, S., & Kulkarni, M. (2024). Implementing "DevSecOps as a Culture"—The Concept, Benefits, Execution Strategies, and Challenges. In *International Conference on Smart Computing and Communication* (pp. 189-197). Singapore: Springer Nature Singapore.

Luz, W., Pinto, G., & Bonifacio, R. (2019). Adopting DevOps in the real world: A theory, a model, and a case study. *Journal of Systems and Software, 157*, 110384. https://doi.org/10.1016/j.jss.2019.07.083

Mansfield-Devine, S. (2018). DevSecOps: Bridging the gap between dev and ops—and security. *Network Security, 2018*(10), 5–11. https://doi.org/10.1016/S1353-4858(18)30110-0

Mao, R., Zhang, H., Dai, Q., Huang, H., Rong, G., Shen, H., & Lu, K. (2020). Preliminary findings about DevSecOps from grey literature. In *Proceedings of the 20th International Conference on Software Quality, Reliability and Security* (pp. 1–8). IEEE.

Metalidou, E., Marinagi, C., Trivellas, P., Eberhagen, N., Skourlas, C., & Giannakopoulos, G. (2014). The human factor of information security: Unintentional damage perspective. *Procedia - Social and Behavioral Sciences, 147*, 424–428. https://doi.org/10.1016/j.sbspro.2014.07.133.

Metin, B., Duran, S., Telli, E., Mutlutürk, M., & Wynn, M. (2024). IT Risk Management: Towards a System for Enhancing Objectivity in Asset Valuation That Engenders a Security Culture. *Information*, *15*(1), 55. https://doi.org/10.3390/info15010055

MITRE Systems Engineering Guide. (n.d.). IT governance. Retrieved February 18, 2021, from https://www.mitre.org/publications/systems-engineering-guide/enterprise-engineering/enterprise-planning-and-management/it-governance

Mohan, V., & Othmane, L. B. (2016, August). SecDevOps: Is it a marketing buzzword?—Mapping research on security in DevOps. In *2016 11th International Conference on Availability, Reliability and Security (ARES)* (pp. 542–547). IEEE. https://doi.org/10.1109/ARES.2016.81

Mutlutürk, M., Wynn, M., & Metin, B. (2024). Phishing and the Human Factor: Insights from a Bibliometric Analysis. *Information*, *15*(10), 643. https://doi.org/10.3390/info15100643

Myrbakken, H., & Colomo-Palacios, R. (2017, October). DevSecOps: A multivocal literature review. In *International Conference on Software Process Improvement and Capability Determination* (pp. 17–29). Springer. https://doi.org/10.1007/978-3-319-67383-7_2

Nagasundari, S., Manja, P., Mathur, P., & Honnavalli, P. B. (2025). Extensive review of threat models for DevSecOps. *IEEE Access*, 13, 45252–45271. https://doi.org/10.1109/ACCESS.2025.3547932

Orehek, Š., & Petrič, G. (2020). A systematic review of scales for measuring information security culture. *Information & Computer Security, 28*(5), 789–808. https://doi.org/10.1108/ICS-06-2019-0070.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Blackwell Publishing. https://doi.org/10.1002/9780470754887.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., & Roen, K. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews*. ESRC Methods Programme. Lancaster University, UK. DOI: 10.13140/2.1.1018.4643.

Racz, N., Weippl, E., & Seufert, A. (2010, July). A process model for integrated IT governance, risk, and compliance management. In *Proceedings of the Ninth Baltic Conference on Databases and Information Systems (DB&IS 2010)* (pp. 155–170).

Rahman, A. A. U., & Williams, L. (2016). Security practices in DevOps. In *Proceedings of the 2016 IEEE Symposium on Security and Privacy Workshops (SPW)* (pp. 94–95). IEEE. https://doi.org/10.1109/SPW.2016.30

Rajapakse, R. N., Zahedi, M., Babar, M. A., & Shen, H. (2022). Challenges and solutions when adopting DevSecOps: A systematic review. *Information and software technology*, *141*, 106700.

Sánchez-Gordón, M., & Colomo-Palacios, R. (2020, June). Security as culture: A systematic literature review of DevSecOps. In *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops* (pp. 266–269). https://doi.org/10.1145/3387940.3392177

Sas, M., Hardyns, W., van Nunen, K., Reniers, G., & Ponnet, K. (2020). Measuring the security culture in organizations: A systematic overview of existing tools. *Security Journal, 33*, 717–740. https://doi.org/10.1057/s41284-020-00229-4

Shin, D., Kim, J., Pawana, I. W. A. J., & You, I. (2025). Enhancing cloud-native DevSecOps: A Zero Trust approach for the financial sector. *Computer Standards & Interfaces*, 103975.

SMART criteria. (2021, January 8). In *Wikipedia*. Retrieved January 30, 2021, from https://en.wikipedia.org/wiki/SMART_criteria

Tomas, N., Li, J., & Huang, H. (2019, June). An empirical study on culture, automation, measurement, and sharing of DevSecOps. In *2019 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)* (pp. 1–8). IEEE. https://doi.org/10.1109/CyberSecPODS.2019.8885146

Wiedemann, A. (2018, January). IT governance mechanisms for DevOps teams—How incumbent companies achieve competitive advantages. In *Proceedings of the 51st Hawaii International Conference on System Sciences*.

Zaydi, M., & Nassereddine, B. (2020). DevSecOps practices for an agile and secure IT service management. *Journal of Management Information and Decision Sciences, 23*(2), 1–16.

Zhu, L., Bass, L., & Champlin-Scharff, G. (2016). DevOps and its practices. *IEEE Software, 33*(3), 32–34. https://doi.org/10.1109/MS.2016.81

4. BRIDGING THE TWO SCIENCES: TAPPING INTO THE CONFLUENCE OF NETWORK SCIENCE AND DESIGN SCIENCE RESEARCH

Samrat Gupta
IIM Ahmedabad, University of Agder
samratg@iima.ac.in, samrat.gupta@uia.no

Thorsten Schoormann
Roskilde University
tschoormann@ruc.dk

Mukul Gupta
IIM Indore
mukulg@iimidr.ac.in

Leona Chandra Kruse
University of Agder
leona.chandra@uia.no

Abstract
Network Science is getting growing attention across a diverse range of disciplines. In this study, we investigate the role of design science research (DSR) in developing and applying network science theories and artifacts. We are especially motivated by the ability of DSR to advance theory and practice in designing systems and the potential of network science to address various real-world problems. By surveying information systems literature at the confluence of network science and DSR, we observed a focus on enhancing the design outcome of network science approaches mainly through deductive approaches (as compared to inductive approaches). Only a few papers in our sample provide insights into how DSR can support network science from problem understanding through artifact building to evaluation. Our work provides an area of inquiry for blending the two forces of DSR with network science to address important problems.

Keywords
Network Science, Design Science Research, Network Analysis, Network Theory

1. Introduction
Network Science (NetSci) is a field that transcends disciplinary boundaries and provides insights into diverse phenomena, such as social dynamics (online communication), biological systems (human brain), economic markets (credit relationships among banks), and technological systems (blockchain). NetSci emphasizes algorithmic modelling and hypothesis testing to tackle problems of unprecedented scale and complexity. Given its power, it has been used for solving societal challenges including systemic risks in banking systems, knowledge sharing in online platforms, and urban-rural health disparities (Cao & Wang, 2018; Gupta & Kumar, 2021; Hu et al., 2012; Lu et al., 2017).

Researchers in the field of NetSci are therefore confronted with the challenging task of developing novel and reusable artifacts beyond disciplinary boundaries, capable of addressing major problems, a challenge that is well-known in Design Science Research (DSR). DSR solves relevant issues by employing research methods and relying on the existing knowledge base to ensure rigor across the entire process of building new design artifacts (Hevner et al., 2004; Peffers et al., 2007). Additionally, it emphasizes the need to extract design knowledge, such as in the form of theories and principles to ensure subsequent design journeys make use of existing insights (Gregor et al., 2020). Due to its ability, DSR offers a vast potential for network scientists, and we argue that both streams NetSci and DSR should work hand in hand to break divisions between different scientific streams.

While researchers in the field of computational science have made significant advances in employing DSR to craft artifacts for artificial intelligence and machine learning (Schoormann et al., 2025), the focus on the integration of DSR with NetSci has been lacking. Only a few studies underscore the use of DSR for this particular field (Shangguan et al., 2022; Velichety & Ram, 2020). This is problematic because currently, network scientists need to formally develop theories, principles, and guidelines for designing and using artifacts for network-based modelling and analysis that can be reused in different projects. Recognizing the critical role that networks play in today's world, the objective of this paper is to investigate how the rigor developed in DSR could be translated to NetSci. Therefore, we pose the following question:

How can design science research support the advances in network science?

The contribution of this research is to propose how to support NetSci, as networks across domains are being mapped digitally, by using what has been known in DSR. For around three decades, network scholars have conducted research with mathematical models, based on theories of statistical physics and complex systems (Albert & Barabási, 2002; Cimini et al., 2019). We propose that the evolution of NetSci, which recognizes the critical role of computing could benefit from the scientific rigor of DSR.

To carry out the research, we review examples of important digital artifacts and theoretical developments in NetSci to identify their development challenges. We reviewed a total of 64 papers on NetSci published in IS outlets out of which 10 use DSR. Our analysis suggests that DSR can potentially address the challenges in NetSci by offering established methods to ensure reliability and validity while enabling the reuse of design knowledge derived from developed theories and artifacts. This study aims to provide valuable insights into the evolution of NetSci in an increasingly digital landscape.

This paper proceeds as follows. Section 2 discusses the background of NetSci and DSR, respectively. Section 3 reviews the existing IS literature at the intersection of NetSci and DSR. Section 4 discusses how network scientists leverage DSR. Section 5 and 6 theorize the potential of DSR in NetSci. Section 7 discusses the implications and possible contributions while concluding the paper.

2. Research Background

2.1. Network Science

Network Science (NetSci) has emerged as a potent interdisciplinary field, leveraging computational modelling, statistical mechanics, and graph theory to comprehend complex real-world systems through their relational and structural properties (Albert & Barabási, 2002; Cimini et al., 2019). Some of the prominent computational problems in the field of NetSci have been around the role and detection of networked communities, link analysis, influential node identification, and dynamics of information spreading (Chelmis & Prasanna, 2013; Gupta & Mishra, 2021; Kumar et al., 2017; Schroeder et al., 2007). In the context of information systems (IS), NetSci offers theoretical lens and methods for analysing the interconnected components that define modern digital ecosystems (Lee & Kim, 2018). In this section, we discuss the foundational concepts and relevance of NetSci to the domain of IS.

NetSci examines systems as networks of nodes (entities) connected by edges (links). Real-world systems constituting various entities, such as individuals, information flows, technologies, and organizations can be represented by networks. Many real-world systems are

based on pioneering concepts such as small-world phenomena and scale-free networks which set the tone for future developments in the field of NetSci (Watts & Strogatz, 1998; Barabási & Albert, 1999). Such theoretical and methodological advances in NetSci are invaluable for modelling digital systems such as social media platforms, knowledge-sharing platforms, and organizational communication structures.

The advent of computational and analytical approaches supplemented by large-scale data collection has bolstered the cross-pollination of NetSci in IS. Leveraging techniques for temporal network analysis and community discovery, IS researchers can capture hierarchical and dynamic relationships within digital ecosystems (Li & Yoo, 2022; Zhang et al., 2016). To maximize user engagement and improve knowledge construction, social and professional networks emanating from platforms such as YouTube and LinkedIn have been modelled and analysed using NetSci principles (Baruffaldi et al., 2017; Gupta et al., 2023). Other examples of NetSci research in the realm of IS are provided in the online Appendix.

Advanced methodologies in NetSci like Exponential Random Graph Modelling (ERGM), Stochastic Actor- Oriented Modelling (SAOM), and Multiple Regression Quadratic Assignment Procedure (MRQAP) offer analytical support for examining intricate relationships in real-world complex networks (Robins et al., 2007; Snijders et al., 2010; Rubineau et al., 2019). By capturing dependencies such as homophily and triadic closure ERGM enables modelling of complex network structures, offering insights into the factors that drive network formation (Robins et al., 2007). Since SAOM focuses on the dynamic evolution of networks, it allows the investigation of the co-evolution of network structures and individual behaviours over time (Snijders et al., 2010). MRQAP facilitates hypothesis testing by addressing the dependence inherent in network ties (Rubineau et al., 2019). Moreover, the mixed method approach to NetSci by combining digital trace data and qualitative approaches for better addressing the problems related to technology-enabled environments has been proposed (Whelan et al., 2016). Methods such as epidemic spreading models and influence propagation models are useful for forecasting the reach of communication campaigns or the adoption of innovations in IS (Bae et al., 2021; Fang et al., 2013). Study of open-source software development also relies on NetSci (Oh & Jeon, 2004; Van Antwerp & Madey, 2010). Tackling the involuntary default behaviour of enterprises requires an understanding of concepts such as network embeddedness and networked community structure (Chen et al., 2025; Gupta & Kumar, 2016). Additionally, by optimizing performance and identifying critical vulnerabilities, NetSci helps dependency networks in software systems and IT architectures (Temizkan et al., 2017). The integration of machine learning and NetSci has enabled IS researchers to develop complex decision-support systems (Chung & Lai, 2023), for example, for fraud detection and customer relationship management (Ma et al., 2024; Zhang et al., 2022).

2.2 Rigor in Design Science Research

The DSR paradigm strives to advance science and technology by developing novel and useful artifacts such as methods, constructs, models, and instantiations (Nabavian & Parsons, 2024). Numerous complex issues that arise in the increasingly digital world, so-called wicked problems with contradictory or evolving problem requirements, are being addressed by DSR (Hevner et al., 2004). Building on Herbert Simon's idea of the Science of the Artificial, DSR provides well-established methodologies (Schoormann et al., 2024) and evaluation frameworks (Venable & Baskerville, 2012) to craft such relevant artifacts. While acknowledging the engineering-driven approach to building something new, it also considers the important role of extracting and codifying design knowledge (Storey & Baskerville, 2025). Among the well-accepted forms of design knowledge codification are design principles (Gregor et al., 2020).

They offer prescriptive assertions that guide academics and practitioners in drawing on prior experiences and avoiding mistakes while developing new information technology (IT) artifacts (Chandra Kruse et al., 2022).

Besides some adoptions of DSR to the contextualized requirements of a certain domain or application type, the IS community has not yet offered DSR-based insights for knowledge production in NetSci, even though the necessity of knowledge contribution of DSR in applied disciplines has been widely acknowledged (Nunamaker & Briggs, 2012). There are no frameworks available to help researchers produce and convey knowledge derived from network-based theory and artifact designs.

3 Research Approach

We adopted a two-phased approach to (1) determine the status of research at the intersection of NetSci and DSR as well as (2) explore and reflect on emerging and dormant higher-order configurations of DSR-driven NetSci research.

In Phase 1, we conducted a literature review using Google Scholar and AISeL to build a corpus of IS NetSci articles. With an emphasis on NetSci, we investigated and validated a variety of search keywords before settling on the following search phrase: “social network” OR “network analysis” OR “complex network” OR “empirical network” OR “network science”. Through the analysis of 64 papers, a final sample of 10 papers that used the DSR paradigm could be extracted. In this process, we excluded the papers that were not positioned as IS NetSci or DSR, non-peer-reviewed preprints, and those without the availability of full-text. For analysis, the first author coded the sample to disclose, among others, the primary term used, the design outcome, and the type of data utilized in a study (see Section 4).

In Phase 2, based on the understanding of the identified literature and its focus on the design outcome, we explored inductive and deductive characteristics of research that can benefit future design problems with similar characteristics. We conducted a workshop within the team of authors in which we brainstormed possible DSR interventions and configurations within NetSci (see Sections 5 and 6).

4 How Researchers in NetSci Leverage DSR

To explore how DSR has been used in IS NetSci research, we conducted a literature review (see Section 3 for methodical explanations and Figure 1 for the final sample). The analysis reveals that a majority of studies focus on building new artifacts incorporating computational elements from NetSci. 6/10 papers can be categorized as design science artifacts and 4/10 papers can be classified as theoretical aspects of design science. Although widely accepted, only 3/10 papers in our sample derive design knowledge to guide the design of social recommender systems, service platforms, and labelling of edges for network analysis (Arazy et al., 2010; Gupta et al., 2025; Janiesch et al., 2019).

Another observation is that most of these papers seek to address major societal and business problems, for instance:

- Assisting crime investigators in identity matching problems, an issue in organized crimes such as arms smuggling, drug trafficking, and money laundering (Xu et al., 2007).
- Prediction of multiple levels of information usefulness in online knowledge communities as the online knowledge base increases, such as communities for sharing profession, work, and special interest knowledge (Liu et al., 2020).

Reference	Primary NetSci term						Primary DSR term				Description
	(Social) Network analysis	Social relationships	Community/cluster analysis	Graph theory	Eigen attention centrality	Network effects	Design (science) artifact	Design theory	Design principles	Design/theory-driven design	
Velichety & Ram (2020)	-	-	●	-	-	-	●	-	-	-	Addresses the problem of recommending online communities for subscription on social media platforms.
Janiesch et al. (2019)	-	-	-	-	-	●	-	●	●	-	Provides a theory of IS design to guide the design of methods that help in the development of service platforms.
Howison et al. (2011)	●	-	-	-	-	-	-	-	-	●	Examines validity concerns related to combining NetSci techniques with digital trace data.
Arazy et al. (2010)	-	●	-	-	-	-	-	-	-	●	Develops a design theory that could improve the performance of social recommenders.
John et al. (2016)	-	-	●		-	-	●	-	-	-	Develops a method for grouping related questions based on a network of social connections between the questions, answers, askers, and answerers.
Liu et al. (2020)	-	-	●	-	-	-	-	-	-	●	Develops a new text analytic framework to identify key features in online forums and apply them to classify the utility of solutions included in the knowledge base
Shangguan et al. (2022)	●	-	-	-	●	-	●	-	-	-	Develops a new composite metric to predict the development of a trading strategy that yields substantial excess returns.
Xu et al. (2007)	●	●	●	-	-	-	●	-	-	-	Develops a method to match criminal identity records to address the matching problem which is a subproblem of drug offenses.
Gupta & Tiwari (2022)	●	-	●	-	-	-	●	-	-	-	Proposes an educational framework based on the problem of community detection in social networks for developing computational thinking skills.
Gupta et al. (2025)	●	●	-	●	-	-	●	-	●	-	Proposes a method for sign prediction of ties in social networks leveraging the tenets of graph augmentation and graph regularization for information diffusion.

Figure 1. Overview of papers using Network Science and Design Science Research

- Building an information systems design theory that offers prescriptions and principles of implementation for engineering a method to design service platforms for network effects (Janiesch et al., 2019).
- Aiming to develop a pedagogical framework for developing computational thinking skills among students (Gupta & Tiwari, 2022).

Other outcomes that emerge from blending NetSci with DSR include a recommender system to identify relevant online communities with social media platforms (Velichety & Ram, 2020), the validity of using digital trace data for social network analysis to ensure the accurate interpretation and representation of network dynamics (Howison et al., 2011). Arazy et al. (2010) designed a framework for designing social recommender systems to enhance user experiences through personalized recommendations. John et al. (2016) showed how design-based techniques improve network clustering and information retrieval by using a graph-based clustering approach to find related questions on online platforms. Another study made synergistic use of network science and DSR to develop a composite measure of information flow in financial networks thereby offering useful insights into the dynamics of the stock market (Shangguan et al., 2022). Gupta et al. (2025) design a method for sign prediction of ties in social networks and demonstrate its utility for predicting trust/distrust on a user network emanating from IMDb platform. Collectively, these studies emphasize how DSR serves as a methodological paradigm that enables network scientists to develop novel, theoretically informed and practically applicable solutions. Our analysis reveals that there has been a clear inclination towards the use of deductive approaches of DSR for NetSci research. The adoption of inductive approaches of DSR in NetSci remains an uncharted area. This suggests that research on NetSci has predominantly relied on deriving new insights and solutions from established theory and models. However, deriving general principles from empirical data and specific observations has not yet been explored in this context.

5 Emergence of Deductive Configurations of DSR in NetSci

Deductive configurations of DSR instantiate an artifact in a practical context and test if it delivers a specific result in a use case (Daase et al., 2024). Deduction is usually employed during the design and iterative refinement phases. Since deduction does not always lead to absolute claims, a design outcome based on generalized descriptive knowledge may indicate the extent to which a property aids in the resolution of the problem, allowing for further design iterations and improvements (Winter, 2013).

5.1 Configuration for Artefact-oriented Design

One of the configurations of DSR in NetSci follows a structured approach focusing on the development of computational artefacts that address specific social and business problems. In these cases, the actual goal of the artefact is specified beforehand, and the researchers systematically work towards achieving this prespecified goal through DSR.

The initial phase in artefact-oriented design in NetSci involves defining the research problem and its associated objectives. This could mean addressing inefficiencies in social network analysis or optimizing business supply chains through network modelling. To ensure the artefact is well-founded, researchers engage in theoretical grounding and empirical analysis. This involves drawing insights from existing NetSci theories, computational models, and domain-specific literature. Artefact-oriented design is inherently iterative and is reflected by Baustein's third configuration (Schoormann et al., 2024) (see Figure 2). Researchers experiment with various computational techniques such as machine learning for sign prediction

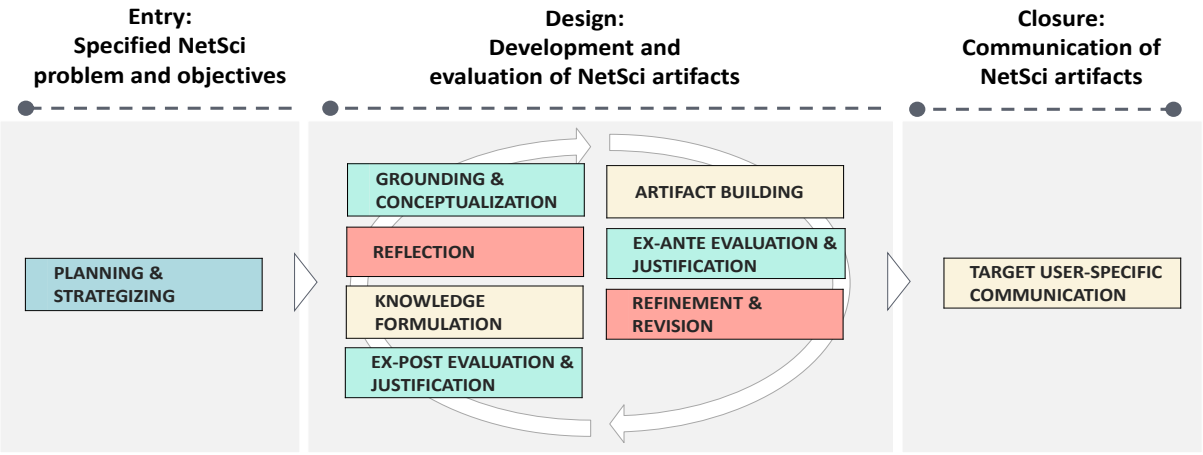

Figure 2. Artefact-oriented design configuration (deductive)

of links, or soft computing for community detection to create a functional prototype (Gupta et al., 2025; Gupta & Kumar, 2020). The choice of methods depends on the complexity and nature of the problem. Once an initial artefact is developed, it undergoes rigorous evaluation to assess its effectiveness and applicability. Network-based metrics, such as conductance for community detection, and optimized precision for sign prediction indicate the success of artefact. Artefact-oriented design typically emphasizes contributions after the artefact has been developed and validated. These contributions may include novel algorithms, frameworks, or methodologies that advance network science and have practical implications for businesses, policymakers, or social systems. Researchers disseminate their findings not only through academic publications, but also through open-source tools, and industry collaborations to ensure a broader impact. In each phase, artefact's design is informed by incorporating iterative refinements based on experimental testing, stakeholder feedback, and real-world applications.

By taking a network-centric perspective, the configuration of artefact-oriented design enables researchers to develop targeted computational solutions that address complex real-world problems while contributing to theoretical advancements and practical applications in NetSci.

5.2 Configuration for Theory-driven Design

Another configuration that emerged from the existing research at the confluence of NetSci and DSR involves solving social and business problems by learning from abstract theoretical knowledge drawn from social or natural sciences (referred to as kernel theories). The theory of strength of weak ties and homophily are some examples of such kernel theories (Granovetter, 1983; McPherson et al., 2001). This configuration is characterized by an emphasis on theory-ingrained artefacts (Schoormann et al., 2024; Iivari, 2015; Peffers et al., 2018); see related configuration in Figure 3. Unlike artefact-oriented design, where the goal is predefined, theory-driven design is primarily motivated by the researcher's curiosity about the applicability of theoretical constructions, the testing of theoretical predictions, and the translation of abstract insights into practical design knowledge.

An important activity in theory-driven design is deriving generalized design imperatives from the underlying theory (referred as meta-requirements). For example, if a theory suggests that network effects in an online platform emerge because of homophily among users, the meta-

requirements may include designing an artefact that can detect and model homophilic structures in the social network emerging from the platform. Such a process ensures that the design outcome is theoretically grounded and addresses practical scenarios as per expectations.

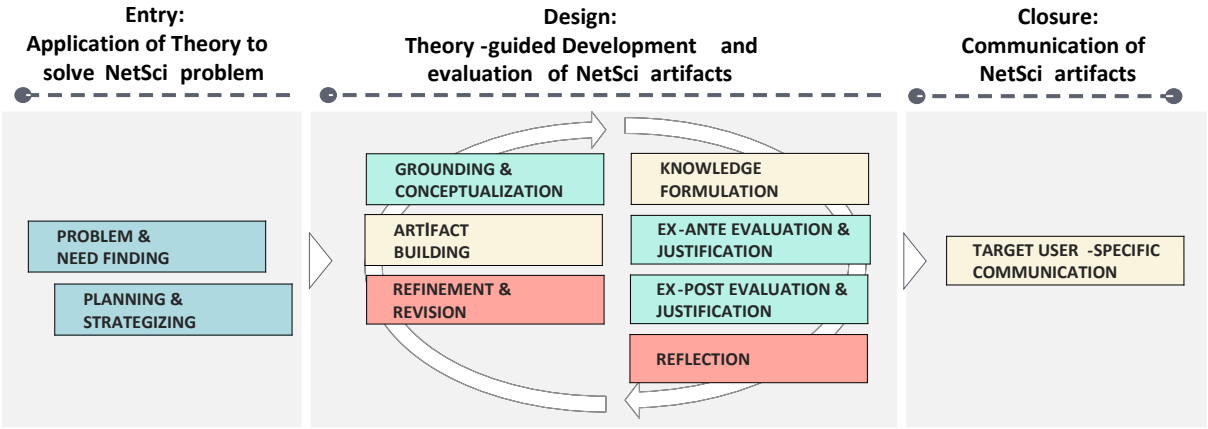

Figure 3. Theory-driven design configuration (deductive).

Secondly, based on the kernel theory and empirical findings, design principles that guide the transformation of theoretical knowledge into explicit prescriptions are formulated. An example of design principle is that on online platforms a feature to indicate the implicit polarity of relationships among users should consider structural balance mechanisms. Such design principles serve as blueprints that can be instantiated in different situational artefacts. For example, a gaming platform that recommends collaboration opportunities among online players should consider structural balance among the competitor teams. Such instantiations evaluate the relevance of kernel theory in real-world thereby allowing for empirical validation and potential refinements to both theory and artefact.

Through systematic progression from abstract theory to design principles to situational artefacts, the configuration of theory-driven design in NetSci provides a paradigm for bridging computational artefacts with theoretical advancements.

6 Dormancy of Inductive Configurations of DSR in NetSci

Inductive configurations of DSR in NetSci generally refer to intervention-based research genres that require a formalized degree of academia-practice collaboration, often across disciplines (Schoormann et al., 2024). Such approaches might be built upon previously completed deductive design projects to advance design knowledge and also reflect the first strategy of DSR wherein design ideas are implemented through artefacts and knowledge is generalized afterwards (Iivari, 2015) (see related configuration in Figure 4). However, despite the potential of such inductive approaches of DSR to generate new design knowledge, they have remained dormant in the landscape of NetSci.

In DSR, induction can occur in the initial phases of a project, which entails synthesizing a body of knowledge or evaluating an artifact by applying it and observing the outcomes (Daase et al.,

2024). In contrast to deductive approaches where objectives are explicitly defined, inductive approaches rely on open-ended exploration and are often sidelined due to their perceived lack of initial direction. They involve analysing complex datasets without predefined hypotheses, making the research process resource-intensive. Also, the lack of predefined goals in inductive configurations poses challenges in evaluation wherein evolving evaluation criteria is required. Leveraging inductive configurations of DSR in NetSci requires advancements that balance the need for systematic evaluation with open-ended discovery. Despite its dormancy, inductive configurations offer significant potential to contribute to NetSci, particularly in areas where network structures and behaviours are poorly understood.

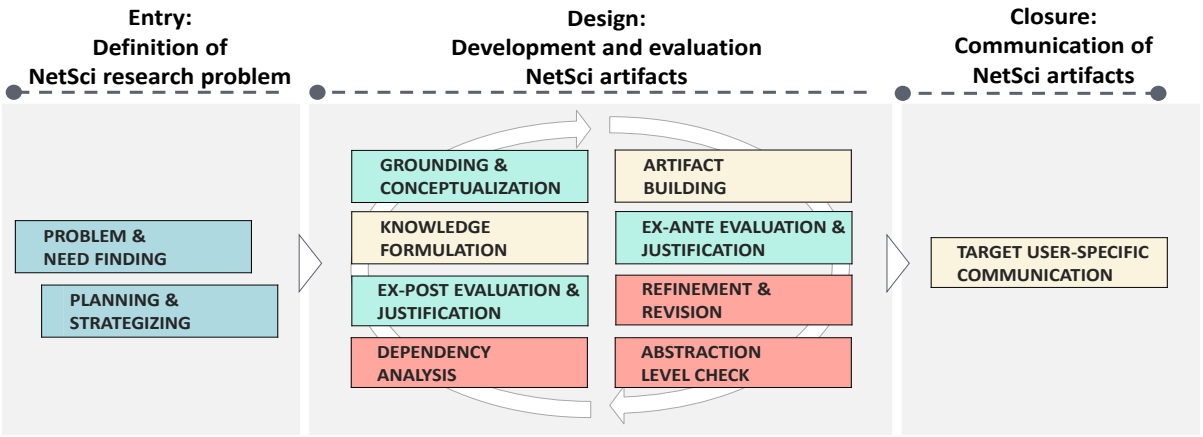

Figure 4. Problem-solving design configuration (inductive).

7 Conclusion

This paper investigates how design science research can support the advances in network science. Our premise is that DSR brings methodological rigor that can be translated to NetSci to improve design outcomes, which often address major societal problems and thus are of great relevance for scholars and practitioners. We surveyed IS literature related to NetSci in which we found that only 10 out of 64 papers lie at the confluence of NetSci and DSR. Our findings point to the tendency to use deductive DSR configurations and approaches in network science studies, applying or validating specific theories and assumptions about artifactual solutions. However, inductive configurations have been largely dormant, though they could provide significant insights, especially where network structures are complex or not well understood. Specifically, this paper calls for leveraging inductive DSR methods, despite their resource- intensive nature to facilitate innovation, discovery and knowledge accumulation in NetSci.

Although this study provides insights on using DSR in NetSci, it is not devoid of limitations thereby opening avenues for future research. First, the analysis of literature is restricted by the search terms used which can be extended to possibly collect a broader range of papers. Second, our study focuses on research within the realm of IS. During the explorative search for the use of DSR for research in network science, we also came across studies published in other research communities. Future research may extend this study to borrow and learn ideas from other fields, such as computer science and operations research. Third, we mainly use Baustein to reflect on common activities in design science research. However, in the future, similar tools can be

extended to emphasize how DSR can be used for advancing network science.

Nonetheless, this study provides reflections on how DSR can orchestrate the future design of network-centric theories and artifacts. With that, it can contribute to both the advancement of rigor in NetSci as well as to computational design in which scholars are increasingly interested in applying novel methods to the building and evaluation of useful artifacts.

References

Albert, R., & Barabási, A. L. (2002). Statistical mechanics of complex networks. Reviews of modern physics, 74(1), 47.

Arazy, O., Kumar, N., & Shapira, B. (2010). A theory-driven design framework for social recommender systems. Journal of the Association for Information Systems, 11(9), 2. Bae, S., Sung, E., & Kwon, O. (2021). Accounting for social media effects to improve the accuracy of infection models: combatting the COVID-19 pandemic and infodemic. European Journal of Information Systems, 30(3), 342-355.

Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. Science, 286(5439), 509-512.

Baruffaldi, S. H., Di Maio, G., & Landoni, P. (2017). Determinants of PhD holders' use of social networking sites: An analysis based on LinkedIn. Research policy, 46(4), 740-750.

Cao, X., & Wang, D. (2018). The role of online communities in reducing urban–rural health disparities in China. Journal of the Association for Information Science and Technology, 69(7), 890-899.

Chandra Kruse, L., Purao, S., & Seidel, S. (2022). How designers use design principles: design behaviors and application modes. Journal of the Association for Information Systems, 23(5), 1235-1270.

Chelmis, C., & Prasanna, V. K. (2013). Social link prediction in online social tagging systems. ACM Transactions on Information Systems, 31(4), 1-27.

Chen, X., & Li, W. L. (2025). Can B2B online communities reduce enterprises' default behavior-From social network perspective. Information & Management, 62(2), 104079.

Chung, W., & Lai, V. S. (2023). A temporal graph framework for intelligence extraction in social media networks. Information & Management, 60(4), 103773.

Cimini, G., Squartini, T., Saracco, F., Garlaschelli, D., Gabrielli, A., & Caldarelli, G. (2019). The statistical physics of real-world networks. Nature Reviews Physics, 1(1), 58-71.

Daase, C., Haertel, C., Nahhas, A., & Turowski, K. (2024). Classifying Design Science Research in Terms of Types of Reasoning from an Epistemological Perspective. In International Conference on Design Science Research in Information Systems and Technology (pp. 155-167). Springer, Cham

Fang, X., Hu, P. J. H., Li, Z., & Tsai, W. (2013). Predicting adoption probabilities in social networks. Information Systems Research, 24(1), 128-145.

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. Sociological theory, 201-233.

Gregor, S., Chandra Kruse, L., & Seidel, S. (2020). Research perspectives: the anatomy of a design principle. Journal of the Association for Information Systems, 21(6), 2.

Gupta, M., Gupta, S., and G. K. Tayi (2025). GASP: A Graph Augmentation-Based Approach for Sign Prediction of Ties in Social Networks. Journal of the Association for Information Systems (Forthcoming)

Gupta, M., & Mishra, R. (2021). Spreading the information in complex networks: Identifying a set of top-N influential nodes using network structure. Decision Support Systems, 149, 113608.

Gupta, S., & Kumar, P. (2016). Community detection in heterogenous networks using incremental seed expansion. In 2016 international conference on data science and engineering (ICDSE) (pp. 1-5). IEEE.

Gupta, S., & Kumar, P. (2021). A constrained agglomerative clustering approach for unipartite and bipartite networks with application to credit networks. Information Sciences, 557, 332-354.

Gupta, S., & Tiwari, A. A. (2022). A design-based pedagogical framework for developing computational thinking skills. Journal of Decision Systems, 31(4), 433-450.

Gupta, S., Jain, G., & Tiwari, A. A. (2023). Polarised social media discourse during COVID- 19 pandemic: evidence from YouTube. Behaviour & Information Technology, 42(2), 227-248

Gupta, S., & Kumar, P. (2020). An overlapping community detection algorithm based on rough clustering of links. Data & Knowledge Engineering, 125, 101777.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS quarterly, 75-105.

Howison, J., Wiggins, A., & Crowston, K. (2011). Validity issues in the use of social network analysis with digital trace data. Journal of the Association for Information Systems, 12(12), 2

Hu, D., Zhao, J. L., Hua, Z., & Wong, M. C. (2012). Network-based modeling and analysis of systemic risk in banking systems. MIS quarterly, 1269-1291.

Iivari, J. (2015). Distinguishing and contrasting two strategies for design science research. European Journal of Information Systems, 24(1), 107-115.

Janiesch, C., Rosenkranz, C., & Scholten, U. (2019). An information systems design theory for service network effects. Journal of the Association for Information Systems: forthcoming

John, B. M., Chua, A. Y., Goh, D. H. L., & Wickramasinghe, N. (2016). Graph-based cluster analysis to identify similar questions: A design science approach. Journal of the

Association for Information Systems, 17(9), 2. Kumar, P., Gupta, S., & Bhasker, B. (2017). An upper approximation based community detection algorithm for complex networks. Decision Support Systems, 96, 103-118.

Lee, C., & Kim, H. (2018). The evolutionary trajectory of an ICT ecosystem: A network analysis based on media users' data. Information & Management, 55(6), 795-805.

Li, H., & Yoo, S. (2022). From information systems resources to effective use: Moderating effect of network embeddedness. The Journal of Strategic Information Systems, 31(3), 101735.

Liu, X., Wang, G. A., Fan, W., & Zhang, Z. (2020). Finding useful solutions in online knowledge communities: A theory-driven design and multilevel analysis. Information Systems Research, 31(3), 731-752.

Lu, Y., Singh, P. V., & Sun, B. (2017). Is a core–periphery network good for knowledge sharing? A structural model of endogenous network formation on a crowdsourced customer support forum. MIS Quarterly, 41(2), 607-628.

Ma, T., Hu, Y., Lu, Y., & Bhattacharyya, S. (2024). Customer engagement prediction on social media: A graph neural network method. Information Systems Research.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. Annual review of sociology, 27(1), 415-444.

Nabavian, S., & Parsons, J. (2024). A Design-Principle-Friendly Conceptual Model of Observational Crowdsourcing. In International Conference on Design Science Research in Information Systems and Technology (pp. 95-108). Cham: Springer Nature Switzerland.

Nunamaker, Jr, J. F., & Briggs, R. O. (2012). Toward a broader vision for information systems. ACM Transactions on Management Information Systems, 2(4), 1-12.

Oh, W. & Jeon, S. (2004). Membership Dynamics and Network Stability in the Open-Source Community: The Ising Perspective. ICIS 2004 Proceedings. Paper 33.

Peffers, K., Tuunanen, T., & Niehaves, B. (2018). Design science research genres: introduction to the special issue on exemplars and criteria for applicable design science research. European Journal of Information Systems, 27(2), 129-139.

Peffers, 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.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. Social networks, 29(2), 173-191.

Rubineau, B., Lim, Y., & Neblo, M. (2019). Low status rejection: How status hierarchies influence negative tie formation. Social Networks, 56, 33-44.

Schoormann, T., Gupta, S., Möller, F., & Chandra Kruse, L. (2025). With Great Power Comes Great Responsibility: Responsible Management of Artificial Intelligence in Supporting Design Research Activities. Proceedings of the 58th Hawaii International Conference on System Sciences, 5405-5414.

Schoormann, T., Möller, F., Chandra Kruse, L., & Otto, B. (2024). BAUSTEIN—a design tool for configuring and representing design research. Information Systems Journal, 34(6), 1871-1901.

Schroeder, J., Xu, J., Chen, H., & Chau, M. (2007). Automated criminal link analysis based on domain knowledge. Journal of the American society for information science and technology, 58(6), 842-855.

Shangguan, W., Leung, A. C. M., Agarwal, A., Konana, P., & Chen, X. (2022). Developing a composite measure to represent information flows in networks: Evidence from a stock market. Information Systems Research, 33(2), 413-428.

Snijders, T. A., Van de Bunt, G. G., & Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics. Social networks, 32(1), 44-60.

Storey, V. C., & Baskerville, R. L. (2025). Digitalization of the natural sciences: Design science research and computational science. Decision Support Systems, 189, 114368.

Temizkan, O., Park, S., & Saydam, C. (2017). Software diversity for improved network security: optimal distribution of software-based shared vulnerabilities. Information Systems Research, 28(4), 828-849.

Van Antwerp, M., & Madey, G. (2010). The importance of social network structure in the open source software developer community. In 2010 43rd Hawaii International Conference on System Sciences (pp. 1-10).

Velichety, S., & Ram, S. (2020). Finding a needle in the haystack-recommending online communities on social media platforms using network and design science. Forthcoming, Journal of Association for Information Systems.

Venable, J., & Baskerville, R. (2012). Eating our own cooking: Toward a more rigorous design science of research methods. Electronic Journal of Business Research Methods, 10(2), 141-153.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. Nature, 393(6684), 440-442.
Whelan, E., Teigland, R., Vaast, E., & Butler, B. (2016). Expanding the horizons of digital social networks: Mixing big trace datasets with qualitative approaches. Information and Organization, 26(1-2), 1-12.
Winter, R. (2013). Towards a framework for evidence-based and inductive design in information systems research. In European Design Science Symposium (pp. 1-20). Cham: Springer International Publishing.
Xu, J., Wang, G. A., Li, J., & Chau, M. (2007). Complex problem solving: identity matching based on social contextual information. Journal of the Association for Information Systems, 8(10), 31.
Zhang, G., Li, Z., Huang, J., Wu, J., Zhou, C., Yang, J., & Gao, J. (2022). efraudcom: An e-commerce fraud detection system via competitive graph neural networks. ACM Transactions on Information Systems, 40(3), 1-29.
Zhang, K., Bhattacharyya, S., & Ram, S. (2016). Large-scale network analysis for online social brand advertising. MIS Quarterly, 40(4), 849-868.

Online Appendix:

The online appendix is available at the following link: http://bit.ly/4jKSBSa

Acknowledgement:

This research is supported by CHERISH knowledge co-creation hub at the Department of Information Systems, University of Agder, Norway

5 CORPORATE DATA DIOGENES: ANALYZING COMPANIES THAT BEHAVE LIKE DIGITAL ACCUMULATOR

Ariel I. La Paz
University of Chile
lapaz@fen.uchile.cl

Josué Salinas
University of Chile
jsalinasc@fen.uchile.cl

David López
University of Chile
dlopezc@fen.uchile.cl

Abstract
Data and technological capacity drive individuals and companies to accumulate digital assets—documents, photos, emails, apps, and more—often without clear purpose, a behavior termed "digital hoarding" or "data sprawl." While individuals hoard for personal reasons, organizations pile up data aiming to collect customer insights, boost efficiency, and inform the business strategy. Yet, without clear data strategies, these repositories become inefficient, incurring high financial and non-financial costs, and hindering knowledge creation and retrieval. Unlike visible physical hoarding, corporate digital accumulation often goes unnoticed until critical needs arise. This article examines whether companies are crafting effective data strategies, exploring how unstructured data growth complicates operations and escalates maintenance costs for unused technology. The study advocates for robust data curation and governance to convert digital assets into actionable value.

Keywords
Data hoarding, Data Governance, Data Strategy, Ontological framework

1 Introduction
Most of us save and store thousands of digital pictures in our cell phones, cloud storage spaces, stick memories, old laptops, and digital cameras, with the hope to later do something with them. Similar behavior with old emails, mobile apps, accounts in different websites, songs, memes, videos, contacts and followers on the social networks, and some hardware too, that keep accumulating forgotten technological “assets”. The literature has classified individual’s behavior about digital accumulation, recognizing motivations and personality characteristics in the physical and the digital spaces (Vitale et al, 2018), identifying as well, some personal and organizational costs associated with digital hoarding, such as time to search and recover data, use of IT resources, storage used in duplication of files, security risks, etc. (McKellar et al, 2020). A study published in 2012 reported the impact of data in terms of useful lifetime, effectiveness, productivity, and more, highlighting that the indiscriminate accumulation of data negatively impacts on data understanding, structured knowledge generation, lack of validity and reliability of data as support for decision making, and the capacity to share knowledge (Gormley, & Gormley, 2012). The unpurposeful accumulation in the physical world is usually diagnosed as a mental disorder, similar

to the Diogenes syndrome, when individuals collect and keep objects and trash without purpose or limits, while in the digital space has been referred to as “digital hoarding” and “data sprawl” (McKellar et al, 2024; Seshadri, 2021). Even though companies are different from individuals, the lack of appropriate technological planning and data strategies facilitated the appearance of symptoms like those of hoarding.

Companies have it even worse about the data curation (capture-storage-retrieval-management-exploitation). Many organizations embarked in the construction of data warehouses, data lakes (or data swamps?) and cloud storage spaces to collect and save the “new gold/new oil” data (The Economist 2017), with the hope to better understand their customers, improve process efficiency, monitor organizational mission, and vision achievement, assess organizational maturity, and to track a long list of indicators. Others do not even try, just keeping the data spread all over user’s emails, personal virtual drives, or hard copies kept somewhere in the furniture or a safe box. Both extreme approaches could be wrong if they do not follow a well-crafted data strategy and governance. One recurrent response to the data storage problem is to 'buy more space' (virtual or physical). A few other times, the reaction is to strategically think about it, define objectives and purposes, 'clean the house', and be realistic about the usability and value of the 'digital treasures', or even better, to implement some data framework or best practices for data curation and governance. Maybe, the virtual nature of the storage spaces makes it invisible for the own company to realize the mess and difficulty to extract value out of the data, or just because data actually works (assuming inefficiencies) for the daily operations and transactions keeping the shop running, the data issues remain unnoticed. In the analogy, the disorder and unproductive accumulation is evident for the neighbors and family of a person affected by the *Diogenes Syndrome*, because it is visible and disturbing, but at the corporate level, the symptoms may be a bit different and remain unknown and undetected, until a new report or specific information is required. This article explores the problems related to the companies’ strategy towards data, aiming to answer: ***are companies designing and implementing data strategies?*** By answering the research question, the authors seek to highlight the importance of corporate data strategies that align its use to the business model and open the discussion about extreme perspectives on corporate data strategy, analyzing common practices applied by companies today.

2 State of the art

It was found an incipient body of literature about data hoarding and accumulation, but there seems to be a much faster implementation of data and technological assets in organizations, than theories guiding its investment and exploitation. The unpurposeful data accumulation has been studied in the literature from a psychological and sociological perspective, mainly at the individual level, or in the company context, still focused on individual’s behavior at work (Sedera et al, 2022). Another research stream considers data collection and accumulation in the personal domain (Sillence et al, 2023), and lastly, the topic of technostress has gained attention, in part related to the behavior of individuals over information overload (Tarafdar et al, 2011). A case study, documented in the context of one company, the perceptions of users about digital data, assessing what are the data that people use and keep over the years (Vitale et al, 2018). Along with the “hoarder” profile, the authors identified “minimalist” users who avoid saving data they regarded as unnecessary and keep only what they consider necessary for their daily performance. More recently, McKellar et al (2020) described characteristics and motivations of individuals accruing data, labelling them as driven by anxiety, compliance, disengagement, and collection, and attempted to explain such

behaviors based on the organizational culture, individuals' awareness, types of data managed by companies, and the degrees of responsibility that individuals and organizations convey in data management.

Digitally storing data is not perceived as a huge financial cost, but just like it occurs to physical accumulators when attempting to bring out an object, the cost of recovery and retrieval of valuable information and knowledge increases along with the unstructured storage, obsolescence and data aging. Then, it can be claimed that the effective costs of data hoarding at the organizational level are not only the direct financial costs, but also indirect costs (time to manage and exploit, inefficiency in the use of resources and poor performance, or the under-exploitation of business value), and qualitative costs (erratic strategies, disengagement, concentration of power over data ownership, and reality distortions about the value of data), which many times are unnoticed. The Information Governance Maturity Index Report 2021, revealed that 45% of information professionals defined their profile as 'record managers', evidencing the use of skilled labor hours on low value tasks, such as keeping records and controlling backups (ARMA, 2024), but for which the curation processes could simplify or partially automate, if data strategies were well defined.

2.1 Data Strategy

Since 2005, the importance of data strategy has grown due to challenges in managing records, including dirty, redundant, and inconsistent data, difficulties integrating diverse data sources, poor reporting performance, data unavailability, and lack of accountability (Adelman et al., 2005). To address these, data strategy has evolved to encompass key dimensions such as data modeling layers, business intelligence, measuring data value through ROI, and defining roles and responsibilities within organizations (Brous et al, 2020; Pothineni, 2023). The exponential growth of data and advancements in processing and storage technologies have increased the complexity of these dimensions, as organizations rely on data curation for decision-making. A robust data strategy aligns with corporate objectives through three core components (SAS, 2017). First, is the strategic objective determined in the data strategy, which aims to support the organizational purposes, whether it is generating revenue, increasing profitability, reducing costs, or positioning and exploiting business niches. The second component is the scope, which determines the types of data to be curated, and the third component is the competitive advantage to be developed in the organization, using data, information and knowledge. In parallel, the integration of structured, semi-structured, and unstructured data along with processing speed, data volume, and deriving value from these data gave rise to the concept of big data. This new data paradigm led companies to pursue multiple technological initiatives related to big data, often losing sight of the vision and strategy to leverage the potential of these data (Mazzei and Noble, 2017).

The data strategy is never isolated and is constantly permeated by technological and organizational changes. Its concept evolves and is described as the strategy to organize, govern, analyze, and deploy the information assets (DalleMule & Davenport, 2017). Identifying information assets implies that economic benefit can be derived and that must also be secured. The balance of offensive strategies, exploiting data to achieve important organizational objectives, and defensive strategies to properly safeguard assets from potential threats is also required (DalleMule & Davenport, 2017). Thus, data strategy can be defined as a coordinated approach to managing multiple data domains to support strategic business goals (Fleckenstein et al., 2018).

Data strategy encompasses dimensions evaluated by their depth (specificity or complexity) and breadth (application across organizational units). These dimensions include data architecture, governance, quality, privacy and security, integration, metadata management, and master data management. While data governance and lifecycle management are often tied to operational execution (DalleMule & Davenport, 2017), they are also integral to strategy formulation (Fleckenstein et al., 2018). Ultimately, a well-defined data strategy emerges through the development and integration of these dimensions.

2.2 Data Governance frameworks

In recent decades, the emergence of data management governance frameworks has been driven by new technologies and evolving challenges in data curation, protection, and utilization. We revised some of the most important practice-oriented literature, as they have a direct impact on business architecture and processes for they are used as best practices, implementation standards and compliance strategies. Published studies compile and classify several data governance and management frameworks (Marcucci et al, 2023; Ekundayo et al, 2023; Chukwurah et al, 2024), from which we discuss DAMA and COBIT for their extensive prevalence in public and private organizations, but the state of the art still needs to further develop and test governance mechanisms, scope of data governance, its antecedents and consequences (Abraham, et al, 2019).

DAMA - DMBOK (Data management body of knowledge) is a comprehensive reference guide that coordinates industry practices, concepts, and approaches for effective data management. DAMA outlines the key functions and processes involved in managing data assets across an organization. DMBOK provides a structure and methodology for establishing, executing, and improving organizational data management practices, and serves as an essential toolkit for professionals aiming to ensure data accuracy, usability, integrity, security, and compliance. DAMA is useful to emphasize the strategic importance of data as a critical asset within the organization, and is helpful to foster a data management strategy aligned with broader business goals. The framework encourages the adoption of best practices through a structured governance approach, encouraging a culture of continuous improvement and accountability in data management.

Control Objectives for Information and Related Technologies (COBIT), incorporates a specific governance/management objective dedicated to Data Management, called “Managed Data - APO 014”. This objective underscores the necessity of managing data as a valuable asset, ensuring its availability, usability, consistency, integrity, and security across the organization and encompasses the whole data lifecycle. As every other management objective in this framework, COBIT will define not only its activities, but also responsibilities, information flows, necessary enterprise architecture, responsibilities and needed skills for achieving the best organizational results while applying Data Management principles and practices. APO 014 purpose is to facilitate the establishment of appropriate policies, standards, and procedures that govern the entire lifecycle of data. Not only the process itself, but also the extended tools provided within COBIT 2019 are designed to effectively leverage this Data Management objective. These tools include maturity models, process capability assessments, and governance best practices that help organizations measure and improve the management of their data assets.

2.3 Construction of an ontological framework of data strategy

Ontological frameworks have been designed and applied to study complex phenomena such as IT requirements engineering (Vasquez and La Paz, 2019), e-Commerce (La Paz et al, 2015), innovation (Cancino et al, 2018), and more. Such technique is a qualitative approach that identifies representative dimensions of a problem or topic under study, and describes the dimensions with the appropriate taxonomies, which can be combined to form sentences that identify specific aspects and mechanisms of the problem at hand. It is worth noting that in the ontological methods the dimensions are not absolute, but practical conceptualizations of the elements, and could be expanded by adding more dimensions and categories or simplified by reducing the ones presented here for the study of data strategy. Also, the ontological frameworks method aims at the visualization of an ample scope of the phenomena, which occurs when the reader combines any number of elements and make sense of a particular instance of the problem. In this regard, the purpose of the ontology is to identify as many data strategies as it can be differentiated, to later test its validity and usability with empirical data. From the literature revised, business theories, and good practices in the formulation of strategies, we identified five well-known and widely used dimensions to introduce the design of the Data Strategy Ontology, namely: 1) *Dependence*, 2) *Governance*, 3) *Curation Process*, 4) *People*, and 5) *Technological capacity*.

Dependence refers to how intensive in the use and exploitation of data is the business model. Then, business strategies may be intensive on data curation for business value, or non-intensive on data for best business performance. *Governance* means the policies and procedures implemented in an organization to manage and exploit data. The data governance for a company can be defined or non-defined. *Curation process* is the series of actions, activities, and steps taken to achieve a particular end, and the taxonomy used in our ontology is a three components taxonomy indicating a curation process could be optimized, deficient, or inexistent. In the context of data strategy, *People* need to learn and develop knowledge, skills, and competences for the right use of data. This dimension will be described with a simple taxonomy of trained and untrained people/users. *Technological capacity*, in the context of data strategy, can be regarded as an enabler of data curation, or a limiting condition for value creation from data.

Connecting words/phrases have been added to the framework to facilitate the instantiation and reading of the 48 possible combinations, calculated as the product of the number of categories in the dimensions (2x2x3x2x2=48). Appendix A presents the complete list of combinations, and after Table 1, three instances of the Data Strategy Ontology are illustrated.

Dependence		**Governance**		**Curation Process**		**People**		**Technological capacity**
Data intensive	*strategy with*	Defined	*Governance,*	Optimized	*Processes,*	Trained	*People, and*	Enabling *technology*
Non-intensive		Non-defined		Deficient		Untrained		Limiting
				Non-existent				

Table 1. Ontology of Data Strategy

1. *Data intensive strategy with defined governance, optimized processes, trained people, and enabling technology.*
2. *Data intensive strategy with non-defined governance, deficient processes, trained people, and limiting technology.*
3. *Non-intensive strategy, with non-defined governance, non-existent processes, untrained people, and limiting technology*

3 Validation of the ontology with empirical data

The assessment and evaluation of a data strategy can be a complex issue, given the fact that data is not classified just as good/bad, useful/unuseful or other binary classifications, but a complex multidimensional construct. Such construct later produces (or not) information, knowledge, and value as the product of resources and practices for data curation and exploitation. To answer the question whether companies are designing and implementing data strategies, this paper first elaborated an ontological framework to describe and systematically analyze data strategies, secondly proposes a questionnaire to collect empirical data of companies and business units regarding their existence and use of strategies, and thirdly explores the results.

3.1 Questionnaire design

To explore data strategies and validate the ontology, a questionnaire was designed using frameworks for data governance and the Deloitte insight-driven organization framework. The questionnaire presents six sections aiming at the collection of empirical data from organizations interested in the assessment of their data strategies. The first dimension collects information on the respondents' profiles, capturing elements such as age, gender, academic degrees obtained, job location, industry, and details about their position in the organization.

The questionnaire evaluates five key dimensions—Data Strategy, Data Governance, Data Processes, People, and Technology—to operationalize and gather insights on data management. The first dimension assesses the clarity of data purpose, the rationale for data storage, resource availability, and considerations like security, privacy, and regulatory compliance. This dimension is critical for understanding how organizations derive value from data by aligning with strategic objectives and adhering to governance principles (Günther et al., 2017). The data governance dimension, examines whether the organization has a defined governance structure, including designated roles or positions responsible for establishing standards, procedures, policies, and incident-response plans for data-related issues.

The remaining dimensions—Data Processes, People, and Technology—are widely employed in studies relating to data usage (Davenport, 1998). To gather information on these dimensions, the Deloitte Insight-Driven Organization framework was the baseline, with adaptations made to identify elements linked to data storage, in contrast to the original predominant focus on data analytics. The Processes dimension collects information on clarity regarding data usage, risk management, and the speed at which data are leveraged. The People dimension addresses leadership, talent management, and the cultivation of a data-driven culture. Finally, the Technology dimension aims to identify the definition of a data architecture, storage structures, capacity management, and operational continuity. The questionnaire is presented in Appendix B.

3.2 Data collection

The instrument was subjected to a pilot test with the members of the research team and invited scholars to ensure clarity, contextual alignment, and relevance of the items. The questionnaire was administered on an online survey platform and shared with a purposive sample of individuals in positions of IT professionals and data specialists in an array of roles, hierarchical levels, and industries. Part of the sample was contacted from alumni records and professional digital communities of interest in data analytics and data governance. The database consists of 66 former postgraduate students, 612 members of a professional IT social network groups, and was also

shared across two professional LinkedIn forums, which collectively comprise 216 members as of the submission date. From the database, 60 participants began the survey, but 19 of them were considered valid responses and analyzed in the final dataset, following the exclusion of those who left more than 20% of questions unanswered and the identification of duplicate entries based on IP address. Invitation to participate was voluntary and based on non-probabilistic purposive and convenient sample, with no incentives provided. The data collection remained open for 17 days, during which participants were informed and reminded of the study. All responses were treated with strict confidentiality and anonymity, in accordance with standard practices for survey-based research.

4 Results

19 complete responses to the questionnaire were obtained from a sample of professionals in their early careers (74% younger than 40 years old), mostly masculine with 84% of males. Half of our sample consisted by data analysts (52%), and the rest declared to be area leaders, managers, supervisors and director or executive officers. Areas of economic activity sectors represented in the sample are communications and IT, construction, entertainment, mining, educational services, financial services, consulting, and logistics and transportation. As a test for the questionnaire, Cronbach alpha were calculated per dimension, obtaining the following values: *Dependence 0.85; Governance 0.89; Curation Process 0.91; People 0.88; Technological capacity 0.93.* Although these values indicate consistency of the instrument, must be cautiously interpreted, due to the small sample size, and the imputation of mean values for partial responses.

The answers were grouped around the five dimensions presented in the ontology of Data Strategy, and the general (mean) scores per dimension are: *Dependence 3.4; Governance 3.2; Curation Process 3.1; People 3.0; Technological capacity 3.2.* The average scores in the Likert scale do not reveal current particular strategies by the organizations in the sample, but for a better description and explanation of the multidimensional nature of the data strategies, the scores per dimension for each observation were used to classify the responses with the taxonomies included in the ontological framework and to associate each organization with a strategy archetype.

In the *Dependence* dimension, scores above 4 were considered as Data Intensive organizations, and below 4 as Non-intensive. For the *Governance* dimension, values 3-5 considered the organization as with a Defined data governance, while less than 3 indicated a Non-defined governance. The *Curation process* considered to be Optimized for values above 4, Deficient less than 4 but more than 2, and Inexistent below that. *People* was regarded as Trained for the data use and curation in companies with an average over 3, and Untrained for scores below that number. Finally, Technological capacity considered to be Enabling when average scores in the dimension surpassed the value 3, and to be Limiting for lower values. Figure 1 presents a flow chart, from which reveals that most companies represented in the sample are non-intensive in the data strategy, have not defined the governance, present deficient curation processes, have trained people, but limiting technology.

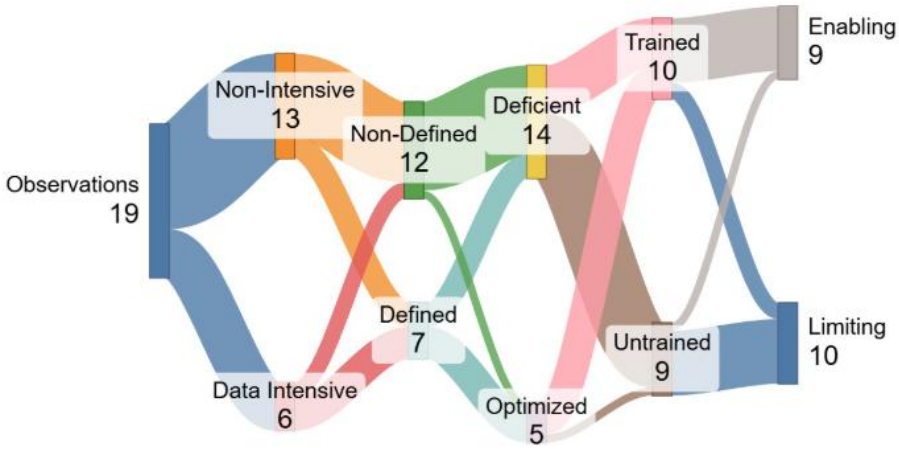

Fig 1: Flow chart of data strategy components

Table 2 further details the data strategy composition for each organization in the sample, including the column "strategy archetype". Strategy Archetypes represent a series of definitions and organizational practices about data. Based on the ontological instances and the responses that validate the use of the ontology to describe the empirical definitions and the questionnaire to operationalize the data collection, this study explores three archetypes to describe what organizations are doing regarding their data.

ID	Dependence	Governance	Curation Processes	People	Technological Capacity	Strategy Archetype
1	Non-intensive	Defined	Deficient	Trained	Enabling	*Data Strategy & Governance*
2	Data Intensive	Defined	Optimized	Trained	Enabling	*Data Strategy & Governance*
3	Data Intensive	Defined	Optimized	Trained	Enabling	*Data Strategy & Governance*
4	Data Intensive	Defined	Optimized	Trained	Enabling	*Data Strategy & Governance*
5	Data Intensive	Defined	Optimized	Trained	Enabling	*Data Strategy & Governance*
6	Non-intensive	Defined	Deficient	Trained	Enabling	*Data Strategy & Governance*
7	Data Intensive	Non-defined	Deficient	Trained	Enabling	*Just in case*
8	Non-intensive	Non-defined	Deficient	Trained	Limiting	*Just in case*
9	Non-intensive	Non-defined	Deficient	Trained	Enabling	*Just in case*
10	Non-intensive	Non-defined	Deficient	Trained	Limiting	*Just in case*
11	Non-intensive	Defined	Deficient	Untrained	Limiting	*Just in case*
12	Non-intensive	Non-defined	Deficient	Untrained	Limiting	*No data strategy*
13	Data Intensive	Non-defined	Deficient	Untrained	Limiting	*No data strategy*
14	Non-intensive	Non-defined	Deficient	Untrained	Limiting	*No data strategy*
15	Non-intensive	Non-defined	Deficient	Untrained	Limiting	*No data strategy*
16	Non-intensive	Non-defined	Deficient	Untrained	Limiting	*No data strategy*
17	Non-intensive	Non-defined	Deficient	Untrained	Limiting	*No data strategy*
18	Non-intensive	Non-defined	Deficient	Untrained	Limiting	*No data strategy*
19	Non-intensive	Non-defined	Optimized	Untrained	Enabling	*No data strategy*

Table 2. Ontological mapping of responses for data strategy description

Data Strategy and Governance

Having available data for decision making processes is important at the strategic level (top to bottom), at the operational base (bottom up) and across the functional areas. The key is 'available', meaning a set of properties that make it valuable (opportune, useful, efficient, economic, reliable, etc.). Achieving the status of efficient and meaningful data repositories and methods for its exploitation is not casual or random, but requires dedication of resources in business know-how and the definition of how dependent a company will be on data, designing a proper governance structure, planning and managing the processes for data curation, acquiring and implementing robust technology, communicating the strategy, and recruiting and training the right people for specific job positions. All of these (and maybe other forces) must be aligned and coordinated, to produce the expected value and synergy.

The "Just in Case" data strategy

A second archetype involves organizations that create and store data beyond transactional and operational needs, adopting data exploration and visualization tools. However, their approach to data strategy, accountability, and governance remains embryonic. This often leads to "Data Diogenes Syndrome," where companies collect and hoard data, investing heavily in technology, personnel, and tools—such as data centers, lakes, or warehouses, and hiring data analysts—to enhance decision-making or explore emergent strategies. These efforts may stem from a desire to keep pace with competitors, but they often lack a clear purpose, an efficient data architecture, or effective methods to transform data into actionable insights for strategic, tactical, or operational decisions. Implementing tools and infrastructure before defining a robust data strategy is risky, as it may fail to deliver the anticipated value from these investments.

The 'No-Data-Strategy' problem

The absence of a data strategy is a common issue in many organizations. This does not imply a lack of data but rather the absence of a purposeful, long-term approach to managing and leveraging it. For example, many companies use information systems and databases for transactional purposes—such as processing payroll, taxes, supplier payments, or maintaining accounting, employee, and inventory records. However, beyond meeting legal compliance requirements, data is often archived and forgotten or lost entirely. Other internal processes, such as sales reporting, marketing plans, production forecasting, or IT portfolio management, rely on fragmented micro-databases and worksheets shared via email or cloud storage. These serve short-term tasks without a broader strategy to extract ongoing value.

In such organizations, accountability on data management and data governance are unknown terms, and a simple task such as 'creating a report of the evolution over the last 2 years of incentives paid to employees', would take several hours or days collecting past data, cleaning and organizing records, processing and validating patterns to finally arrive to a one-time-use new report. The hypothetical example provided may not be a business problem, unless the type of processed data is a required response to an urgent mandatory request, or a key piece of information to make a relevant and opportune business decision. Unfortunately, these types of situations are becoming

more and more frequent, stressing organizations and frustrating executives with the long data processing times.

5 Conclusions and limitations

The Diogenes syndrome is an abnormal behavior of individuals, affecting the personal quality of life of the person suffering from it, and for others surrounding him/her. The present article used the analogy to understand why are companies also behaving as hoarders with data, and how the lack of data strategy and governance affects their operations and results. The causes of disorders in companies can be many and based on frameworks for IT/data management and frameworks for governance the authors proposed an ontological framework that can assist in the description and assessment of the current state of an organization regarding the data assets. Exploratory work with empirical data revealed that more companies either not have a data strategy, or the data strategy is insufficient, leading to inefficiencies such as data hoarding, inconsistencies between the strategic definitions and the processes for curation and training of users. The ontological framework introduced was useful to systematically analyze a small sample, and to condense an emerging literature into a system of models and could be a basis for training robust analytical algorithms capable of analyzing large sets of data.

This study has several limitations related to sample size and representativeness. While the survey aims to capture data on the described phenomena, the limited number of responses restricts the generalizability of findings across diverse countries or industries, necessitating cautious interpretation. Additionally, self-selection bias may occur, as respondents with a particular interest in data strategy and governance could skew results, potentially overrepresenting certain archetypes. Furthermore, the rapid pace of technological advancements and regulatory changes can influence outcomes, potentially altering the defined archetypes over time. These dynamics highlight opportunities for future research to explore the factors driving changes in data strategies and identify elements critical for maintaining business security and continuity.

References

Abraham, R., Schneider, J., & Vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International journal of information management,* 49, 424-438.

Adelman, S., L. T. Moss and M. Abai (2005). *Data strategy*, Addison-Wesley.

ARMA International. “Information Governance Maturity Index Report—2023.” April 2024.

Brous, P., Janssen, M., & Krans, R. (2020, April). Data governance as success factor for data science. In *Conference on e-Business, e-Services and e-Society* (pp. 431-442).

Cancino, C., La Paz, A., Ramaprasad, A., & Syn, T. (2018). Technological innovation for sustainable growth: An ontological perspective. *Journal of Cleaner Production*, 179, 31-41.

Chukwurah, N., Ige, A. B., Adebayo, V. I., & Eyieyien, O. G. (2024). Frameworks for effective data governance: best practices, challenges, and implementation strategies across industries. *Computer Science & IT Research Journal*, 5(7), 1666-1679.

DalleMule, L., & Davenport, T. H. (2017). What’s your data strategy. *Harvard business review*, 95(3), 112-121.

Davenport, T. H. (1998). Working knowledge: How organizations manage what they know. NewYork Harvard Business School.

Ekundayo, T., Bhaumik, A., & Chinoperekweyi, J. (2023). Identifying the core data governance framework principle: a framework comparative analysis. *Organization Leadership and Development Quarterly*, 5(1), 30-53.

Fleckenstein, M., Fellows, L., & Ferrante, K. (2018). Modern data strategy (pp. 63-76). New York, USA: Springer International Publishing.

Gormley, C. J., & Gormley, S. J. (2012). Data hoarding and information clutter: The impact on cost, life span of data, effectiveness, sharing, productivity, and knowledge management culture. *Issues in Information Systems*, 13(2), 90-95.

Günther, W. A., Mehrizi, M. H. R., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *The Journal of Strategic Information Systems*, 26(3), 191-209.

La Paz, A. I., Ramaprasad, A., Syn, T., & Vasquez, J. (2015). An ontology of E-commerce-mapping a relevant corpus of knowledge. *Journal of Theoretical and Applied Electronic Commerce Research,* 10(2), i-ix.

Marcucci, S., Alarcon, N. G., Verhulst, S. G., & Wullhorst, E. (2023). Mapping and Comparing Data Governance Frameworks: A benchmarking exercise to inform global data governance deliberations. *arXiv preprint arXiv:2302.13731*.

Mazzei, M. J., & Noble, D. (2017). Big data dreams: A framework for corporate strategy. *Business Horizons*, *60*(3), 405-414.

McKellar, K., Sillence, E., Neave, N., & Briggs, P. (2020). There is more than one type of hoarder: collecting, managing and hoarding digital data in the workplace. *Interacting with Computers*, 32(1), 209-220.

McKellar, K., Sillence, E., Neave, N., & Briggs, P. (2024). Digital accumulation behaviours and information management in the workplace: exploring the tensions between digital data hoarding, organisational culture and policy. *Behaviour & Information Technology*, 43(6), 1206-1218.

Pothineni, S. (2023). The Impact of Data Strategy and Emerging Technologies on Business Performance. *International Journal of Business Strategy and Automation*, 4(1), 1-19.

Sedera, D., Lokuge, S., & Grover, V. (2022). Modern-day hoarding: a model for understanding and measuring digital hoarding. *Information & Management*, 59(8).

Seshadri, P. (2021). Data is the new oil and data sprawl is the new oil spill. *Dataquest*, Retrieved from https://www.proquest.com/trade-journals/data-is-new-oil-sprawl-spill/docview/2585439954/se-2

Sillence, E., Dawson, J. A., Brown, R. D., McKellar, K., & Neave, N. (2023). Digital hoarding and personal use digital data. *Human–Computer Interaction*, 1-20.

The Economist (2017) Fuel of the future; the data economy. (2017/05/06/, 2017 May 06). *The Economist*, 423, 17.

Tarafdar, M., Tu, Q., Ragu-Nathan, T. S., & Ragu-Nathan, B. S. (2011). Crossing to the dark side: examining creators, outcomes, and inhibitors of technostress. *Communications of the ACM*, 54(9), 113-120.

Vásquez, Jonathan, & La Paz, Ariel. (2019). The Knowledge Body of Requirement Engineering in IST Innovations: An Ontological Analysis. *Journal of Technology Management & Innovation*, 14(4), 78-84.

Vitale, F., Janzen, I., & McGrenere, J. (2018, April). Hoarding and minimalism: Tendencies in digital data preservation. In *Proceedings of the Conference on Human Factors in Computing Systems* (pp. 1-12).

Appendix A. List of combinations in the Data Strategy Ontology

1 Data intensive strategy with Defined governance, Optimized processes, Trained people, and Enabling technology
2 Data intensive strategy with Defined governance, Optimized processes, Trained people, and Limiting technology
3 Data intensive strategy with Defined governance, Optimized processes, Untrained people, and Enabling technology
4 Data intensive strategy with Defined governance, Optimized processes, Untrained people, and Limiting technology
5 Data intensive strategy with Defined governance, Deficient processes, Trained people, and Enabling technology
6 Data intensive strategy with Defined governance, Deficient processes, Trained people, and Limiting technology
7 Data intensive strategy with Defined governance, Deficient processes, Untrained people, and Enabling technology
8 Data intensive strategy with Defined governance, Deficient processes, Untrained people, and Limiting technology
9 Data intensive strategy with Defined governance, Non-existent processes, Trained people, and Enabling technology
10 Data intensive strategy with Defined governance, Non-existent processes, Trained people, and Limiting technology
11 Data intensive strategy with Defined governance, Non-existent processes, Untrained people, and Enabling technology
12 Data intensive strategy with Defined governance, Non-existent processes, Untrained people, and Limiting technology
13 Data intensive strategy with Non-defined governance, Optimized processes, Trained people, and Enabling technology
14 Data intensive strategy with Non-defined governance, Optimized processes, Trained people, and Limiting technology
15 Data intensive strategy with Non-defined governance, Optimized processes, Untrained people, and Enabling technology
16 Data intensive strategy with Non-defined governance, Optimized processes, Untrained people, and Limiting technology
17 Data intensive strategy with Non-defined governance, Deficient processes, Trained people, and Enabling technology
18 Data intensive strategy with Non-defined governance, Deficient processes, Trained people, and Limiting technology
19 Data intensive strategy with Non-defined governance, Deficient processes, Untrained people, and Enabling technology
20 Data intensive strategy with Non-defined governance, Deficient processes, Untrained people, and Limiting technology
21 Data intensive strategy with Non-defined governance, Non-existent processes, Trained people, and Enabling technology
22 Data intensive strategy with Non-defined governance, Non-existent processes, Trained people, and Limiting technology
23 Data intensive strategy with Non-defined governance, Non-existent processes, Untrained people, and Enabling technology
24 Data intensive strategy with Non-defined governance, Non-existent processes, Untrained people, and Limiting technology
25 Non-intensive strategy, with Defined governance, Optimized processes, Trained people, and Enabling technology
26 Non-intensive strategy, with Defined governance, Optimized processes, Trained people, and Limiting technology
27 Non-intensive strategy, with Defined governance, Optimized processes, Untrained people, and Enabling technology
28 Non-intensive strategy, with Defined governance, Optimized processes, Untrained people, and Limiting technology
29 Non-intensive strategy, with Defined governance, Deficient processes, Trained people, and Enabling technology
30 Non-intensive strategy, with Defined governance, Deficient processes, Trained people, and Limiting technology
31 Non-intensive strategy, with Defined governance, Deficient processes, Untrained people, and Enabling technology
32 Non-intensive strategy, with Defined governance, Deficient processes, Untrained people, and Limiting technology
33 Non-intensive strategy, with Defined governance, Non-existent processes, Trained people, and Enabling technology
34 Non-intensive strategy, with Defined governance, Non-existent processes, Trained people, and Limiting technology
35 Non-intensive strategy, with Defined governance, Non-existent processes, Untrained people, and Enabling technology
36 Non-intensive strategy, with Defined governance, Non-existent processes, Untrained people, and Limiting technology
37 Non-intensive strategy, with Non-defined governance, Optimized processes, Trained people, and Enabling technology
38 Non-intensive strategy, with Non-defined governance, Optimized processes, Trained people, and Limiting technology
39 Non-intensive strategy, with Non-defined governance, Optimized processes, Untrained people, and Enabling technology
40 Non-intensive strategy, with Non-defined governance, Optimized processes, Untrained people, and Limiting technology
41 Non-intensive strategy, with Non-defined governance, Deficient processes, Trained people, and Enabling technology
42 Non-intensive strategy, with Non-defined governance, Deficient processes, Trained people, and Limiting technology
43 Non-intensive strategy, with Non-defined governance, Deficient processes, Untrained people, and Enabling technology
44 Non-intensive strategy, with Non-defined governance, Deficient processes, Untrained people, and Limiting technology
45 Non-intensive strategy, with Non-defined governance, Non-existent processes, Trained people, and Enabling technology
46 Non-intensive strategy, with Non-defined governance, Non-existent processes, Trained people, and Limiting technology
47 Non-intensive strategy, with Non-defined governance, Non-existent processes, Untrained people, and Enabling technology
48 Non-intensive strategy, with Non-defined governance, Non-existent processes, Untrained people, and Limiting technology

Appendix B. Questionnaire for Data Strategy Description

Dimension	Metric	Field
Data Strategy	Likert 1 to 5	My organization/company has defined the purpose of data storage in its business model
Data Strategy	Likert 1 to 5	My organization/company has defined the purpose of data usage in its business model
Data Strategy	Likert 1 to 5	My organization sets its strategic objectives by taking into account the available data resources
Data Strategy	Likert 1 to 5	Business units identify their role in data usage
Data Strategy	Likert 1 to 5	Business units identify their role in data storage and data quality
Data Strategy	Likert 1 to 5	My organization/company incorporates privacy, security, and regulatory compliance of the stored data as a key objective
Data Governance	Likert 1 to 5	My organization has a defined data governance framework
Data Governance	Likert 1 to 5	There is a person/role/unit within the organization responsible for data management
Data Governance	Likert 1 to 5	My business unit has specific data quality standards and procedures to comply with
Data Governance	Likert 1 to 5	My business unit classifies and manages different types of data (e.g., public, confidential, personal)
Data Governance	Likert 1 to 5	There are clear policies and procedures for data access and sharing within your business unit
Data Governance	Likert 1 to 5	My business unit is aware of the laws related to data storage, protection, and security
Data Governance	Likert 1 to 5	Employees are trained on data governance policies and the importance of data security
Data Governance	Likert 1 to 5	My business unit has a data incident response plan (data loss or theft, security breaches, inconsistencies)
Process	Likert 1 to 5	My business unit has clear processes for data storage and usage
Process	Likert 1 to 5	My business unit has processes in place to manage new data storage and usage requests
Process	Likert 1 to 5	My business unit knows the data curation processes (modeling, visualization, transformation, and optimization)
Process	Likert 1 to 5	My business unit is clear about data extraction, standardization, and storage processes
Process	Likert 1 to 5	My business unit has appropriate risk management for data storage and usage
Process	Likert 1 to 5	My unit has monitoring (observability) and traceability applications for data flows up to storage
Process	Likert 1 to 5	In my business unit, data exploitation is streamlined
People	Likert 1 to 5	The business unit has leadership that promotes data management
People	Likert 1 to 5	The business unit has leadership that promotes careful and proper data storage
People	Likert 1 to 5	There is a data-driven decision-making culture
People	Likert 1 to 5	There are change management processes aimed at improving data usage
People	Likert 1 to 5	Employees are trained in the proper use of data
People	Likert 1 to 5	There is talent management for employees who demonstrate strong data practices
Technology	Likert 1 to 5	In my unit, there is a defined data architecture
Technology	Likert 1 to 5	My unit has a well-defined storage structure (folders, databases, backups, among others)
Technology	Likert 1 to 5	The data architecture allows for the exploitation of data and the derivation of value
Technology	Likert 1 to 5	The data architecture includes the necessary technological safeguards for data storage
Technology	Likert 1 to 5	There is capacity management in place linked to the projection of the business model
Technology	Likert 1 to 5	There are operational continuity plans and backups related to data

6. CYBERBULLYING DETECTION WITH MACHINE LEARNING & PSYCHOLOGY: A SYSTEMATIC REVIEW

Lubaina Navaid
Institute of Business Administration, Karachi
lnavaid@khi.iba.edu.pk

Faisal Iradat
Institute of Business Administration, Karachi
firadat@iba.edu.pk

Nazim Taskin
Boğaziçi University
nazim.taskin@bogazici.edu.tr

Belgin Metin
Boğaziçi University
belgin.metin@bogazici.edu.tr

Abstract
With the rise of communication technologies, cyberbullying has become a pressing problem. Victims of cyberbullying often find themselves struggling with the physical and psychological effects for many years to come, some even going as far as to take their own lives as a result. Thus, detecting and preventing cyberbullying is crucial. Machine learning and artificial intelligence models have been used for this purpose, however, there is room for improvement regarding considering context and training of these models on additional features. Taking into account the psychological traits and user profile of offenders to train detection can improve performance and help authorities in implementing better detection and prevention systems, thus protecting potential victims from being bullied on online platforms.

Keywords
cyberbullying, cyberbullying detection, cyber bullying, machine learning, deep learning

1. Introduction
The increased prevalence of social media and internet use has unfortunately led to a rise in cyberbullying - a form of bullying that occurs in digital spaces. Cyberbullying is defined as intentional, aggressive acts repeatedly carried out using information technology against victims who are unable to defend themselves (Smith, 2008), which builds on the definition of bullying itself as repetitive negative actions by one or more individuals over time (Olweus, 1993).

While numerous studies utilize machine learning and deep learning algorithms for cyberbullying detection, achieving significant results (e.g., Mehendale, Rajpara, & Shah, 2002; Balakrishnan & Kaity, 2023), there is a notable gap in research exploring the underlying psychology driving this behavior. When it comes to real world application, the datasets these models will be working on will also be largely imbalanced – with cyberbullying instances making up a lower percentage of the data while non-cyberbullying instances make up the majority. While oversampling and synthetic data can help reduce this issue in a research context, these techniques do not improve results in real world applications (Elor & Averbuch-Elor, 2022) and thus, existing models need additional information to more accurately detect cyberbullying instances.

In that vein, although extensive psychological research has investigated the factors contributing to bullying behavior in general, there is a need to understand the specific psychological traits, characteristics, and attributes common among cyberbullies that motivate their online actions. This paper aims to address this gap by systematically reviewing existing literature on two key areas:

1. The application of machine learning in automatic cyberbullying detection and
2. The psychological factors contributing to cyberbullying behavior.

This review aims to connect these two areas by exploring how an understanding of cyberbully psychology can inform and improve detection methods.

2. Background

Due to its digital nature, cyberbullying can take place around the clock and is not limited to the bullies and victims being within the same geographical location. Common behaviors associated with cyberbullying involve repeated targeting of victims, sending threatening or harassing messages (Peebles, 2014), posting embarrassing photos or videos of the victims (Hinduja & Patchin, 2010), deliberately excluding victims from online groups or activities, badmouthing victims (Peebles, 2014) and stalking. In cyberspace, these activities catch on rapidly, and provide bullies constant access to the victims, thus posing several consequences.

Cyberbullying has severe consequences on the victims. These include physical as well as psychological consequences (Geel, Goemans, & Toprak, 2017). Some common psychological effects of cyberbullying include stress, anxiety (Li, Li, Li, 2018), worthlessness, depression (Wright, 2018), reduced self-esteem, suicide ideation (Kowalski, Toth, & Morgan, 2018), sleep disorders, post-traumatic stress disorder (PTSD) (Mateu, 2020), psychosis (Catone, Marwaha & Lennox, 2017), and in some extreme cases has even led to the victim taking their own life. Other cases of cyberbullying show that victims may turn towards drug abuse to cope with their victimhood, which further escalates to physical harm, behavioral regression and psychosocial challenges (Foody, McGuire, & Kuldas, 2019). Therefore, detecting instances of cyberbullying as early as possible is vital to intervene and protect victims and resolve the problem at its earliest.

With advancements in artificial intelligence (AI), the, automatic detection of text-based cyberbullying instances has become possible. Artificial intelligence can analyze large amounts of data to identify patterns that indicate bullying behavior, while NLP techniques allow these systems to understand the nuances of human language to accurately identify context and meaning. Using these tools, cyberbullying detection can, in theory, be carried out without the need for human supervision. This detection is useful because the large volumes of data generated on a daily basis by social platforms makes manual detection of such behavior impossible.

However, despite the benefit of using technology to carry out this process, concerns around the detection of cyberbullying exist. These lie mainly in the fact that this behavior itself is quite subjective and dependent on various factors, such as the three core aspects of cyberbullying. These are: repetition (rather than a one-time incident), intentionality (with the intent to bully, harass or intimidate), and power imbalance (where the victim is unable to sufficiently defend themselves) (Slonje and Smith, 2007). Each of these factors have to be present for an interaction to be considered cyberbullying. As of yet, most work in cyberbullying detection works to detect aggressive or harassing content that could very well be a one-time instance, rather than repetitive

behavior. Therefore, it may not constitute cyberbullying, but rather an individual and isolated case of abusive behavior. Similarly, there may be cases where the intention to harm may or may not exist, and a power imbalance may or may not exist, and therefore may not be classified as cyberbullying.
As such, this paper attempts to find the gaps in the existing research regarding cyberbullying detection, and how the psychology of bullying can be used to improve detection models in a way that informs the detection process.

3. Methodology

A systematic literature review (SLR) involves identifying and evaluating high quality studies within a given research domain while adhering to well-defined criteria and employing a formal approach that addresses specific research questions. This section goes over the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) model that guided the selection of relevant papers.

3.1 Research Questions

This review aims to answer the following research questions (RQs):

RQ 1 - What machine learning and deep learning algorithms are commonly used for detecting cyberbullying in social media interactions?
RQ 2 - What psychological traits or patterns do experts associate with cyberbullies?
RQ 3 – Can psychological traits be used as a training feature to help improve cyberbullying detection models?

Therefore, we will identify the work that has been done in the field and what gaps currently exist. We will also understand the extent to which cyberbullying psychology has been implemented in assessing whether an instance is classified as cyberbullying or not, and how it can be used to develop models for this purpose.

3.2 Search Strategy

This review aims to answer questions that lie within the overlap of two domains, i.e., information systems and psychology. As such, the papers reviewed followed a different strategy for each domain. The first research question (RQ1) is concerned with information systems. To answer this, we looked at two main academic databases focused on information technology, IEEE Xplore, and Web of Science. Publications selected were done between a 10 year period from 2013 to 2023, as most developments regarding machine learning and deep learning have been made over the past decade (Cioffi & Travaglioni, 2020). The initial search was carried out using specific key terms (‘machine learning’ OR ‘deep learning’ AND ‘cyberbullying’ OR ‘cyber bullying’ AND ‘online’). This was to make sure models using both types of algorithms would be included, and cyberbullying instances that took place on online platforms were looked at. Boolean operators were used to enhance the search and gain more literature related to the study objectives. Forward and backward search based on the studies was part of this process. The search for this domain yielded a total of 386 articles across the databases selected.

For RQ2, the focus was on the psychology of bullying, which falls within the domain of psychology. For this, the time period searched was reduced to eight years, from 2015 to 2023 and

only review papers were considered. This is because cyberbullying has increased dramatically since 2015 and is on the rise due to the increased use of social media (Renee, Lynwood, & Sean, 2016). Two databases were referred, ScienceDirect and Web of Science, where more articles relevant to the field of psychology could be found. The search was carried out using keywords related to the field ('psychology' OR 'psychological traits' OR 'traits' OR 'characteristics' AND 'bullying' OR 'cyberbullying'). This is to make sure all papers that go over the psychology of bullying and mention psychological traits will be included. Table 1 gives a summary of the search strategy employed.

Since these two research questions address different fields of study (with RQ1 focusing on information systems and RQ2 focusing on psychology), the search strategy employed for each was different. This was to make sure that relevant information was gathered to answer each research question. Using this information, we can answer RQ3, which looks at how to combine the two fields to improve cyberbullying detection processes.

Focus	**RQ1**	**RQ2**
Domain	Information Systems	Psychology
Databases Searched	IEEE Xplore, Web of Science	ScienceDirect, Web of Science
Time Period	2013-2023	2015-2023
Search Terms	'machine learning' OR 'deep learning' AND 'cyberbullying' OR 'cyber bullying' AND 'online'	'psychology' OR 'psychological traits' OR 'traits' OR 'characteristics' AND 'bullying' OR 'cyberbullying'
Boolean Operators Used	Yes	Yes

Table 3: Selection Strategy

The selection process used aims to find studies based on inclusion and exclusion criteria that can better answer our research questions by providing relevant information.

3.2.1 Inclusion Criteria

- Journals, conference papers and early access articles published in the last 10 years from 2013 to 2023 to answer RQ1,
- Review articles published in the last eight years from 2015 to 2023 to answer RQ2,
- Papers that deal with cyberbullying detection carried out using machine learning or deep learning algorithms,
- Papers and articles written in English,
- Datasets taken from various platforms as long as the cyberbullying instances being looked at take place online,
- Cyberbullying detection studies focused on textual data.

3.2.2 Exclusion Criteria

- For RQ1, research that does not use main machine learning algorithms,

- Research that is subjective in nature (i.e., surveys, case studies or interviews),
- Research that focuses on cybercrime that is not specific to cyberbullying (e.g., racism, hate speech, etc.)
- For RQ2, research that focuses on external factors such as environmental or academic factors, and not the psychology of the individual involved in bullying themselves.
- Research that does not focus entirely on other types of bullying

3.3 Quality Assessment

For a systematic literature review, a quality assessment (QA) is imperative to make sure the studies selected for review are worthy and relevant. From the total 386 articles found in the domain of information systems across the three databases, papers were selected based on inclusion and exclusion criteria specified above. Papers that met the criteria were manually screened based on their titles and abstracts. In some cases, the full text had to be screened to ensure that the papers were relevant. The quality was also determined based on the methodology specified. A final 19 papers were selected for this study.

For RQ2, which deals with the domain of psychology, there were a total of 1597 articles on both databases. From here, duplicates were removed, and papers were selected based on the inclusion and exclusion criteria. A final five papers were selected based on the above-mentioned criteria. These were manually screened based on the full content of the articles to determine relevance. The QA phase and the trend of publications are illustrated in Figures 1 and 2 respectively. In Figure 1, n refers to the papers addressing RQ1 and m refers to papers addressing RQ2.

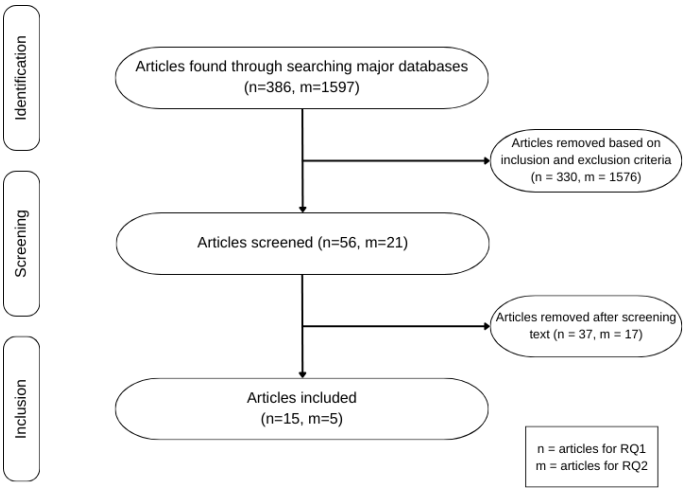

Figure 6: Selection Strategy

A majority of the articles reviewed were published in 2020. Of the six articles published in 2020, two were review articles on psychological traits of cyberbullies, while the other four were on the use of algorithms for detecting cyberbullying instances.

4. Discussion

The articles selected for this paper were examined to ensure that they could effectively answer the specified research questions.

4.1 The Use of Machine Learning in Identifying Cyberbullying

The articles selected for RQ1 all use either machine learning or deep learning models for detecting cyberbullying on social media platforms. Table 2 shows a summary of the articles focusing on cyberbullying detection included in this review.

Author(s)	Platform Assessed	Model Used	Performance & Results	Limitations Identified
Zhao, Zhou & Mao (2016)	Twitter	SVM	The model had precision, recall and f1score of 75.6%, 78.04% and 76.68% respectively	Context not considered and only a fixed dictionary used
Noviantho, Isa & Ashianti (2017)	Not Specified	Naïve Bayes, SVM	SVM performed better than Naïve Bayes	Limited hyperparameter tuning
Agrawal & Awekar (2018)	Formspring, Twitter, Wikipedia	Logistic regression, SVM, Naïve Bayes, Random Forest, CNN, BLSTM, BLSTM with attention, LSTM	Deep learning models perform better than traditional machine learning models.	No information on severity of bullying or the users
Al-Ajlan & Ykhlef (2018)	Not Specified	CNN	The algorithm had a 95% accuracy.	Model considered tweets with foul language but does not consider context.
Bu & Cho (2018)	Twitter	Character level CNN, LRCN	The model had an accuracy of 87.22%	Output is averaged out, but could possibly perform better if an ensemble method was used.
Rosa et al (2018)	GoogleNews, Twitter and Formspri ng	CNN, a hybrid CNN-LSTM and a mixed CNN-LSTM- DNN	CNN performed better than other models.	Dataset was imbalanced and model did not perform very well with minority class.
Hani et al (2019)	Formspring	Neural Networks, SVM	SVM – 89.87% accuracy; NN – 91.76% accuracy	Limited hyperparameter tuning
Balakrishnan, Khan & Fernandez (2019)	Twitter	Random Forest	The model had a precision of 91.5% and a recall of 91.6% when using personality classification.	Dataset was from a very specific community on one platform, so results are not very generalizable
Balakrishnan, Khan & Arabnia (2020)	Twitter	Naïve Bayes, Random Forest, J48	J48 model has a high accuracy after accounting for personality and psychological features	Dataset was too small to be generalizable
Islam, Uddin & Islam (2020)	Twitter, Facebook	Naïve Bayes, SVM, Decision Trees, Random Forest	SVM shows higher accuracy in comparison to other algorithms.	Does not consider other dimensions of cyberbullying or context of interaction
Muneer & Fati (2020)	Twitter	Logistic Regression, Light Gradient Boosting Machine, Stochastic Gradient Descent, Random Forest, AdaBoost, Naive Bayes, SVM	Logistic regression had the highest F1 score while SGD had the best precision	Study did not consider feature extraction techniques.
Yadav, Kumar & Chauhan (2020)	Formspring, Wikipedia	BERT	The model showed different results for each dataset.	Model does not consider any features besides text

Author(s)	Platform Assessed	Model Used	Performance & Results	Limitations Identified
Van, Huang & Inkpen (2020)	Instagram, Facebook, Pinterest, Twitter and YouTube	CNN	The model had an accuracy of 84.5% and F1-score of 86.7%	Model does not consider context
Alotaibi, Alotaibi & Razaque (2021)	Twitter	Multichannel model - BiGRU, CNN, Transformers	The model had an accuracy of 88%.	Dataset was too small.
Aldhyani, Al- Adhaileh & Alsubari (2022)	Wikipedia, Twitter	CNN-BLSTM, BLSTM	Binary classification assessment resulted in equal accuracy of both algorithms at 94%.	Context of interaction not considered
Dhingra et al (2023)	Twitter	SVM, Random Forest, Adaboost, MLP	SVM has the highest F1score of the assessed models	Limited to text data
Hudda, Misha & Kumar (2023)	Not Specified	BERT, kimCNN, SVM	The neural network algorithms performed better	Training dataset was too small
Kadam (2023)	Twitter	Decision Tree, Random Forest, SVM, MLP	Multilayer perceptrons had the highest recall, while decision trees had the most precision	Only considers the language aspect of cyberbullying and not the other dimensions.
Obamiyi et al (2023)	Not Specified	SVM, Naïve Bayes, KNN - Ensemble Model	The ensemble model used had a 95% accuracy	Model overlooks context in which language is used.

Table 2: Literature assessed for RQ1

In Table 2, we see some of the most commonly used machine learning algorithms. These include Logistic Regression, Naive Bayes, Support Vector Machines (SVM), Random Forest and Decision Trees, while commonly used deep learning algorithms include convolutional neural networks (CNN), long short-term memory (LSTM), bidirectional long short-term memory (BLSTM), bidirectional gated recurrent units (BGRUs) and other neural networks and transformers models such as BERT.

Both deep learning and machine learning models performed well, with most algorithms showing scores greater than 80%. However, it is difficult to compare performance between studies due to the difference in metrics used: some studies use accuracy as a measure, while others use F1 score, precision and recall. Similarly, each study also has differences in terms of the size, nature and diversity of the dataset used. Studies using datasets with high class imbalances or diversity in types of cyberbullying may perform differently from a studies that do not have these features. Therefore, while the performance of an algorithm may be high within the study, it is difficult to make a comparison with other studies around the same subject i.e. cyberbullying detection. However, since the algorithm performs well within the study, it can be assumed that it does manage to detect cyberbullying to a reasonable extent. In this review, the algorithms that have performed well over multiple studies have been considered.

Most datasets used do not have information on the severity of bullying or any information about the users. A few studies, such as those by Balakrishnan, Khan & Arabnia (2020) and Balakrishnan, Khan & Fernandez (2019) account for differences in personality and psychological features between users, but these are carried out on a much smaller dataset and are difficult to generalize.

Similarly, a few studies (Al-Ajlan & Ykhlef, 2018, Kadam, 2023) do not consider the context in which cyberbullying takes place, but rather focus on just the text itself. As a whole, it can be assumed that studies largely take into consideration only the linguistic aspect of cyberbullying, rather than the people involved and the specific social and environmental context in which it occurs. Through this review, we found that while detecting isolated instances of aggressive behavior has been addressed, there are still several gaps in the field. For example, as mentioned in Smith (2008), other dimensions of cyberbullying, such as intentionality, repetition and power imbalance are not assessed when detecting individual circumstances of aggressive behavior. This suggests that cases of disagreement between two users, cases where aggressive language is used in a sarcastic manner or as friendly banter may be incorrectly flagged as cyberbullying behavior.

Table 2 shows datasets that were extracted from various platforms for the detection of cyberbullying. These include Twitter, Instagram, Facebook, Formspring, and more. Among these, 68.4% used data from Twitter. Some datasets combined data from different platforms while others do not mention the platform used. This may suggest that instances of cyberbullying may be more prevalent on Twitter, which is a microblogging site with real-time interactions possible between users. However, collection of data from other social platforms (particularly messenger platforms where conversations are private, such as Discord, WeChat, Whatsapp, etc.) may show different results. While steps have been taken to reduce cyberbullying and aggressive behavior on the various platforms, it is not possible to completely eliminate the possibility. Furthermore, users of social media platforms have been using word camouflage to get past content moderation detection systems through the use of abbreviations, images containing bullying text, using symbols instead of letters, and other such methods (Huerta-Gracias, 2023). Cyberbullying detection systems do not, as of yet, incorporate these dimensions into their operations. While our review included studies analyzing textual data, a notable limitation observed across the reviewed literature is the predominant focus solely on textual data. This singular focus limits the comprehensive detection of cyberbullying, as real-world instances frequently involve multimodal elements such as images, videos, and audio (Dhingra et al., 2023) which were largely unaddressed in the included studies. These models also did not consider the element of repetition, where comments and statements that are otherwise harmless and not considered aggressive may be part of cyberbullying behavior if they are repeatedly being sent to a single user by another individual or group with malicious intentions. Furthermore, language on social platforms is generally unstructured and does not follow the rules of language. Not only does the social nature of the platform allow for new slang to be created, which further complicates detection since newer words cannot be categorized as positive or negative (Farzand, 2023).

Most detection models also do not differentiate in terms of severity. While hate speech, harassment and bullying can be identified, the classification as a positive or negative instance does not take into account the time period for which the bullying has been going on, nor the severity of the instance itself. As such, while machine learning algorithms do offer solutions in terms of identifying and addressing aggression, there are several gaps in the research that make cyberbullying detection efforts fall short. To create better detection and prevention systems, these gaps will have to be addressed. Furthermore, the introduction of large-language models (LLMs) in detection processes may have improved performance in comparison to traditional machine learning and deep learning models. However, LLMs, which are transformers-based models, will

still largely use the text of an interaction itself to identify cyberbullying instances, and therefore, the addition of psychological features may still be useful.

4.2 Psychological Features That Lead to Cyberbullying

To answer RQ2, several papers regarding psychology were also looked at to determine what traits are likely to be found in cyberbullies. To determine their usefulness in training models, it was necessary to first identify what the traits were. Table 3 shows a summary of the traits identified as being associated with cyberbullying behavior, as well as other factors to note regarding bullying.

From Table 3 it is evident that most cases of bullying and cyberbullying will have a few traits in common. For example, aggression, low empathy, depression, impulsivity, antisocial behavior and self-esteem problems seem to be common in bullies. It is also worth noting that many cases of cyberbullying involve a bully who has previously been victimized. This may be because many traits found in bullies are also found in victims. That is, the positive relationship between the perpetrator and victim subtypes regarding aggression, anxiety, depression, etc. results in individuals who have previously been victimized becoming bullies, and vice versa (Chan & Wong, 2015).

4.3 The Use of Psychological Traits to Improve Cyberbullying Detection

Despite the relationship between psychology and cyberbullying, there has been minimal research dedicated to integrating psychological features into cyberbullying detection models. This represents a promising field of further study into detection systems to improve their performance.

Current models (as seen in section 4.1) rely primarily on the lexical and semantic features extracted directly from the text of the interaction itself. While this is effective to a degree, these models will often struggle with the nuances of human communication such as implicit threats, irony, sarcasm and context-dependent malice. An understanding of the underlying psychological drivers can offer an advantage in overcoming these limitations. For instance, interactions with users displaying traits common in cyberbullies (as outlined in section 4.2) may carry a higher probability of being cyberbullying, though the linguistic cues may be ambiguous without this added context. These traits can be integrated into detection models in various ways. For example, by translating them into quantifiable features, they can be used for feature engineering within existing models. This might involve developing classifiers to infer user traits from their broader online behavior (e.g., posting patterns, topics of interest, interaction styles) or even through self-reported data in specific contexts. These inferred traits could then serve as additional input features for detection models. They can also be used for contextual augmentation, such as to provide each interaction with a risk score based on the traits a user exhibits. Though the incorporation of psychology into detection models is largely unexplored, having additional context could potentially improve the ability to differentiate between banter and aggression. This can reduce the identification of false positives and false negatives, and thus improve performance, and allow for proactive prevention and intervention strategies.

5. Conclusion

While communication technology has provided users all over the world with the opportunity to connect with each other in ways that were previously impossible, it has also allowed for the negative social behaviors of these individuals to spread over digital spaces. Particularly due to the

allowance of anonymity on the internet and thus, a reduced chance of facing consequences, this problem has become widespread and requires a solution. Machine learning and deep learning models, used alongside the psychological traits used to identify bullies can help identify instances of cyberbullying and determine methods for preventing them, which can address the problem of cyberbullying in digital spaces.

Authors	Psychological Traits Identified	Other Factors to Note
Mitsopoulou & Giovazolias (2015)	Low agreeableness, high extraversion, high impulsivity, anxious, moody, depressed, low empathy	Demographic factors like gender and age had a role to play in determining behavior and psychological traits.
Alvarez-Garcia, Garcia & Nunez (2015)	Impulsivity, hyperactivity, aggressiveness, low empathy, antisocial behavior, low self esteem, depression, extraversion, sensation-seeking	Sociodemographic factors, physical factors and environmental factors such as school and family circumstances also play a role in determining bullying behavior.
Swearer & Hymel (2015)	Pychopathic tendencies, antisocial behavior, anxiety, susceptibility to peer pressure	Other influences such as family, peers and school, as well as cultural/societal context also have a role to play in affecting bullying behavior.
Bashir, Rehman & Amin (2020)	Emotional problems, aggression, low self-esteem, insecurity, anxiety, low empathy, low emotional intelligence, high stress levels, antisocial behavior, depression, frustration, loneliness, jealousy, revengeful	Other factors to note include personal factors such as demographics, socio-cognitive factors, and environmental factors such as technological skill, family and household circumstances, academic performance, etc.
Marin & Guachichullca Guaman (2022)	Aggressiveness, antisocial behavior, impulsivity, high degree of anxiety and isolation, psychopathy Moral disconnection, lack of self-confidence, arrogance, jealous and envious	Risk factors that affect bullying behavior include personality, individual traits (such as psychological or emotional state), social, technological and economic factors.

Table 4: Assessed Studies for RQ2

.

References

Agrawal, S., & Awekar, A. (2018). Deep learning for detecting cyberbullying across multiple social media platforms. In G. Pasi, B. Piwowarski, L. Azzopardi, et al. (Eds.), *Advances in information retrieval* (pp. 141–153). Springer International Publishing.

Al-Ajlan, M. A., & Ykhlef, M. (2018). Deep learning algorithm for cyberbullying detection. *International Journal of Advanced Computer Science and Applications (IJACSA), 9*. https://doi.org/10.14569/IJACSA.2018.090927

Aldhyani, T. H. H., Al-Adhaileh, M. H., & Alsubari, S. N. (2022). Cyberbullying identification system based deep learning algorithms. *Electronics, 11*, 3273.

Alotaibi, M., Alotaibi, B., & Razaque, A. (2021). A multichannel deep learning framework for cyberbullying detection on social media. *Electronics, 10*, 2664.

Álvarez-García, D., García, T., & Núñez, J. C. (2015). Predictors of school bullying perpetration in adolescence: A systematic review. *Aggression and Violent Behavior, 23*, 126–136.
Balakrishnan, V., Khan, S., & Arabnia, H. R. (2020). Improving cyberbullying detection using Twitter users' psychological features and machine learning. *Computers & Security, 90*, 101710.
Balakrishnan, V., Khan, S., Fernandez, T., et al. (2019). Cyberbullying detection on twitter using Big Five and Dark Triad features. *Personality and Individual Differences, 141*, 252–257.
Balakrishnan, V., & Kaity, M. (2023). Cyberbullying detection and machine learning: A systematic literature review. *Artificial Intelligence Review, 56*, 1–42.
Bashir Shaikh, F., Rehman, M., & Amin, A. (2020). Cyberbullying: A systematic literature review to identify the factors impelling university students towards cyberbullying. *IEEE Access, 8*, 148031–148051.
Bu, S.-J., & Cho, S.-B. (2018). A hybrid deep learning system of CNN and LRCN to detect cyberbullying from SNS comments. In F. J. de Cos Juez, J. R. Villar, E. A. de la Cal, et al. (Eds.), *Hybrid artificial intelligent systems* (pp. 561–572). Springer International Publishing.
Catone, G., Marwaha, S., Lennox, B., et al. (2017). Bullying victimization and psychosis: The interdependence and independence of risk trajectories. *BJPsych Advances, 23*, 397–406.
Cioffi, R., Travaglioni, M., Piscitelli, G., et al. (2020). Artificial intelligence and machine learning applications in smart production: Progress, trends, and directions. *Sustainability, 12*, 492.
Dhingra, N., Chawla, S., Saini, O., et al. (2023). An improved detection of cyberbullying on social media using randomized sampling. *International Journal of Bullying Prevention*. https://doi.org/10.1007/s42380-023-00188-4
Elor, Yotam, and Hadar Averbuch-Elor. ″To SMOTE, or Not to SMOTE?″ arXiv, 2022. https://doi.org/10.48550/ARXIV.2201.08528.
Farzand, M. (2023). *Evolution of internet slang and its impact on English language communication*.
Foody, M., McGuire, L., Kuldas, S., et al. (2019). Friendship quality and gender differences in association with cyberbullying involvement and psychological well-being. *Frontiers in Psychology, 10*.
Fu, Q., Gao, Z., Zhou, J., et al. (2021). CLSA: A novel deep learning model for MOOC dropout prediction. *Computers & Electrical Engineering, 94*. https://doi.org/10.1016/j.compeleceng.2021.107315
Geel, M., Goemans, A., Toprak, F., et al. (2016). Which personality traits are related to traditional bullying and cyberbullying? A study with the Big Five, Dark Triad and sadism. *Personality and Individual Differences, 106*. https://doi.org/10.1016/j.paid.2016.10.063
Gredler, G. R. (2003). Review of *Bullying at school: What we know and what we can do*. *Psychology in the Schools, 40*, 699–700.
Hani, J., Nashaat, M., Ahmed, M., et al. (2019). Social media cyberbullying detection using machine learning. *International Journal of Advanced Computer Science and Applications (IJACSA), 10*. https://doi.org/10.14569/IJACSA.2019.0100587
Hinduja, S., & Patchin, J. W. (2010). Bullying, cyberbullying, and suicide. *Archives of Suicide Research, 14*, 206–221.
Howland, J., Wright, T., Boughan, R., et al. (2009). How scholarly is Google Scholar? A comparison to library databases. *College & Research Libraries, 70*, 227–234.

Hudda, N., Mishra, A., & Kumar, S. (2023). Detection of cyberbullying on social media. *International Journal of Computer Applications, 185*, 43–47.
Huertas-García, Á., Martín, A., Huertas-Tato, J., et al. (2023). Countering malicious content moderation evasion in online social networks: Simulation and detection of word camouflage. *Applied Soft Computing, 145*, 110552.
Islam, M. M., Uddin, M. A., Islam, L., et al. (2020). Cyberbullying detection on social networks using machine learning approaches. *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, 1–6.
Jacovi, A., Sar Shalom, O., & Goldberg, Y. (2018). Understanding convolutional neural networks for text classification. In T. Linzen, G. Chrupała, & A. Alishahi (Eds.), *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP* (pp. 56–65). Association for Computational Linguistics.
Joachims, T. (1998). Text categorization with support vector machines. *Proc European Conf Machine Learning (ECML'98)*. https://doi.org/10.17877/DE290R-5097
Joyce, B., & Deng, J. (2019). Sentiment analysis using naïve Bayes approach with weighted reviews - A case study. *2019 IEEE Global Communications Conference (GLOBECOM)*, 1–6.
Kadam, B. (2023). Cyberbullying detection using machine learning algorithms. *IJRASET, 11*, 1326–1328.
Kowalski, R. M., Toth, A., & Morgan, M. (2018). Bullying and cyberbullying in adulthood and the workplace. *Journal of Social Psychology, 158*, 64–81.
Li, Y., Li, D., Li, X., et al. (2018). Cyber victimization and adolescent depression: The mediating role of psychological insecurity and the moderating role of perceived social support. *Children and Youth Services Review, 94*. https://doi.org/10.1016/j.childyouth.2018.09.027
Marín Balcázar, S., & Guachichullca Guamán, M. P. (2022). A systematic literature review on bullying and cyberbullying models, programs, and intervention strategies. *Maskana, 13*, 21–33.
Mateu, A., Pascual-Sánchez, A., Martinez-Herves, M., et al. (2020). Cyberbullying and post-traumatic stress symptoms in UK adolescents. *Archives of Disease in Childhood, 105*, 951–956.
Mehendale, N., Rajpara, K., Shah, K., et al. (2022). A review on cyberbullying detection using machine learning. https://doi.org/10.2139/ssrn.4116153
Mitsopoulou, E., & Giovazolias, T. (2015). Personality traits, empathy and bullying behavior: A meta-analytic approach. *Aggression and Violent Behavior, 21*, 61–72.
Muneer, A., & Fati, S. (2020). A comparative analysis of machine learning techniques for cyberbullying detection on Twitter. *Future Internet, 12*, 187.
Noviantho, I. S. M., & Ashianti, L. (2017). Cyberbullying classification using text mining. *2017 1st International Conference on Informatics and Computational Sciences (ICICoS)*, 241–246.
Obamiyi, S. E., Badeji-Ajisafe, B., Oguntimilehin, A., et al. (2023). An ensemble approach to cyberbullying detection and prevention on social media. *ABUAD International Journal of Natural and Applied Sciences, 3*, 47–52.
Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing, 26*, 217–222.
Peebles, E. (2014). Cyberbullying: Hiding behind the screen. *Paediatrics & Child Health, 19*, 527–528.

Peng, J., Lee, K., & Ingersoll, G. (2002). An introduction to logistic regression analysis and reporting. *Journal of Educational Research, 96*, 3–14.
Rosa, H., Matos, D., Ribeiro, R., et al. (2018). A “deeper” look at detecting cyberbullying in social networks. *2018 International Joint Conference on Neural Networks (IJCNN)*, 1–8.
Slonje, R., & Smith, P. K. (2008). Cyberbullying: Another main type of bullying? *Scandinavian Journal of Psychology, 49*, 147–154.
Smith, P. K., Mahdavi, J., Carvalho, M., et al. (2008). Cyberbullying: Its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry, 49*, 376–385.
Srivastava, D., & Bhambhu, L. (2010). Data classification using support vector machine. *Journal of Theoretical and Applied Information Technology, 12*, 1–7.
Strzelecka, A., Kurdyś-Kujawska, A., & Zawadzka, D. (2020). Application of logistic regression models to assess household financial decisions regarding debt. *Procedia Computer Science, 176*, 3418–3427.
Swearer, S. M., & Hymel, S. (2015). Understanding the psychology of bullying: Moving toward a social-ecological diathesis–stress model. *American Psychologist, 70*, 344–353.
Van Bruwaene, D., Huang, Q., & Inkpen, D. (2020). A multi-platform dataset for detecting cyberbullying in social media. *Language Resources and Evaluation, 54*, 851–874.
Van Huynh, T., Nguyen, V. D., Van Nguyen, K., et al. (2019). Hate speech detection on Vietnamese social media text using the Bi-GRU-LSTM-CNN model. https://doi.org/10.48550/arXiv.1911.03644.
Wright, M. F. (2018). Cyberstalking victimization, depression, and academic performance: The role of perceived social support from parents. *Cyberpsychology, Behavior, and Social Networking, 21*, 110–116.
Yadav, J., Kumar, D., & Chauhan, D. (2020). Cyberbullying detection using pre-trained BERT model. *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 1096–1100.
Yuan, J., Wang, H., Lin, C., et al. (2019). A novel GRU-RNN network model for dynamic path planning of mobile robot. *IEEE Access, 7*, 15140–15151.
Zhao, R., Zhou, A., & Mao, K. (2016). Automatic detection of cyberbullying on social networks based on bullying features. *Proceedings of the 17th International Conference on Distributed Computing and Networking*, 1–6.

7. EXPLAINING PERCEIVED QUALITY OF A TERTIARY EDUCATION RECOMMENDATION SYSTEM THROUGH USER STUDIES

Unathi September
University of Cape Town
sptuna001@myuct.ac.za

Lisa Seymour
University of Cape Town
lisa.seymour@uct.ac.za

Abstract

This study investigates the factors influencing perceived quality in tertiary education Recommender Systems (RS) by examining structural relationships among user-centric evaluation constructs. The research developed an evaluation framework integrating objective system data with subjective user feedback to ensure that recommendations aid in selecting educational pathways. Using a Design Science Research approach, the framework underwent three DSR iterations with sample sizes of N = 166, 357, and 719, respectively. Quantitative results from the final iteration revealed several statistically robust significant relationships amongst the evaluated constructs: Perceived Recommendation Set Variety (PRSV) positively influenced Perceived System Usefulness (PSU) ($\beta = 0.36$, $t = 7.92$, $p < 0.001$) and Perceived Recommendation Quality (PRQ) also significantly impacted PSU ($\beta = 0.36$, $t = 8.86$, $p < 0.001$). Furthermore, Perceived System Enjoyment (PSE) and Perceived Ease of Use (PEU) contributed to Overall Satisfaction (OS) ($\beta = 0.16$, $t = 3.03$, $p < 0.01$; $\beta = 0.27$, $t = 5.28$, $p < 0.001$, respectively), which in turn strongly drove user Interaction (INT) ($\beta = 0.59$, $t = 19.02$, $p < 0.001$). Additional significant relationships included PSU's positive effect on Overall Trust in the RS (OTRS) ($\beta = 0.22$, $t = 5.18$, $p < 0.001$), as well as modest yet significant effects of PRSV and PRQ on OTRS. These findings, derived from exploratory and confirmatory analyses, validate the framework's effectiveness in capturing behavioural metrics and user perceptions, offering RS developers a robust tool to enhance RS quality in tertiary education.

Keywords

Recommender Systems, Design Science Research, Information Systems

1. Introduction

Recommender Systems (RS) have emerged as a solution for filtering data and facilitating efficient decision-making (O'Donovan & Smyth, 2005). The phenomena of information overload and the need to make informed, data-driven decisions are prevalent in the education field, where high school learners are required to make decisions about institutions and qualifications after completing their high school education. One of the RS's objectives is to proactively suggest items based on observed user behaviours or interests explicitly stated in a system (Pu et al., 2011). RS outputs help users easily filter and find information in large databases with numerous options (Adomavicius et al., 2019).

Researchers increasingly agree that RS's aim goes beyond prediction accuracy; its primary real-world purpose is to offer personalised assistance in finding relevant content or items. (Ricci et al.,

2015). Thus, the RS needs to be evaluated in multiple dimensions, including user-centric perspectives such as trust, novelty, serendipity, utility, usability, and risk. In this research context, perceived quality refers to the users' subjective assessment of the RS value and effectiveness. This perception is understood to be shaped by the user's experiences with the RS, such as its usefulness, the quality of its recommendations, ease of use, and the enjoyment derived from interacting with it. This is distinct from the system quality that may be measured through metrics like MRR, MAE, nDCG or even RMSE. While this study incorporates objective system data, its primary focus is capturing and modelling these user-centric perceptions of quality. Considering the importance of quality in education RS, this study aims to address the question: What are the structural relationships among user-centric constructs, and how are they associated with perceived quality in a tertiary education recommender system?

2. Related work

2.1. Recommender Systems Evaluation Experiments

The RS field defines three foundational methods in RS experiments: User Studies, Online and Offline evaluations. User studies rely on users' explicit selection or satisfaction measures, which are traditionally obtained through rating procedures where test users are asked to complete a series of activities that require interaction with the RS (Ricci et al., 2011). Researchers record users' behaviours during tasks to gather quantitative data. Lab studies involve participants being aware of their participation in simulations, while real-world studies have participants unaware of the research, rating recommendations as they use the RS. (Beel & Langer, 2015). In other forms of User studies, researchers may conduct qualitative interviews before, during, and after completing a task. Such inquiries can elicit information that is not immediately apparent, such as whether the user appreciated the user interface or found the activity simple to accomplish (Ricci et al., 2015). Unlike user studies, Online evaluation measures aspects such as users' acceptance rates of recommendations in real-world RS. In contrast, offline evaluations are conducted utilising a pre-compiled data collection of user selections or ratings. Offline studies use metrics such as MRR, Recall, F-measure, MAE, nDCG, and RMSE to measure recommendation accuracy (Beel & Langer, 2015).

2.2. Existing User Studies Frameworks for RS Evaluation

Literature on user study evaluation frameworks reveals two widely accepted evaluation frameworks in RS evaluation: ResQue (Pu et al., 2012) and The Framework for User-centric Evaluation of Recommender Systems (B. Knijnenburg, Willemsen, & Kobsa, 2011). The Framework for User-Centric Evaluation of Recommender Systems is designed to facilitate direct control over user experience evaluations. Grounded in its robust theoretical foundations, the Technology Acceptance Model (TAM), the framework adopts a comprehensive approach to assessing the relationships among the RS's Objective System Aspects (OSA), Subjective System Aspects (SSA), User Experience (EXP), and Interaction (INT). This framework was found to be a more suitable framework to build on for this study, as demonstrated in Figure 1.

The framework emphasises user-centric constructs for measurement of Perceived Recommendations Set Variety (PRSV), Perceived Recommendation Quality (PRQ), Trust, Usage effort, Choice difficulty, Perceived System Effectiveness (PSE), Satisfaction, System-specific Privacy Concerns (SPC), and Intent to provide feedback. It considers these aspects to mediate the link between objective system features and user experience, while also acknowledging the user's

personal and situational traits, including Demographics, Domain knowledge, Trust, System Privacy Concerns, and usage frequency. These factors are essential for moderating the RS user experience. The framework connects objective elements and subjective user behaviour, mediated by selected constructs that clarify how and why the user interacts with the RS (Knijnenburg, Willemsen, & Kobsa, 2011). Through the related studies, the identified gap in the literature was the lack of user studies that examine and model the structural relationships in the evaluation frameworks for measuring perceived quality in education RS. Consequently, the study aims to answer an explanatory question: What are the structural relationships among user-centric constructs, and how are they associated with perceived quality in a tertiary education recommender system?

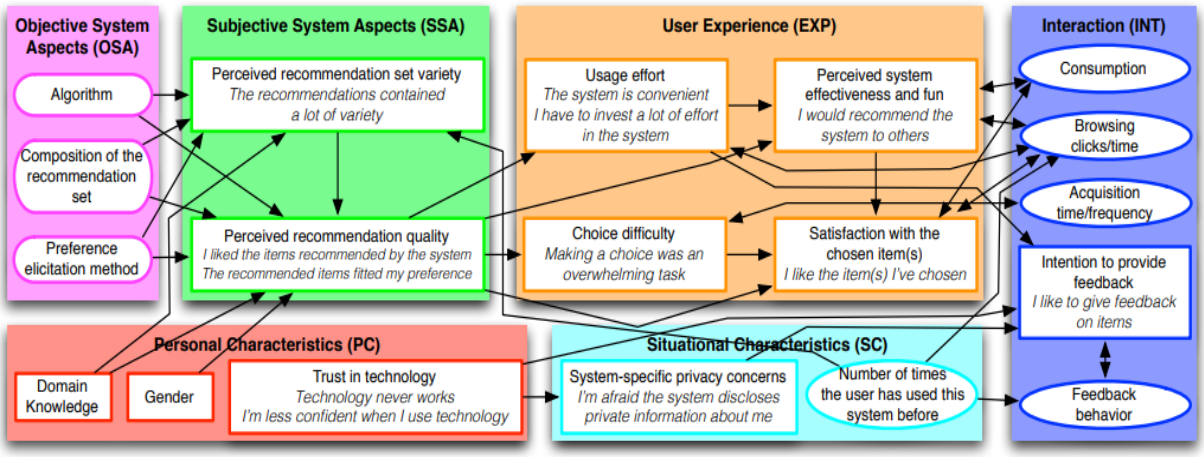

Fig. 1. The framework for user-centric evaluation of recommender systems (B. P. Knijnenburg et al., 2011)

3. Method

This study addresses an explanatory research question aimed at elucidating the tested and validated associations between user-centric constructs within the proposed evaluation framework for tertiary education recommender systems. This study is situated within a pragmatic philosophical paradigm, allowing the researcher to use both deductive and abductive reasoning throughout the iterative design cycles of the study, viewed through research onion lenses (Saunders et al., 2019). The study employs Structural Equation Modelling (SEM) as the primary analytical technique to examine how specific user-centric constructs are associated with users' perceptions of education RS quality. The SEM was performed using the seminr package in R. The explanatory focus provides empirical validation of the proposed framework, enhancing its generalizability and utility for future applications in both research and practice. DSR was selected as the most appropriate research strategy for building artefacts from existing RS theories and evaluation models. The DSR allowed for iterative artefact development, testing, and refinement across three research iterations. Figure 2 illustrates the adaptation of DSR reasoning process, highlighting the knowledge flows, process steps, outputs, and related cognitive processes employed to develop the final framework (Vaishnavi et al., 2019). The framework incorporates user-centric constructs that explain the perceived quality of an RS from a user's perspective.

This study employed mixed-methods data, utilising qualitative and quantitative research approaches to address the research problem from a user's perspective. Mixed techniques enable

more detailed and nuanced data insights by integrating quantitative and qualitative patterns (Vaishnavi et al., 2019). This comprehensive approach reduces the biases and limitations of single-method studies, thereby improving the trustworthiness of research (Venkatesh et al., 2013). The research utilised questionnaires and semi-structured interviews to explore user input and validate quantitative results. Twenty-six volunteer university students who completed the first iteration of questionnaires were interviewed across South African universities. These interviews enhanced the research by providing flexible, in-depth data and capturing nuanced responses. Students offered insights about other information sources for exploring institutions and qualifications. Additionally, the interviews validated findings and highlighted discrepancies from the first two iterations. The results improved the assessment model's resilience and reliability, leading to a focused third iteration on significant findings instead of the general hypothesis. Thematic analysis of the interview transcripts followed the six-stage process; involving familiarisation with the data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report, to identify patterns relevant to user perceptions (Braun & Clarke, 2006). NVivo was used to manage this process. Moreover, an objective system dataset, in the form of the user's audit trail, was used to understand how frequently students performed tasks, leading to the generation of final recommendations and the management of their profiles on the education RS. The introduction of objective user data for each participating user, collected through system logs, was mapped to each user by the system's User_Id, along with the questionnaire answers from the students, to form a comprehensive view of the user's RS usage and the feedback provided through the forms.

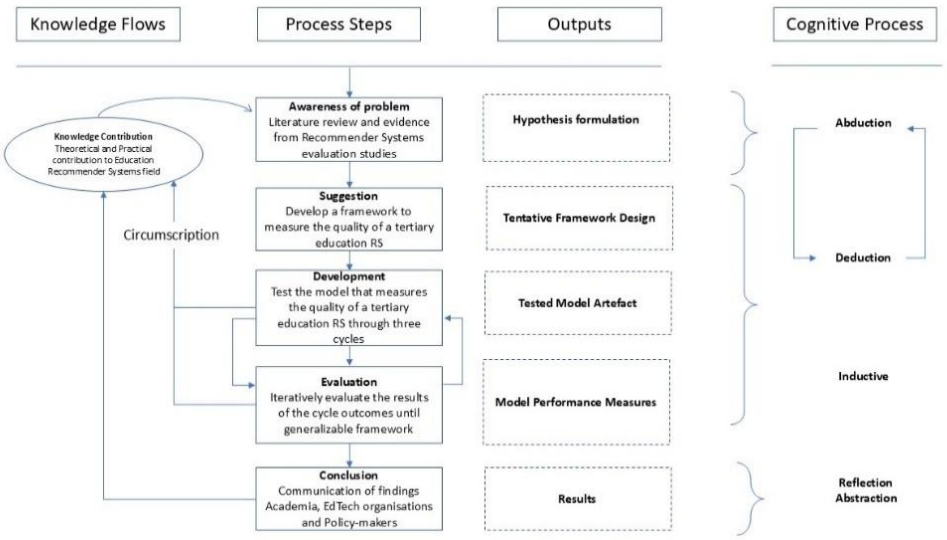

Fig. 2. Adaptation of DSR reasoning and process steps for the research

This study gathered data from tertiary students over three years, assessing their RS experiences at the end of each academic year, after completing six to ten months in their chosen institutions. SEM, an integrative statistical procedure, was consistently applied to test relationships throughout the three iterations simultaneously. SEM lets researchers model and estimate complex interactions between several dependent and independent variables (Cole & Preacher, 2014) . SEM compensates for observed variable measurement error in relationship estimates. Thus, the technique measures

theoretical notions more precisely (Cole & Preacher, 2014). Before SEM analysis, Confirmatory Factor Analysis (CFA) with the lavaan package in R validated the measurement model, ensuring observed variables reliably represented their latent constructs. SEM using the seminr package, then tested the hypothesised structural relationships between these constructs. Model fit was assessed using common indices (CFI, TLI, RMSEA, DF, Chi-Square). Table 1 shows the study's hypothesised relationships. Recognising that individual differences influence technology perception, based on The Framework for User-Centric Evaluation of Recommender Systems, we suggested that personal characteristics like Age, School Quintile, and Domain Knowledge might moderate key relationships (H10a, H10b, H10c), but these effects were not supported in the final data.

4. Results

This section presents the results of three iterations and highlights users' steps to obtain recommendations for tertiary and qualification options, focusing on the third iteration of the user experiment.

4.1. Results of Recommendation Composition Set Change

Manipulating the Recommendation Composite Set variables is crucial for evaluating OSA relationships with SSA, EXP, and INT (B. P. Knijnenburg et al., 2011). The study manipulated variables related to the Number of Recommendations (NRec) given to students in two iterations. In the first, students received seven institution recommendations, each with three qualifications, totaling twenty-one qualifications. The second iteration presented eleven institutions, each with three qualifications, resulting in a total of thirty-three recommended qualifications. This increase aimed to improve RS, providing a wider variety of options and clearer distinctions between institutions and qualifications. This led to Hypothesis 1: the composition of the recommendation set will positively impact Perceived Recommendation Quality (PRQ).

To test this hypothesis, the study compared the mean PRQ scores of the 21-item group (Iteration 1) and the 33-item group (Iteration 2) using the Welch two-sample t-test and the Wilcoxon Rank-Sum Test, analysing the distribution of PRQ scores in the two iterations. The Welch t-test yielded a non-significant p-value ($t = -0.6024$, $df = 345.05$, $p = 0.5473$, with a 95% confidence interval of [-0.2888450, 0.1533975]), indicating no statistically significant evidence that the 33-item group had a higher or lower PRQ mean compared to the 21-item group. Furthermore, the sample estimates showed mean PRQ scores of 4.02259 for Iteration 1 and 4.09031 for Iteration 2. The Wilcoxon rank-sum test found no significant difference in the median PRQ between the two groups. The data showed that increasing the recommendation set size did not increase students' PRQ. Quantitatively, the mean PRQ difference was negligible and non-significant.

4.2. Iteration 3 Structural Equation Model Results

The RS was evaluated using objective system interaction data and user perspective, and the paths presented in Table 3 emerged. Including only the hypothesised effects in reporting results may lead to overlooking other significant effects (B. P. Knijnenburg et al., 2012b). Unlike the first two iterations, the third iteration of the SEM model results includes only significant relationships to trim the model and focus on the key relationships discussed in the subsequent discussion section. Hypotheses 9b, 9c, 10a, 10b, and 10c were eliminated as there was no significant evidence from

the final iteration data to support their continuation. A summary of the significant relationships for iteration three is presented in Table 2, which includes standardised path coefficients (β) representing the strength and direction of the relationships, along with their statistical significance from the final iteration (Iteration 3).

Hypothesis	Hypothesis detail	Reference
H1	The composition of the recommendation set will positively affect the Perceived Recommendation Quality (PRQ).	(B. Knijnenburg, Willemsen, & Kobsa, 2011; Pu et al., 2011; Ricci et al., 2015)
H2	The Perceived Recommendation Set Variety (PRSV) will positively affect the RS's Perceived System Usefulness (PSU).	(Adomavicius et al., 2019; B. Knijnenburg, Willemsen, Gantner, et al., 2011; McNee et al., 2006; Pu et al., 2011)
H3	The Perceived Recommendation Quality (PRQ) will positively affect the Perceived System Usefulness (PSU) of the RS.	(Wang & Benbasat, 2007; Xu & Cenfetelli, 2011)
H4	The Perceived System Enjoyment (PSE) of the system will have a positive effect on the Overall Satisfaction (OS) of the system	(Lee et al., 2003; Thong et al., 2006; Van Der Heijden, 2004)
H5	The RS Perceived Ease of Use (PEU) will have a positive effect on the Overall Satisfaction (OS) of the system.	(Venkatesh & Davis, 2000; Xu & Cenfetelli, 2013)
H6	The Perceived System Usefulness (PSU) will positively affect the Overall Trust in the RS (OTRS)	(Komiak & Benbasat, 2006; Koufaris & Hampton-Sosa, 2004; Wang & Benbasat, 2007)
H7	The Overall Trust in the RS (OTRS) will positively affect the RS Interaction (INT)	(Benlian et al., 2012; Komiak & Benbasat, 2006; Söllner et al., 2016)
H8	The Overall Satisfaction (OS) will have a positive effect on Interaction (INT)	(B. P. Knijnenburg et al., 2012a; Pu et al., 2011)
H9a	The User's Situational Characteristics (SC); System Privacy Concerns (SPC) will negatively affect the Perceived Recommendation Quality (PRQ)	(Awad & Krishnan, 2006, 2006; Beldad et al., 2010; B. Knijnenburg & Kobsa, 2013)
H9b	The User's Situational Characteristics (SC); System Prior-Knowledge (SPK) will positively affect the Overall Satisfaction (OS)	(B. P. Knijnenburg et al., 2012a; Pu et al., 2012; Wang & Benbasat, 2007)
H9c	The User's Situational Characteristics (SC); Choice Goal (CG), will positively affect the Overall Trust in the RS (OTRS)	(Benbasat, 2005; B. P. Knijnenburg et al., 2012a; Pu & Chen, 2007; Söllner et al., 2014)

H10a	The effect of OTRS on the RS utility will be moderated by the user's Personal Characteristics (PC), such as Age.	(B. Knijnenburg & Kobsa, 2013; McKnight et al., 2002; Venkatesh et al., 2003)
H10b	The effect of OTRS on the RS utility will be moderated by the user's Personal Characteristics (PC) like School Quintile.	(Van Deursen & Van Dijk, 2014)
H10c	The effect of Overall satisfaction on the RS utility will be moderated by the user's Personal Characteristics (PC), such as Domain Knowledge.	(B. P. Knijnenburg et al., 2012a; Pu et al., 2012)

Table 1. Hypothesis table

Relationship	Iteration Original	3 Bootstrap Mean	Bootstrap SD	T Stat.	2.5% CI	97.5% CI	Significance
PRSV → PSU	0.3645	0.3656	0.0461	7.9158	0.2719	0.4587	***
PRQ → PSU	0.3558	0.3548	0.0402	8.8566	0.2828	0.4354	***
PSE → OS [H4]	0.1562	0.1590	0.0516	3.0269	0.0652	0.2586	**
PEU → OS [H5]	0.2669	0.2657	0.0506	5.2770	0.1679	0.3614	***
PSU → OTRS	0.2202	0.2231	0.0425	5.1817	0.1421	0.3107	***
OTRS → INT	0.0956	0.0976	0.0302	3.1657	0.0396	0.1586	**
OS → INT [H8]	0.5865	0.5854	0.0308	19.0241	0.5237	0.6434	***
SPC → PRQ	0.3484	0.3480	0.0453	7.6959	0.2610	0.4368	***
SPC → PSU	0.1240	0.1237	0.0231	5.3642	0.0843	0.1739	***
PEU → INT	0.1565	0.1560	0.0328	4.7721	0.0925	0.2225	***
PRSV → OTRS	0.0803	0.0814	0.0179	4.4920	0.0489	0.1178	***
PSU → INT	0.1414	0.1406	0.0319	4.4338	0.0777	0.1990	***
PRQ → OTRS	0.0783	0.0794	0.0189	4.1459	0.0463	0.1199	***
PRSV → INT	0.0516	0.0513	0.0131	3.9334	0.0276	0.0782	***
PSU → OS	0.2052	0.2028	0.0528	3.8838	0.0988	0.3036	***
PRQ → INT	0.0503	0.0501	0.0135	3.7229	0.0263	0.0789	***
PRSV → OS	0.0748	0.0741	0.0214	3.4978	0.0342	0.1182	***
PRQ → OS	0.0730	0.0722	0.0215	3.3994	0.0315	0.1180	***
PSE → INT	0.0916	0.0932	0.0310	2.9526	0.0372	0.1524	**
SPC → OTRS	0.1124	0.1127	0.0407	2.7596	0.0362	0.1918	**
SPC → INT	0.0488	0.0494	0.0210	2.3184	0.0087	0.0915	*

Table 2. Path Coefficients and Significance Levels of SEM in Iteration 3

5. Discussion

The third iteration of the DSR highlighted the validation of key hypotheses, further solidifying the roles of EXP, Trust, and the quality of recommendations in students' decision-making processes when presented with recommendations. A significant finding is the continued confirmation of H2, which states that PRSV positively affects PSU. The assessed relationship remained highly significant ($\beta = 0.3645$, $t = 7.92$, $p < 0.001$) across iterations despite a slight decrease from Iteration

2 ($\beta = 0.4533$). This confirms that students value diverse recommendations when selecting tertiary institutions and qualifications. One student expressed, “I actually discovered some institutions I never knew about. I discovered careers I never knew about. It enlightened me to new careers.” Another reinforced this, stating, “I was able to see multiple qualifications in the same career field, and it allowed me to compare institutions more clearly than before.” These findings suggest that PRSV has a significant impact on PSU. Similarly, H3, which asserts that PRQ positively affects PSU, was strongly supported ($\beta = 0.3558$, $t = 8.86$, $p < 0.001$), even though the effect size slightly decreased from Iteration 2 ($\beta = 0.4441$). Students described how the RS provided high-quality, personalised recommendations aligned with their aspirations. One student explained, "The app showed me the specific qualifications relevant to me. It didn’t just throw random suggestions; they actually made sense." This highlights the system’s effectiveness in providing targeted, meaningful recommendations that enhance decision-making. H4, which examines the impact of PSE on OS, was also validated. The relationship became statistically significant ($\beta = 0.1562$, $t = 3.03$, $p < 0.01$), increasing from Iteration 2 ($\beta = 0.0788$). Students emphasised how an engaging, interactive interface contributed to their satisfaction. One student noted, “I actually enjoyed using the app. It felt like an easy process, not something stressful like some of the university websites.”Another added, “It was simple but interactive. It felt like it was designed for students like me.” These findings indicate that a system prioritising ease of use and enjoyment can significantly enhance OS.

H5, which states that PEU positively affects OS, was strongly supported ($\beta = 0.2669$, $t = 5.28$, $p < 0.001$). One student remarked, “The app was very easy to use, no struggle at all.” While another concurred by saying, “The instructions were clear, and I didn’t feel overwhelmed at any point.” H6, which asserts that PSU positively affects OTRS, remained highly significant ($\beta = 0.2202$, $t = 5.18$, $p < 0.001$), despite a slight decline from Iteration 2 ($\beta = 0.2670$). This suggests that students who found the system useful were more likely to trust its recommendations. One student noted, “I felt like I could trust it because it gave me realistic options based on my marks, not just random courses.” Concurring, another student added, “It actually helped me narrow down my choices, which means it wasn’t just giving out random suggestions.” These findings highlight the importance of PSU in building trust. H7, which suggests that OTRS positively affects INT, was statistically significant ($\beta = 0.0956$, $t = 3.17$, $p < 0.01$), confirming that students who trusted the system engaged with it more. One participant stated, “Because I trusted it, I actually followed through with the recommendations.” These findings highlight the strong link between trust and sustained INT. H8, which examines the relationship between OS and INT, was further reinforced in Iteration 3 ($\beta = 0.5865$, $t = 19.02$, $p < 0.001$), showing a substantial increase from Iteration 2 ($\beta = 0.5257$, $t = 7.93$). One student enthusiastically noted, “I actually used the app multiple times because it was just that good.” This finding strongly validates the role of OS in driving INT.

Beyond the validated hypotheses, several newly significant relationships emerged in Iteration 3. The emergence of these significant relationships in Iteration 3, such as PEU → INT and PRSV → OTRS, may be attributed to several factors. The substantially larger sample size in Iteration 3 (N=719) provided greater statistical power to detect these effects. Additionally, iterative refinements to the RS based on feedback from prior iterations could have led to clearer user perceptions of these specific system aspects, making their influence more pronounced. PEU and INT became highly significant ($\beta = 0.1565$, $t = 4.77$, $p < 0.001$) over the Iterations, suggesting that the ease of use enhances OS and directly drives INT. Additionally, PRSV and OTRS emerged as

significant ($\beta = 0.0803$, $t = 4.49$, $p < 0.001$), indicating that students who perceive a greater variety in recommendations are more likely to develop trust in the system. PRQ and OTRS relationship also emerged ($\beta = 0.0783$, $t = 4.15$, $p < 0.001$), highly significant, reinforcing that students trust the system more when they perceive the recommendations as accurate and relevant. In Iteration 3, PSE and INT became statistically significant ($\beta = 0.0916$, $t = 2.95$, $p < 0.01$), suggesting that an enjoyable system increases OS and motivates students to engage more actively. Several hypotheses related to situational and personal characteristics (H9a, H9b, H9c, H10b, H10c) were not supported in the final third iteration, and these relationships remained either insignificant or weak through the iterations. For example, H9a, which examined whether SPC influenced PRQ, was not strongly validated. Many students reported having little to no concerns regarding privacy, with one participant stating, "I didn't see any need for me to be worried.". Similarly, Hypothesis 10a, which examined the moderating effect of Age (AGE) on Interaction (INT), demonstrated weaker empirical support, indicating that age-related differences did not significantly affect engagement with the system.

Although H1 results did not reach statistical significance, it is essential to acknowledge that real-life experiments pose unique challenges compared to laboratory studies. In live settings, altering the size of the recommendation set directly impacts actual users, which restricts the degree to which variables can be modified without affecting real-world outcomes. Consequently, detecting threshold effects in such environments may necessitate prolonged A/B testing or longitudinal studies across additional academic cycles to capture significant changes. Figure 3 illustrates the final SEM model generated from the third iteration of the DSR.

6. Conclusion and Future Work

This study advances the evaluation of tertiary education RS through testing the structural relationships among user-centric evaluation constructs that explain perceived quality in tertiary education recommender systems. The study contributes to practice within Information Systems and the RS fields in education. By integrating objective system data with subjective user feedback through an iterative DSR approach and employing SEM alongside qualitative analyses, we have explored key relationships among constructs such as PRQ, PSU, Trust, and OS. These findings directly inform the design and improvement of educational RS, demonstrating how tailored recommendation strategies can enhance user engagement and decision-making. The framework, with its relationships, clarifies the interplay between critical constructs like PRSV, PSU, and Trust in the RS. It introduces constructs such as Interaction that expand our understanding of RS evaluation in educational contexts, which, in practice, further contributes to a validated, actionable model that education technology developers can use to assess and improve the quality of educational RS. The insights from this study lay the groundwork for developing a comprehensive software prototype that can evaluate RS performance, ensuring that recommendation technologies are measured beyond accuracy and consider user sentiments as well.

This study demonstrated integrating an evaluation of user-centric aspects into a real-life RS, capturing detailed and objective behavioural data from the RS. Further, this research demonstrated the integration of RS evaluation within larger education systems and addressed broader sociotechnical and ethical challenges around RS for decisions with long-term impact. Future studies can explore dynamic feedback loops, integrating student advisor insights for adaptive systems. Further RS development can include co-curricular and extracurricular data for holistic

recommendations, enhancing quality and relevance. RS design can foster student reflection and discovery, challenging assumptions and introducing new opportunities beyond current preferences. Based on the model, future studies can focus on comprehensive online evaluation platforms, ensuring cultural and situational invariance in measurement, as explored in initial study iterations

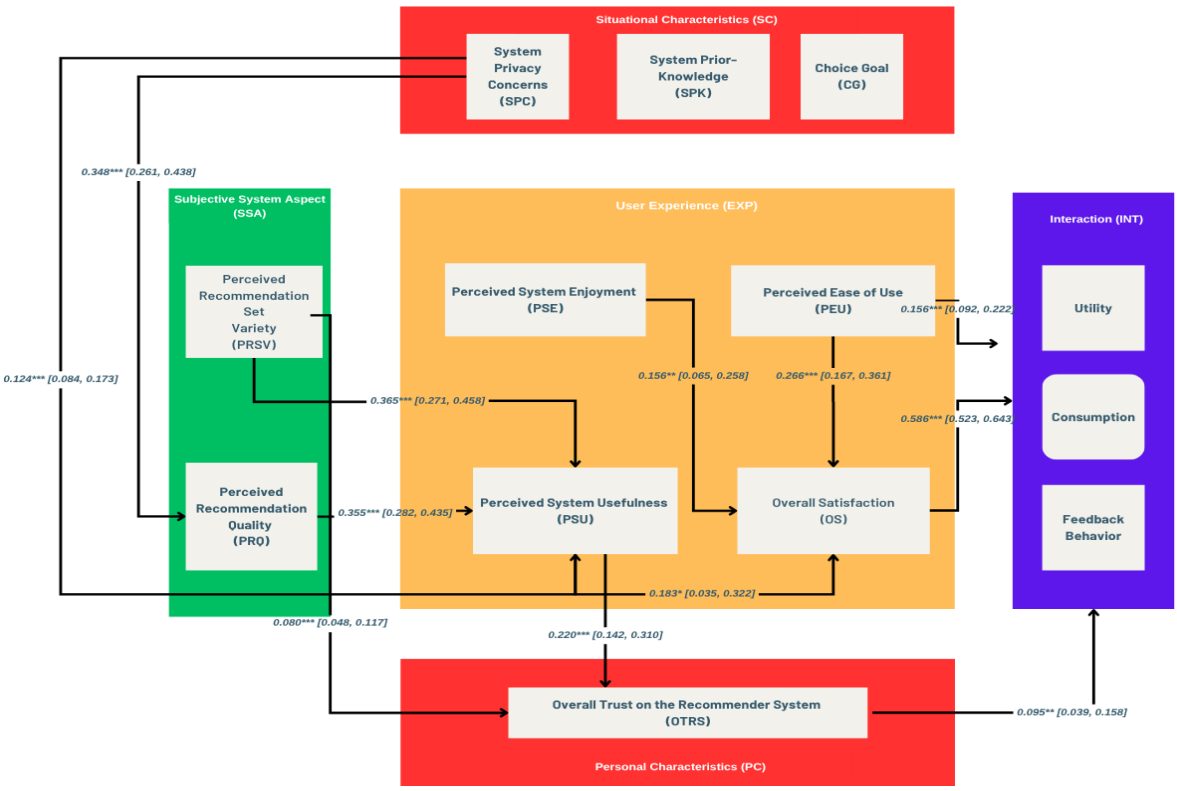

Fig. 3 SEM path model for Iteration 3 tested relationships Path Coefficients and Significance Levels of SEM in Iteration 3.
*** $p < .001$, ** $p < .01$, * $p < .05$, Insignificant = $p \geq .0$

References

Adomavicius, G., Bockstedt, J., Curley, S., & Zang, J. (2019). Reducing Recommender System Biases: An Investigation of Rating Display Designs. MIS Quarterly, 43, 1321–1341. https://doi.org/10.25300/misq/2019/13949

Awad & Krishnan. (2006). The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization. MIS Quarterly, 30(1), 13. https://doi.org/10.2307/25148715

Beel, J., & Langer, S. (2015). A Comparison of Offline Evaluations, Online Evaluations, and User Studies in the Context of Research-Paper Recommender Systems. International Conference on Theory and Practice of Digital Libraries, 153–168.

Beldad, A., De Jong, M., & Steehouder, M. (2010). How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust. Computers in Human Behavior, 26(5), 857–869. https://doi.org/10.1016/j.chb.2010.03.013

Benbasat, I. (2005). Trust In and Adoption of Online Recommendation Agents. Journal of the Association for Information Systems, 6(3), 72–101. https://doi.org/10.17705/1jais.00065

Benlian, A., Titah, R., & Hess, T. (2012). Differential Effects of Provider Recommendations and Consumer Reviews in E-Commerce Transactions: An Experimental Study. Journal of Management Information Systems, 29(1), 237–272. https://doi.org/10.2753/MIS0742-1222290107

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. Psychological Methods, 19(2), 300–315. https://doi.org/10.1037/a0033805

Knijnenburg, B., & Kobsa, A. (2013). Making decisions about privacy: Information disclosure in contextaware recommender systems. ACM Transactions on Interactive Intelligent Systems, 3, 1–23.

Knijnenburg, B. P., Willemsen, M. C., Gantner, Z., Soncu, H., & Newell, C. (2012a). Explaining the user experience of recommender systems. User Modeling and User-Adapted Interaction, 22(4–5), 441–504. https://doi.org/10.1007/s11257-011-9118-4

Knijnenburg, B. P., Willemsen, M. C., Gantner, Z., Soncu, H., & Newell, C. (2012b). Explaining the user experience of recommender systems. User Modeling and User-Adapted Interaction, 22(4–5), 441–504. https://doi.org/10.1007/s11257-011-9118-4

Knijnenburg, B. P., Willemsen, M. C., & Kobsa, A. (2011). A pragmatic procedure to support the user-centric evaluation of recommender systems. Proceedings of the Fifth ACM Conference on Recommender Systems, 321–324. https://doi.org/10.1145/2043932.2043993

Knijnenburg, B., Willemsen, M., Gantner, Z., & Soncu, H. (2011). Explaining the user experience of recommender systems. User Modeling and User-Adapted Interaction Journal (UMUAI), 22, 441–504.

Knijnenburg, B., Willemsen, M., & Kobsa, A. (2011). A Pragmatic Procedure to Support the User-Centric Evaluation of Recommender Systems. Fifth ACM Conference on Recommender Systems.

Komiak & Benbasat. (2006). The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents. MIS Quarterly, 30(4), 941. https://doi.org/10.2307/25148760

Koufaris, M., & Hampton-Sosa, W. (2004). The development of initial trust in an online company by new customers. Information & Management, 41(3), 377–397. https://doi.org/10.1016/j.im.2003.08.004

Lee, Y., Kozar, K. A., & Larsen, K. R. T. (2003). The Technology Acceptance Model: Past, Present, and Future. Communications of the Association for Information Systems, 12. https://doi.org/10.17705/1CAIS.01250

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. Information Systems Research, 13(3), 334–359. https://doi.org/10.1287/isre.13.3.334.81

McNee, S. M., Riedl, J., & Konstan, J. A. (2006). Being accurate is not enough: How accuracy metrics have hurt recommender systems. CHI ’06 Extended Abstracts on Human Factors in Computing Systems, 1097–1101. https://doi.org/10.1145/1125451.1125659

O’Donovan, J., & Smyth, B. (2005). Trust in recommender systems. Proceedings of the 10th International Conference on Intelligent User Interfaces, 167–174.

Pu, P., & Chen, L. (2007). Trust-inspiring explanation interfaces for recommender systems. Knowledge-Based Systems, 20(6), 542–556. https://doi.org/10.1016/j.knosys.2007.04.004
Pu, P., Chen, L., & Hu, R. (2011). A User-Centric Evaluation Framework for Recommender Systems. Fifth ACM Conference on Recommender System, 157–164.
Pu, P., Chen, L., & Hu, R. (2012). Evaluating recommender systems from the user's perspective: Survey of the state of the art. User Modeling and User-Adapted Interaction, 22(4–5), 317–355. https://doi.org/10.1007/s11257-011-9115-7
Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to Recommender Systems Handbook. In F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (Eds.), Recommender Systems Handbook (pp. 1–35). Springer US. https://doi.org/10.1007/978-0-387-85820-3_1
Ricci, F., Rokach, L., Shapira, B., Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender Systems Handbook. Springer.
Saunders, M., Lewis, P., Thornhill, A., Saunders, M., Lewis, P., & Thornhill, A. (2019). Research Methods for Business Students.
Söllner, M., Hoffmann, A., Hoffmann, H., Wacker, A., & Leimeister, J. M. (2014). Understanding the Formation of Trust. In K. David, K. Geihs, J. M. Leimeister, A. Roßnagel, L. Schmidt, G. Stumme, & A. Wacker (Eds.), Socio-technical Design of Ubiquitous Computing Systems (pp. 39–58). Springer International Publishing. https://doi.org/10.1007/978-3-319-05044-7_3
Söllner, M., Hoffmann, A., & Leimeister, J. M. (2016). Why different trust relationships matter for information systems users. European Journal of Information Systems, 25(3), 274–287. https://doi.org/10.1057/ejis.2015.17
Thong, J. Y. L., Hong, S.-J., & Tam, K. Y. (2006). The effects of post-adoption beliefs on the expectation-confirmation model for information technology continuance. International Journal of Human-Computer Studies, 64(9), 799–810. https://doi.org/10.1016/j.ijhcs.2006.05.001
Vaishnavi, V., Kuechler, W., & and Petter, S. (2019). Design Research in Information System. http://www.desrist.org/design-research-in-information-systems
Van Der Heijden. (2004). User Acceptance of Hedonic Information Systems. MIS Quarterly, 28(4), 695. https://doi.org/10.2307/25148660
Van Deursen, A. J., & Van Dijk, J. A. (2014). The digital divide shifts to differences in usage. New Media & Society, 16(3), 507–526. https://doi.org/10.1177/1461444813487959
Venkatesh, Morris, Davis, & Davis. (2003). User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly, 27(3), 425. https://doi.org/10.2307/30036540
Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science, 46(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926
Venkatesh, V., University of Arizona, Bala, H., & Indiana University. (2013). Bridging the Qualitative-Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems. MIS Quarterly, 37(1), 21–54. https://doi.org/10.25300/MISQ/2013/37.1.02
Wang, W., & Benbasat, I. (2007). Recommendation Agents for Electronic Commerce: Effects of Explanation Facilities on Trusting Beliefs. Journal of Management Information Systems, 23(4), 217–246. https://doi.org/10.2753/MIS0742-1222230410

Xu, J., & Cenfetelli, R. (2011). The Effects of Service and Consumer Product Knowledge on Online Customer Loyalty. Journal of the Association for Information Systems, 12(11), 741–766. https://doi.org/10.17705/1jais.00279

Xu, J., & Cenfetelli, R. T. (2013). Integrating Service Quality with System and Information Quality: An Empirical Test in the E-Service Context. MIS Quarterly, 37(3), 777–794. https://doi.org/10.25300/MISQ/2013/37.3.05

8. GENERATIVE ARTIFICIAL INTELLIGENCE AND DIGITAL TRANSFORMATION IN CONTACT CENTER BUSINESSES

Shafiq Alam
School of Management & Marketing,
Massey University, New Zealand
salam1@massey.ac.nz

Windy Dharmawan
Otago Polytechnic, Auckland
International Campus, Auckland, New Zealand
dharw1@student.op.ac.nz

Lorraine Skelton
Otago Polytechnic, Auckland International Campus, Auckland, New Zealand
lorraine.skelton@op.ac.nz

Muhammad Sohaib Ayub
Data Science Institute, University of Galway, Galway, Ireland
0133839s@universityofgalway.ie

Abstract

In today's business landscape, digital transformation is critical to competitiveness, with generative AI (GAI) offering significant potential to improve customer interactions and operational efficiency - yet its adoption in contact centres faces challenges such as technical complexity, data privacy concerns, and resistance to change. In this study, we explore these barriers by surveying contact centre personnel. Our findings reveal that security risks, potential misuse, and process complexity drive reluctance toward GAI adoption, along with a knowledge gap to maximize its benefits. The study also uncovers obstacles at all levels of employees and stresses the need for strong governance, multi-stakeholder collaboration, a focus on data ethics, bias mitigation, and risk management. We recommend addressing these challenges to ensure successful integration of GAI, providing valuable insights for organizations that navigate digital transformation while balancing innovation with responsible implementation.

Keywords:

Generative AI (GAI), Contact Centres, Digital Transformation, Adoption Barriers

1 Introduction

Contemporary businesses have prioritized digital transformation to stay competitive in the rapidly evolving technological landscape (Astapciks, 2023). Digital transformation is crucial for businesses across industries, enabling them to adapt and thrive by integrating digital tools into their operations (Bumann & Peter, 2019). This transformation involves reimagining business models, enhancing customer service, and improving overall productivity (Rogers, 2016). Technology has also heightened customer expectations, driving the need for businesses to innovate their models to meet these demands (Dwivedi et al., 2021). As businesses strive to stay ahead, emerging technologies like Generative Artificial Intelligence (GAI) are playing an increasingly pivotal role in accelerating this digital shift. GAI is emerging as a key technology that facilitates digital transformation. According to McKinsey (2023), GAI could add up to $4.4 trillion to the global economy and significantly reduce the time spent on repetitive tasks, potentially automating 60-70% of the work activities of employees. GAI, capable of generating content such as text, images, and audio based on data, presents transformative opportunities, especially in customer

interactions (Lawton, 2023), with companies such as Salesforce noting its potential to optimize customer engagement.

Contact centers are the key to modern businesses, serving as the primary hubs of customer engagement (Gartner, 2023). GAI offers excellent potential in these environments, improving participation by generating personalized responses and automating tasks such as data entry (Kimindiri, 2023). However, the adoption of GAI remains limited, with surveys indicating that 62% of organizations have not yet integrated it into their operations (DISCO, 2023). GAI's integration, particularly in customer service, is projected to become widespread between 2030 and 2060 (Chui et al., 2023).

The adoption of Generative Artificial Intelligence (GAI) in contact center environments represents a complex technological transformation influenced by multiple critical factors. This study investigates the multifaceted determinants of GAI implementation, including performance expectancy, effort expectancy, social influence, and facilitation conditions. A comprehensive review of the existing literature highlights that while the potential transformative capabilities of GAI are widely acknowledged, organizations face significant barriers in its integration. Challenges such as technological complexity, data privacy concerns, and resistance to change are recurrent themes in prior studies (e.g., Chui et al., 2023; Wach et al., 2023). Moreover, the literature suggests that GAI's impact on business models and operational efficiency is still under exploration, with varying degrees of success across industries. These obstacles encompass intricate technical complexities, nuanced data privacy considerations, and ethical dilemmas related to customer data processing and algorithmic decision-making (Wach et al., 2023). Consequently, robust governance frameworks emerge as imperative mechanisms to ensure organizational accountability, transparency, and responsible implementation of GAI technologies.

This paper explores how GAI can transform contact center operations, focusing on improving customer engagement, business models, and operational efficiency. Through this research, the study seeks to provide actionable insights for organizations navigating the complexities of GAI adoption in the contact center industry (Caliskan, Bryson, & Narayanan, 2017).

2 Background and Evolution of Generative AI

The term AI was first coined by John McCarthy in 1955, and it was initially focused on creating machines that could simulate human behavior (McCarthy et al., 1956). Over the years, AI systems evolved from narrow applications to more complex tasks but faced challenges due to tacit knowledge that was difficult to formalize (Autor, 2015). Traditional AI is focused on specific tasks using defined datasets, whereas GAI uses large, unspecific datasets, allowing it to handle open-ended tasks and generate novel content (IBM, 2023). GAI models, based on neural networks, excel at mimicking human creativity and generating new content. The size of these models has grown exponentially, with Google's PaLM model reaching 540 billion parameters (Mars, 2022). More recent estimates suggest that ChatGPT-4.0 may have trillions of parameters. Before GAI, AI was used in basic customer service tasks, such as decision trees for chatbots in 1966 (Adamopoulou & Moussiades, 2020). However, GAI's advancements in natural language processing have revolutionized customer service, enabling quicker access to real-time information and supporting agents in responding to customer queries more efficiently (Feuerriegel et al., 2023). GAI's ability

to enhance customer service and employee productivity has been shown to increase productivity by up to 37% in certain settings (Noy & Zhang, 2023), with specific tasks being completed faster when GAI tools are implemented (Eloundou et al., 2023).
Despite its potential, GAI implementation faces several barriers, including workers' insecurity about AI, a lack of capital, and insufficient IT infrastructure (Rjab et. al., 2023). These challenges are compounded by concerns over job displacement and AI's potential to perpetuate biases (Caliskan, Bryson, & Narayanan, 2017). Some argue that reskilling workers and developing ethical guidelines could help mitigate these challenges (World Economic Forum, 2020).

As GAI technology becomes more widespread, organizations must develop governance frameworks to ensure its ethical use. Issues such as data privacy, algorithmic bias, and intellectual property rights are critical in GAI adoption (Wach et al., 2023). Companies need to establish ethical standards, minimize biases in AI models, and implement governance practices that ensure accountability and proper oversight of GAI outputs (Mondal, Das, & Vrana, 2023).

3 Methodology

This study examines the role of Generative AI (GAI) in contact centers within the Australasia region, focusing on the preparedness, challenges, and considerations for its successful integration. A mixed-methods approach was employed to provide a comprehensive understanding of the subject. The quantitative data were gathered through closed-ended questions, while the qualitative data were collected through open-ended questions. The mixed-methods design allowed for an independent collection of both data types, ensuring a robust analysis of GAI adoption. The methodology is grounded in existing research frameworks (e.g., Creswell & Creswell, 2017), and the survey questions were developed based on previous studies in AI adoption (Caliskan et al., 2017). We also clarify that the purposive sampling strategy, where participants were selected based on their relevant experience and knowledge, was used to ensure the representation of experts in the field. This concurrent design allows for independent data collection for both components, enabling a more holistic perspective on the research topic. The survey was designed to assess three main themes: the utilization of GAI for business model reinvention, the obstacles faced during GAI implementation, and the governance considerations for its integration. The target population for the survey consists of managers and non-managers working in contact centres across the Australasia region. Managers were included for their strategic insights into the organization's digital transformation efforts, while non-managers provided valuable perspectives based on their daily customer service experiences. Participants were recruited using purposive sampling via email and LinkedIn invitations, ensuring that the most knowledgeable about the phenomenon were included in the study. The survey was anonymous to encourage candid responses and avoid any biases. Data collection was conducted over a one-month period, during which a diverse range of responses were gathered from contact center professionals across the Australasia region. We acknowledge the limitation of the one-month data collection period and note that a longer data collection window could provide a broader perspective. The data analysis combined both quantitative and qualitative methods. Quantitative data from the Likert scale responses were analyzed using descriptive and inferential statistics. In contrast, qualitative data from the open-ended questions were examined through thematic analysis to identify recurring patterns and themes. The use of a mixed-methods strategy enriched the findings, providing a deeper understanding of the challenges and opportunities for implementing GAI in contact centres.

4 Results and Discussion

The final sample comprised 52 valid responses of the survey conducted via a Qualtrics link shared over the email and LinkedIn, predominantly from the insurance, IT, and healthcare sectors. Respondents' managerial levels were split, with 62% in managerial roles and 38% in non-managerial positions. The sector breakdown, shown in Figure 1, highlighted that insurance, information technology, and healthcare were the top contributors. To ensure validity, the survey's content was derived from credible sources, such as Rogers (2016) and Watkins & Weissbeck (2023), while reliability was evaluated using Cronbach's alpha, focusing on six themes: customer, strategy, organization, culture, technology, and people. Table 1 shows the Cronbach's alpha results. The organization theme achieved the highest consistency at 0.85, while the people theme was not assessed due to only one survey item.

4.1 Descriptive Analysis for Quantitative Data

The first question evaluated how well contact centers utilize GAI and reinvent their business models. Results are summarized in Table 2, and key findings indicate a strong emphasis is on data and customer value, with areas of improvement in rapid-cycle innovation and GAI use-case identification.

4.1.1 Obstacles to Implementing Generative AI

The obstacles to GAI implementation were assessed through multiple responses from participants, yielding 273 responses across 52 participants. Key barriers identified include validation and reliability concerns, especially for non-managerial respondents, and resource and expertise issues for managerial respondents. The primary barriers reported are operational reliability and compatibility with existing infrastructure, with managerial respondents also highlighting concerns about resource allocation and economic constraints.

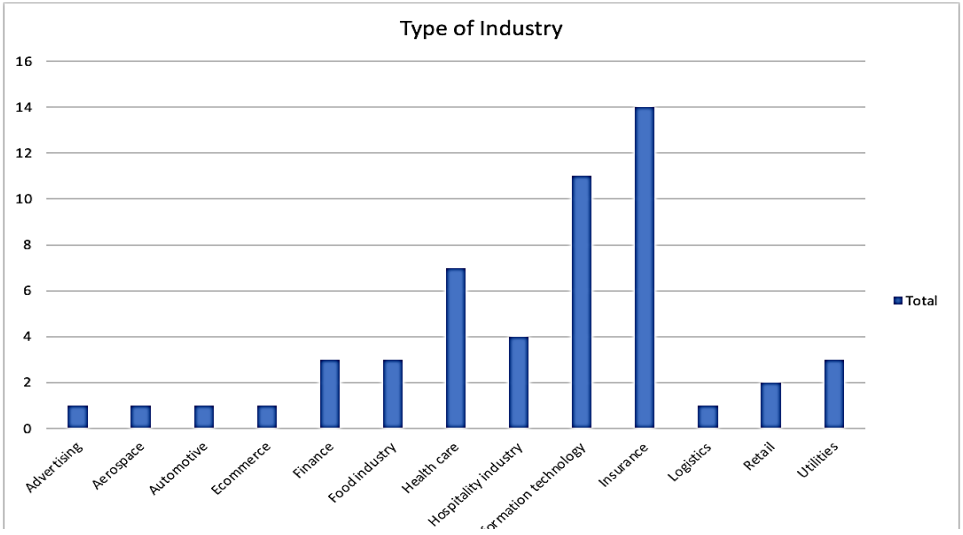

Figure 7 Respondent Composition Based on Type of Industry

4.1.2 Considerations for Implementing Generative AI and Governance

Implementing GAI and its governance is critical for establishing ethical standards, protecting data privacy, maintaining quality control, adhering to legal standards, mitigating emerging risks, and encouraging fairness (New Zealand Digital Government, 2023). The survey data collected in Table 3 help the understanding of the factors in GAI implementation and the governance practices and standards.

The descriptive statistical analysis from both Table 3 and Figure 2 revealed several noteworthy findings about the critical considerations when implementing GAI for digital transformation in contact center businesses:

Primary Considerations: Data governance, privacy, and security emerged as the top primary considerations with a mean score of 4.25 (SD=0.73). A significant majority (approximately 82.69%) of respondents either strongly agreed or somewhat agreed, with this being their key focus.
Fairness and Bias: Attention to assure fairness, bias-free operation, and diversity while employing GAI algorithms was also high, with a mean score of 4.15 (SD=0.72), where 84.61% of respondents demonstrated agreement.
Balance of Governance: Participants recognized the importance of striking the right balance between innovation and risk mitigation in governance, underscored by a mean score of 4.13 (SD=0.71), with 80.77% of respondents favoring this aspect.

.

How well is your organization able to utilize generative AI and reinvent your business model?		
Component	**Theme**	**Cronbach Alpha**
We are focused on our customers' changing digital habits and new ways of engagement.	Customer	0.76
Our value proposition is defined by changing customer needs.	Customer	
We assess new technologies by how they could create new value for our customers.	Customer	
We look to create value through platforms and external networks.	Strategy	0.75
Our data strategy is focused on how to turn data into new value.	Strategy	
We manage our data as a strategic asset that we are building over time.	Strategy	
We make decisions through experiments and testing wherever possible.	Organization	0.85
We innovate in rapid cycles, using prototypes to learn quickly.	Organization	
We aim to adapt early to stay ahead of the curve of change.	Culture	0.70
We are able to seed and develop new ideas that are unusual for our business.	Culture	
We are aware of Generative AI and its potential impact.	Technology	0.70
We are able to identify the actual use case of Generative AI that leverages our business.	Technology	
Generative AI applications help in cost efficiency.	Technology	
Combination of employee experience with Generative AI able to augment productivity and deliver.	People	*

Table 1: Cronbach's Alpha Results for Survey Themes

The composition of Likert scale results is shown in Figure 2 to provide graphical clarity and weight for each component. Based on the results, the least regarded factors are:

- **Intellectual Property:** Intellectual property issues related to GAI sourcing had the lowest mean score of 3.67, with only 7.84% disagreeing, indicating that it might not be a significant concern, illustrating potential gaps in understanding or prioritizing this aspect.
- **Regulatory and Legal Awareness:** Awareness about regulatory and legal considerations while developing the governance framework came second to the bottom (mean score: 3.73), suggesting certain unawareness of regulatory and legal structure for governing AI with 11.54% of respondents somewhat disagree.
- **Monitoring and Evaluation:** Continuous monitoring and evaluation of the performance and impact of GAI systems throughout the digital transformation process was slightly disregarded by the respondents, indicated by a meager disagreement rate of 1.92% and a mean score of 4.00.

#	Field	Minimum	Maximum	Mean	Std Dev	Variance	Count
1	We are focused on our customers' changing digital habits and new ways of engagement.	1.00	5.00	3.69	0.97	0.94	52
2	We look to create value through platforms and external networks.	1.00	5.00	3.79	0.95	0.90	52
3	Our data strategy is focused on how to turn data into new value.	2.00	5.00	3.85	0.91	0.82	52
4	We manage our data as a strategic asset we are building over time.	2.00	5.00	3.96	0.88	0.77	52
5	We make decisions through experiments and testing wherever possible.	1.00	5.00	3.31	1.08	1.17	52
6	We innovate in rapid cycles, using prototypes to learn quickly.	2.00	5.00	3.15	0.99	0.98	52
7	Our value proposition is defined by changing customer needs.	1.00	5.00	3.81	0.94	0.89	52
8	We assess new technologies by how they could create new value for our customers.	2.00	5.00	4.15	0.82	0.67	52
9	We aim to adapt early to stay ahead of the curve of change.	2.00	5.00	3.60	1.00	1.01	52
10	We are able to seed and develop new ideas that are unusual for our business.	1.00	5.00	3.17	0.93	0.87	52
11	We are aware of Generative AI and its potential impact.	2.00	5.00	3.81	0.96	0.92	52
12	We are able to identify the actual use case of generative AI that reinvents our business.	2.00	5.00	3.15	1.06	1.13	52
13	Combination of employee experiences with generative AI able to augment productivity and deliver the expected response.	1.00	5.00	3.52	1.12	1.25	52
14	Generative AI applications help in cost efficiency.	1.00	5.00	3.41	0.93	0.87	51

Table 2: Descriptive Statists on Organizational Use of GenAI and Business Model Reinvention

These findings demonstrate an overarching emphasis on data considerations and the humanistic aspect of AI usage (bias, fairness). At the same time, there exists relative oversight on issues like intellectual property, regulatory considerations, and continuous monitoring.

4.2. Inferential Analysis for Quantitative Data

There may be different considerations for implementing GAI for digital transformation between non-managerial and managerial levels in the data sample. The p-values obtained from independent T-tests will indicate if there are statistically significant differences between the managerial and non-managerial views for each component of the survey questions.

The independent T-test result for the first question, "How well is your organization able to utilize GAI and reinvent your business model?" is shown in Table 4.

What considerations should organizations take into account when implementing generative AI for digital transformation, particularly with regards to its governance?							
#	**Field**	**Min**	**Max**	**Mean**	**Std Dev**	**Variance**	**Count**
1	The risk of malicious information that could affect the output and lead to an incident or spark criticism from society.	1.00	5.00	4.12	0.89	0.79	52
2	The issue of intellectual property sourced by generative AI.	1.00	5.00	3.67	0.98	0.97	51
3	Transparency and responsible decision-making processes for the application of generative AI.	3.00	5.00	4.10	0.60	0.36	52
4	Involvement of all stakeholders in the development of generative AI governance during the digital transformation process.	2.00	5.00	4.04	0.76	0.58	52
5	Striking the correct balance of governance between innovation and risk mitigation.	3.00	5.00	4.13	0.71	0.50	52
6	There are regulatory and legal considerations that we are aware of while developing the governance framework.	2.00	5.00	3.73	0.98	0.97	52
7	Data governance, privacy, and security are the primary considerations.	3.00	5.00	4.25	0.73	0.53	52
8	How to assure fairness, bias-free, and diversity while employing generative AI algorithms.	2.00	5.00	4.15	0.72	0.51	52
9	Continuous monitoring and evaluation of the performance and impact of generative AI systems throughout the digital transformation process.	2.00	5.00	4.00	0.83	0.69	52

Table 3: Descriptive Statistical Results on Considerations for GenAI Implementation and Governance

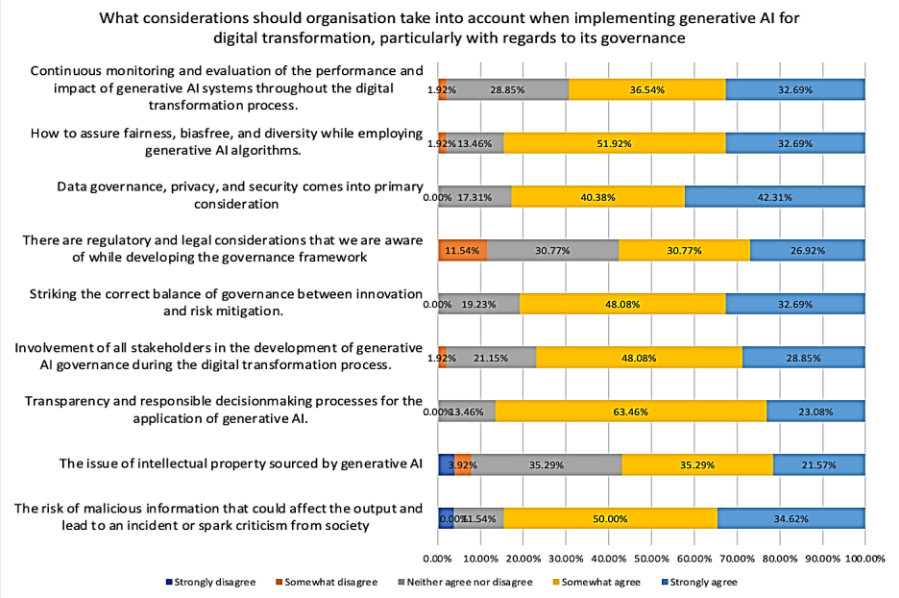

Figure 2: Likert Scale Composition Result of the Fifth Question

The result indicates no significant difference between managerial and non-managerial perception in the utilization of GAI and reinvention of their business model. In conclusion, there is a holistic alignment in the perception of managerial and non-managerial staff on the organizational use of GAI and its potential to reinvent business models. The independent T-test result for the fifth question, "*What considerations should organizations take into account when implementing GAI for digital transformation, particularly with regards to its governance?*" is shown in Table 5.

Field	Non-Managerial			Managerial			P-value
	Mean	Std Dev.	Variance	Mean	Std Dev.	Variance	(independent t-test)
We are focused on our customers' changing digital habits and new ways of engagement.	3.45	1.10	1.21	3.84	0.88	0.78	0.19
We look to create value through platforms and external networks.	3.79	0.96	0.92	3.78	0.97	0.95	0.95
Our data strategy is focused on how to turn data into new value.	3.85	0.92	0.84	3.78	0.94	0.89	0.52
We manage our data as a strategic asset that we are building over time.	3.96	0.88	0.78	3.91	0.93	0.86	0.56
We make decisions through experiments and testing wherever possible.	3.31	1.09	1.20	3.31	1.09	1.19	0.97
We innovate in rapid cycles, using prototypes to learn quickly.	3.15	1.00	1.00	3.19	1.06	1.13	0.75
Our value proposition is defined by changing customer needs.	3.81	0.95	0.90	3.84	0.95	0.91	0.73
We assess new technologies by how they could create new value for our customers.	4.15	0.83	0.68	4.25	0.92	0.84	0.26
We aim to adapt early to stay ahead of the curve of change.	3.60	1.01	1.03	3.53	1.16	1.35	0.52
We are able to seed and develop new ideas that are unusual for our business.	3.17	0.94	0.89	3.09	1.00	0.99	0.43
We are aware of Generative AI and its potential	3.81	0.97	0.94	3.88	0.94	0.89	0.54

impact.							
We are able to identify the actual use case generative AI that reinvents our business.	3.15	1.07	1.15	3.16	1.14	1.30	0.98
Combination of employee experiences with generative AI able to augment productivity and deliver expected response.	3.52	1.13	1.27	3.56	1.16	1.35	0.73
Generative AI applications help in cost efficiency.	3.41	0.94	0.89	3.39	0.80	0.65	0.83

Table 4: Independent T-Test Result on the First Question about use of generative AI.

.

Based on the p-values, all of them are considerably higher than 0.05, the usual cut-off for significance. This suggests there is no substantial difference between the perspectives of the managerial and non-managerial levels on these topics. This uniformity could be positive, as it implies shared understanding and consensus in the organization.

4.3 Thematic Analysis of Qualitative Data

This section presents the thematic analysis of qualitative responses to the open-ended question: 'Can you please elaborate on your perspective regarding the utilization of GAI in reinventing your business model?' Following Braun and Clarke's (2022) framework, we conducted inductive coding to identify emergent themes, which were then analyzed to reveal patterns in participants' perceptions of GAI's role in business transformation. Table 6 summarizes the eight key themes that emerged, ranging from skepticism about GAI's transformative potential to its operational benefits and implementation challenges. Representative quotes from respondents are integrated throughout the analysis to ground these themes in the data, providing nuanced insights into how contact center professionals view GAI's opportunities and limitations

Field	Non-Managerial			Managerial			P-value
	Mean	Std Dev.	Variance	Mean	Std Dev.	Variance	(independent t-test)
The risk of malicious information that could affect the output and lead to an incident or spark criticism from society.	4.10	0.97	0.94	4.13	0.87	0.76	0.93
The issue of intellectual property sourced by generative AI.	3.67	0.99	0.99	3.61	0.95	0.91	0.64
Transparency and responsible decision-making processes for the application of generative AI.	4.10	0.60	0.36	4.09	0.59	0.35	0.97
Involvement of all stakeholders in the development of generative AI governance during the digital transformation process.	4.04	0.77	0.59	4.16	0.72	0.52	0.18
Striking the correct balance of governance between innovation and risk mitigation.	4.13	0.71	0.51	4.25	0.67	0.45	0.16
There are regulatory and legal considerations that we are aware of while developing the governance framework.	3.73	0.99	0.98	3.84	1.05	1.10	0.29
Data governance, privacy, and security are the primary considerations.	4.25	0.74	0.54	4.25	0.72	0.52	1.00
How to assure fairness, bias-free, and diversity while employing generative AI algorithms.	4.15	0.72	0.52	4.13	0.71	0.50	0.73
Continuous monitoring and evaluation of the	4.00	0.84	0.71	4.06	0.84	0.71	0.51

performance and impact of generative AI systems throughout the digital transformation process.							

.

Table 5: Independent T-Test Result on the Fifth Question about GenAI governance.

Codes	Common Themes
Skepticism about GAI's impact on business;	Skepticism about GAI's potential to reinvent business model
Value in timesaving;	GAI as a tool for efficiency
GAI for better omnichannel customer service;	The use of GAI in improving customer service
Old tech, expensive to change;	The cost and technical complexity involved in
Reduction in workforce	GAI's potential in reducing workforce
Applicable for simple cases; The need for human intervention for complex cases	GAI's limited capability on complex cases

Table 6: Codes and Common Themes from the Second Question

- **Skepticism about Generative AI's potential to reinvent business models:** Five respondents indicated doubt about GAI's capacity to alter business models drastically. For example, a participant said, "Not sure about the usage will reinvent our business because it only one part, still have people and process other than technology" (#21).
- **Generative AI as a tool for efficiency**: The other five responses acknowledged the potential of GAI for creating efficiencies within existing structures, such as speeding up service times or streamlining processes. As respondent #17 noted, "It is both an opportunity to have lean operations, hence managing costs efficiently, and to improve customer experience."
- **The use of Generative AI in improving customer service**: Five respondents indicated that GAI was seen as a valuable asset for improving customer service operations. For instance, respondent #4 said, "Generative AI empowers teams to provide omnichannel customer service and better customer experiences using a platform from a major software company."
- **The cost and technical complexity of Generative AI integration:** Four respondents raised concerns about the cost and technical challenges of integrating GAI into existing systems. As respondent #20 mentioned, "Generative AI is promising, but implementing it could be costly and complex, making it unsuitable for SMEs with limited resources."
- **Generative AI's potential to reduce manpower:** Two participants noted that GAI could reduce the need for human personnel. Respondent #14 stated, "Reducing manpower as GAI could handle customer queries."
- **Generative AI's capability with complex cases:** Three respondents expressed skepticism about GAI's ability to handle complex scenarios. As respondent #12 noted, "GAI may struggle with technical complexity and integration into business models, requiring human judgment."
- **Generative AI's role in medium- and long-term strategies:** One participant saw GAI as part of a long-term strategy. Respondent #5 said, "Not now, but in the medium term, GAI could have a significant impact."
- **The need for human judgment given GAI's limitations:** Three respondents emphasized the

need for human intervention in certain tasks. As respondent #27 noted, "For complex tasks like customized terms and conditions, GAI requires human creativity and judgment."

5 Discussion

This section analyzes and discusses the results and findings based on the identified themes from both quantitative and qualitative data. We compare our results with existing studies on GAI adoption in contact centers (e.g., Feuerriegel et al., 2023; Noy & Zhang, 2023) and discuss the implications of these findings in the context of ongoing technological advancements. We also highlight areas where the findings may differ from previous studies, such as the challenges around GAI integration and its impact on business model reinvention.

5.1 Utilization of Generative AI and Business Model Reinvention

The analysis of quantitative data shows that contact centers recognize the potential of data utilization and GAI in enhancing customer service (Mean=4.15). There is a strategic approach to managing data (Mean=3.96) and aligning it to create new customer value (Mean=3.85). This data-driven mindset forms a solid foundation for GAI integration.

However, a significant gap exists in fully exploiting GAI for business transformation. The lower mean scores for aspects such as rapid innovation (Mean=3.15), identifying transformative AI applications (Mean=3.15), and fostering unconventional ideas (Mean=3.17) indicate that contact centers still face challenges in leveraging GAI to its full potential for business model reinvention.

This gap is reflected in the qualitative data, where GAI is primarily seen as an efficiency tool rather than a driver of business transformation. Concerns about security, misuse, and GAI's inability to handle complex issues also highlight barriers to its broader adoption. A reliance on human judgment suggests a persistent preference for traditional decision-making, limiting GAI's potential in reinventing business models.

In conclusion, while contact centers show interest in using GAI for operational efficiencies, doubts persist regarding its ability to drive comprehensive business model reinvention. The existing readiness for GAI integration is promising but insufficient for a complete transformation of traditional contact center business models.

5.2 Obstacles to Implementing Generative AI

Both the quantitative and qualitative data reveal several key obstacles in the implementation of GAI. These challenges include:

- Resource Allocation and Organizational Readiness: Both data sets indicate that sufficient resources and preparedness are essential. Financial concerns were raised by 17 manag rial and 11 non-managerial respondents, underlining the importance of financial resources and expertise.
- Building Competencies and Skills in AI: A significant concern, especially among man- agerial staff (21 responses), relates to the lack of AI skills. The need for in-house expertise or external consultants was highlighted.
- Managing and Leveraging Data Effectively: Data management, including issues of avail- ability, quality, and relevance, emerged as a prominent theme. This was echoed by 11 managerial and

10 non-managerial staff in the quantitative data.

- Aligning GAI Implementation Strategy with Business Model: Both data sets point to the challenge of integrating GAI into existing business models. For example, 16 managerial and 9 non-managerial respondents expressed concerns about compatibility with current IT infrastructure.
- Striking a Balance between Human and AI Roles: The qualitative data reflects the quantitative concern regarding the reliability of GAI outputs. Both managerial and non- managerial staff (15 and 17, respectively) indicated that human supervision is essential for complex issues.

The overlap between managerial and non-managerial perceptions of obstacles suggests shared concerns, such as resources and data management, but also highlights differences in priorities. Addressing these obstacles requires collaboration across levels to ensure successful GAI implementation in contact centers.

5.3 Considerations for Implementing Generative AI and Governance

The findings emphasize critical considerations concerning data governance, privacy, and security in implementation of GAI. With a mean score of 4.25 (SD=0.73), 82.69% of respondents expressed concern about the risks associated with managing large data sets used to train GAI systems. There is a clear awareness of ethical obligations and the potential reputational and financial consequences of data mismanagement.

The importance of fairness, reducing bias, and fostering diversity in GAI systems was also highlighted, with a mean score of 4.15 (SD=0.72), indicating a strong understanding of the societal implications of AI. Furthermore, the emphasis on balancing innovation and risk management (Mean=4.13, SD=0.71) reflects the need for a nuanced governance approach. However, less attention is given to intellectual property (Mean=3.67) and regulatory/legal concerns (Mean=3.73). This suggests that businesses may not fully understand the legal complexities associated with GAI and could face future legal challenges.

In summary, while there is a solid grasp of the ethical and societal aspects of GAI governance, there are gaps in understanding intellectual property, legal frameworks, and continuous monitoring. Addressing these gaps will be crucial for ensuring safe and effective GAI implementation.

6 Conclusions, Recommendations, and Limitations

Contact centers are interested in adopting GAI for operational efficiency, but there are concerns about its role in business model reinvention, its ability to handle complex issues, and potential security and misuse risks. The reliance on human judgment for complex decisions further complicates the transition to GAI-driven business models.

The gaps identified in innovation cycles, use case identification, and the fostering of unconventional ideas suggest that the current state of readiness is not sufficient for a full-scale transformation. To achieve a successful integration of GAI into business models, strategic investments, comprehensive planning, and risk management are required to address the challenges of implementation.

Companies should focus on enhancing awareness of GAI's strategic potential, moving beyond operational efficiency to its value-creating abilities. Upskilling staff to bridge the knowledge gap and to conduct incremental implementation in low-risk projects can help build confidence before scaling up. Collaboration between managerial and non-managerial staff should be encouraged to provide diverse insights.

Developing a robust data strategy to ensure data quality, availability, and compliance with governance regulations is crucial for GAI's success. Furthermore, stronger governance mechanisms should be implemented to address intellectual property, legal compliance, and continuous monitoring of GAI systems.

Future research can address the limitation of relatively small sample size which may not fully capture the breadth of GAI implementation across different industries. Additionally, the complexity of GAI poses challenges in comprehensively assessing its full impact within contact centers. Future studies should consider larger sample sizes and explore longitudinal data to better understand the long-term effects and potential scalability of GAI adoption in various sectors.

References

Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, 2, 100006. https://doi.org/10.1016/j.mlwa.2020.100006

Astapciks, I. (2023). Why do companies need digital transformation? *Forbes*. https://www.forbes.com/sites/forbestechcouncil/2023/03/20/why-do-companies-need-digital-transformation/

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*, 29(3), 3–30. https://doi.org/10.1257/jep.29.3.3

Braun, V., & Clarke, V. (2022). Thematic analysis: A practical guide. *QMiP Bulletin*, 1(33), 46–50. https://doi.org/10.53841/bpsqmip.2022.1.33.46

Bumann, J., & Peter, M. (2019). Action fields of digital transformation: A review and comparative analysis of digital transformation maturity models and frameworks. *Digitalisierung und andere Innovationsformen im Management*, 2, 13–40.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183–186. https://doi.org/10.1126/science.aal4230

Chui, M., et al. (2023). The economic potential of generative AI: The next productivity frontier. *McKinsey & Company*. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

DISCO. (2023). Data shows majority of consumers have been slow to adopt generative AI tools for work-related tasks. *DISCO*. https://csdisco.com/pressrelease/data-shows-majority-of-consumers-have-been-slow-to-adopt-generative-ai

Dwivedi, Y. K., et al. (2021). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management*,

59, 102168. https://doi.org/10.1016/j.ijinfomgt.2020.102168
Eloundou, T., et al. (2023). GPTs are GPTs: An early look at the labor market impact potential of large language models. *arXiv*. https://doi.org/10.48550/arxiv.2303.10130
Feuerriegel, S., et al. (2023). Generative AI. *SSRN*. https://doi.org/10.2139/ssrn.4443189
Gartner. (2023). Definition of contact center. *Gartner Information Technology Glossary*. https://www.gartner.com/en/information-technology/glossary/contact-center
IBM. (2023). What is artificial intelligence (AI)? *IBM*. https://www.ibm.com/topics/artificial-intelligence
Kimindiri, M. (2023). What is generative AI and how does it apply to call centers? *The Level AI*. https://thelevel.ai/blog/generative-ai/
Lawton, G. (2023). What is generative AI? Everything you need to know. *TechTarget*. https://www.techtarget.com/searchenterpriseai/definition/generative-AI
Mars, M. (2022). From word embeddings to pre-trained language models: A state-of-the-art walkthrough. *Applied Sciences*, 12(17), 8805. https://doi.org/10.3390/app12178805
McCarthy, J., et al. (1956). A proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *Dartmouth College*. http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.htm
McKinsey. (2023). The economic potential of generative AI: The next productivity frontier. *McKinsey & Company*. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier
Mondal, S. R., Das, S., & Vrana, V. (2023). How to bell the cat? A theoretical review of generative artificial intelligence towards digital disruption in all walks of life. *Technologies (Basel)*, 11(2), 44. https://doi.org/10.3390/technologies11020044
New Zealand Digital Government. (2023). Managing the risks of GenAI to the public service. *Digital Government*. https://www.digital.govt.nz/standards-and-guidance/technology-and-architecture/interim-generative-ai-guidance-for-the-public-service/managing-the-risks-of-genai-to-the-public-service/
Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *SSRN*. https://doi.org/10.2139/ssrn.4375283
Rjab, A. B., Mellouli, S., & Corbett, J. (2023). Barriers to artificial intelligence adoption in smart cities: A systematic literature review and research agenda. *Government Information Quarterly*, 101814. https://doi.org/10.1016/j.giq.2023.101814
Rogers, D. M. (2016). *The digital transformation playbook: Rethink your business for the digital age*. https://dlibrary.ittelkom-pwt.ac.id/index.php?p=show_detail&id=10834
Wach, K., et al. (2023). The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. *Entrepreneurial Business and Economics Review*, 11(2), 7–30. https://doi.org/10.15678/eber.2023.110201
Watkins, M. D., & Weissbeck, R. (2023). Implementing generative AI: A strategic guide to overcoming the challenges. *IMD*. https://www.imd.org/ibyimd/strategy/implementing-generative-ai-a-strategic-guide-to-overcoming-the-challenges/
World Economic Forum. (2020). We need a global reskilling revolution – here's why. *World Economic Forum*. https://www.weforum.org/agenda/2020/01/reskilling-revolution-jobs-future-skills/

9. LOOKING BEYOND TECHNOLOGY: HIERARCHICAL LIST OF PEOPLE AND CULTURE CAPABILITIES FOR ADVANCED DATA ANALYTICS

Tina Afshar Ghochani
Middle East Technical University
tina.ghochani@metu.edu.tr

Özden Özcan-Top
Middle East Technical University
ozdenoz@metu.edu.tr

Banu Aysolmaz
Eindhoven University of Technology
b.e.aysolmaz@tue.nl

Abstract

Advanced data analytics (ADA), as an emerging field, empowers companies in gaining competitive advantages by offering insights derived from data. To derive business value from ADA, organizations must develop various capabilities across different areas, including infrastructure (technology), data management, people, and culture. Previous literature has consistently emphasized technical aspects over other dimensions, providing comparatively limited information about capabilities related to people and culture. To fill in this gap, we conducted a systematic literature review of studies on capabilities and related concepts including maturity and success factors of ADA. We synthesized the findings from 29 studies into a hierarchical list, categorizing capabilities at two levels and delineating each capability with a set of components. Through the presentation of an extensive and structured list, we aim to clarify how these capabilities are constituted but also guide organizations in identifying necessary capabilities and planning their development, thereby facilitating improved implementation and management of ADA.

.

Keywords

Big Data, Data Science, People, Culture, Capability, Data Analytics

1 Introduction

In today's data-driven world, businesses are dealing with vast amounts of data from various sources such as customer transactions, social media interactions, and operational processes. Making sense of this data and extracting valuable insights has become essential for organizations seeking to gain a competitive edge. This is where advanced data analytics (ADA) comes into play. ADA refers to the use of various advanced analytic techniques on data to extract insights, address business questions, develop solutions and predict the results of those solutions (Bose, 2009). ADA can be leveraged to turn big data into smart data, enabling the extraction of crucial information from large datasets. Consequently, smart data delivers actionable insights and enhances decision-making processes for organizations and companies (Hariri et al., 2019). For example, ADA is widely used in manufacturing to integrate real-time data and expert knowledge, enhancing production efficiency and product quality (Voigt et al., 2021).

Organizations must leverage diverse capabilities within their organizations to create business value from ADA (Günther et al., 2017; Gupta & George, 2016). These capabilities encompass the availability of various resources including tangible ones (e.g., sufficient investments, adequate time, data and technology), human skills (e.g., technical and managerial skills and intangible resources (e.g., organizational learning and data-driven culture) (Korsten et al., 2024). Up until now, research has predominantly emphasized technical capabilities such as infrastructure,

intelligence, and analytics tools, but has provided rather an abstract view on non-technical areas, specifically people and culture in organizations (Demir et al., 2024; Enholm et al., 2021; Mikalef et al., 2018). People capability refers to the skills and proficiency of an organization's workforce, covering a range of abilities such as managerial, technical, business acumen, relational skills, and expertise in analytics (Moseneke & Brown, 2022). Culture capability can be defined as the set of values, beliefs, norms, and behaviors within an organization that shape its approach to big data analytics, including attitudes towards data-driven decision making, innovation, collaboration, and learning (Moseneke & Brown, 2022). While these non-technical capabilities are as important as technical ones (Korsten et al., 2024; Mikalef & Krogstie, 2020), organizations cannot find sufficient guidance on improving them (Mikalef & Krogstie, 2020).

To fill this gap, the objective of this study is to provide a hierarchical list of people and culture capabilities for ADA including their detailed components. To achieve this objective, we performed a Systematic Literature Review (SLR) of people and culture related capabilities for various ADA applications. To ensure that we reveal an extensive set of capabilities and identify their details, i.e., components, we also include literature on maturity models and success factors related to ADA. Based on our thematic coding and structuring of SLR results following the approach of Gioia et al. (2013), we present a hierarchical list of people and culture capabilities, which includes a hierarchical categorization of their components for ADA. This hierarchical list can help practitioners in their decision-making for improving their capabilities in ADA projects. Our study contributes to research by focusing exclusively on the people and culture capabilities of ADA projects and providing a hierarchical and detailed view of these capabilities by capturing knowledge not only from capability literature but also from success factor and maturity model literature.

The rest of this paper is structured as follows. In Section 2, the related background for ADA capabilities is discussed. In Section 3, the research method for conducting the SLR is described. Section 4 provides the hierarchical list. The discussion, conclusions, and suggestions for future work are given in Sections 5 and 6, respectively.

2 Background and Related Work

2.1.Advanced Data Analytics (ADA)

ADA involves transforming raw data, such as textual information and web data; into high-level knowledge by identifying patterns or models from the observed data. This process uses predictive techniques to uncover patterns within these data forms, which are essential for informed decision-making (Bose, 2009). ADA utilizes sophisticated data analytics techniques to extract actionable insights and valuable knowledge from complex datasets (Brinch et al., 2020). It includes advanced statistical methods, machine learning algorithms, and predictive modelling approaches to uncover patterns, trends, and relationships within the data for informed decision-making and business value (Brinch et al., 2020). Advanced analytics-driven data examinations enable businesses to attain a comprehensive understanding of their operations and customers. The knowledge derived from these analyses is subsequently utilized to direct, streamline, and automate decision-making processes, facilitating the successful attainment of organizational objectives (Bose, 2009).

2.2 Data analytics capabilities

Diverse definitions and classifications can be found for ADA in the literature. Gupta et al. defines big data analytics (BDA) capability as “a firm’s ability to assemble, integrate, and deploy its big

data-based resources" (Gupta & George, 2016). These resources are categorized into three including tangible (i.e., data, technology, and basic resources, e.g., sufficient investments, adequate time), human skills (i.e. managerial and technical big data skills), and intangible resources (i.e. data-driven culture and the intensity of organizational learning) (Korsten et al., 2022). Leveraging BDA capabilities in a firm is confirmed to have a positive impact on its market and operational performance (Korsten et al., 2024).

Data analytics capabilities have been conceptualized based on IT capabilities and a measurement tool is provided to assess the effects of data analytics capabilities (Saltz & Shamshurin, 2016). Data analytics capabilities can be defined as constructs encompassing three major dimensions: infrastructure, personnel expertise, and relationship (Saltz & Shamshurin, 2016). Six capability areas—IT, process, performance, human, strategic, and organizational capabilities— were identified through an extensive review of literature across various domains, including big data analytics, ADA, IT, and business process management (Arunachalam et al., 2018). These capability areas were refined into the following categories in the capability maturity model for ADA: People & Culture, Performance & Value, Strategy, Data & Governance, and Process Design & Collaboration (Gupta & George, 2016).

In summary, people and culture capabilities have been defined as separate and prominent capabilities required for ADA, distinguished from technical capabilities (Enholm et al., 2021; Mikalef et al., 2018). However, literature has provided less details on people and culture capabilities, leaving these areas understudied (Korsten et al., 2024; Mikalef & Krogstie, 2020).

2.3 Related work

We have identified nine SLRs about the use and success of ADA and examined the extent they address our research objective. Saltz & Shamshurin (2016) focused on critical success factors (CSFs) for the successful implementation of data analytics rather than capabilities. Arunachalam et al. (2018) proposed a maturity model for BDA capabilities in supply chain management, addressing a similar objective but in a specific domain. Mikalef et al. (2018) presented a theoretical *framework* for enhancing business value and competitive performance through big data utilization, but their study did not focus on other ADA applications and may not be generally applicable due to industry specific differences. Al-Sai *et al.* (2019) examined the challenges of applying big data from various aspects including people and organization dimensions, but they do not introduce a capability perspective and do not report the details of the examined papers. Hattingh et al. (2019) proposed a model to develop a successful data science workforce, partially addressing our objective with its competency focus. Surbakti et al. (2020) explored the factors influencing the effective use of big data and synthesized them into a research framework; However, they encompass a narrow focus solely on organizational-level factors. Al-Sai et al. (2020) identified the CSFs of BDA projects categorized into different dimensions. Demir et al. (2024) also identified CSFs for a broad coverage of data analytics applications. While CSF is a related topic, organizations cannot directly derive capabilities for ADA success based on CSFs. Korherr & Kanbach (2021) presented a holistic framework of human-related capabilities essential for effective BDA implementation, with a focus on facilitating decision-making processes but they lack a coverage of the taxonomy across different ADA fields and the exploration of detailed capabilities. Therefore, relevant SLR studies do not fully address our objective and are not recent, warranting the need for an SLR.

3 Research Design

The SLR process we followed includes three phases: planning, conducting and reporting (Kitchenham & Charters, 2007). In the planning phase, we defined the search libraries and search criteria. Scopus, ACM, IEEE Xplore, and Google Scholar digital libraries were chosen since they are widely acknowledged for their reliability and credibility. The following keyword search was utilized across these libraries: ("data analytic*" OR "data science" OR "data driven") AND (capability OR maturity OR "success factor*") AND (people OR culture OR team). In total, 1551 articles were retrieved (i.e., 551 from Scopus, 500 from ACM, 200 from IEEE, and 300 from Google Scholar). We included peer-reviewed studies that directly refer to people and culture aspects about ADA in a variety of ways, including BDA/AI capabilities and resource-based view of the firms/dynamic capabilities, maturity models for Artificial Intelligence (AI), BDA, Business Analytics (BA), and similar technologies, and CSFs and challenges related to ADA. We excluded studies focusing on the relationship between people and organizational elements. For instance, studies investigating the relationship between components of firm performance and BDA capabilities were excluded because they were not directly facilitating our research objective to reveal related capabilities. We also removed studies which were SLRs themselves but used them for snowballing. Finally, we excluded studies which did not include an empirical research method but were rather conceptual in nature.

We followed a two-step screening process. After removing the duplicates, each article was screened using the inclusion and exclusion criteria in two rounds. At each round, a sample of 10 articles were screened by all authors separately, and the decisions were compared and discussed. This selective evaluation ensured a mutual agreement and understanding of the relevance. The rest of the screening was performed by the first author. First, the articles were screened based on the title and abstract, which reduced the set to 162 articles. Finally, the remaining articles were reviewed based on their full text, which lead to a final set of 29 articles. We analyzed and synthesized the selected articles through content analysis and coding (Krippendorff, 2018). We used the Gioia methodology (Gioia et al., 2013) to structure the resulting codes into three-level components. Initially, one of the authors developed a set of preliminary codes, drawing from existing literature on people and culture components of ADA. Then, the same author identified statements mentioning people and culture-related aspects of ADA, which can be associated with capabilities, and either assigned statements to existing codes or created new ones. Based on the detail level of the statement, the code could be related to any of the levels. Subsequently, another author refined these codes, collaboratively enhancing the depth and breadth of the analysis in multiple sessions. Finally, the third author reviewed and suggested changes to the components, which were performed based on consensus by all authors. This collaborative approach allowed us to thoroughly examine capability-related aspects, gain a comprehensive understanding of the various existing components of related capabilities for ADA, and organize them into detailed components.

4 Results

For the final paper included 29 articles, Among these articles, 23 are journal papers, and the remaining six are conference papers. We present the results in Table 5 and Table 6, which were developed following the coding approach that we described in Section 3. The left most column presents the high-level capabilities that organizations need to develop, which are detailed as

second-level capabilities in the second column. The 3rd column presents the components that make up each capability, for example, specific components of the technical knowledge and skills capability include programming, statistical analysis and data management. While capabilities identify competencies that an organization should possess, components address the specific aspects of these capabilities and in this way, can provide guidance on achieving these capabilities. The range of components that should be possessed related to a capability depends on the organizational context and business domain. For example, the set of components for technical knowledge and skills depends on the type of ADA. Capabilities and components are ordered in decreasing number of papers. For some papers, we extracted information on a capability level while no specific information in component level was found. In this case, references are indicated only for the capability.

4.1. People capabilities and components for ADA

The hierarchical list of people capabilities and components are listed in Table 5, along with the corresponding reference numbers. Three high-level capabilities are described in total by 15 capabilities. The capabilities are further divided into a total of 57 components.

High-level Capabilities	Capabilities	Components
Knowledge and Skills 10 capabilities, 15 references	Technical knowledge and skills 15 components, 12 references	Programming (Ashaari et al., 2021; Kim et al., 2018; Mandal, 2018; Upadhyay & Kumar, 2020; Vidgen, 2014; Vidgen et al., 2017; H. J. Watson, 2014) Distributed computing (Akhtar et al., 2019; Ashaari et al., 2021; Kim et al., 2018; Mandal, 2018; Upadhyay & Kumar, 2020) Statistical analysis (Akhtar et al., 2019; Kim et al., 2018; Popovič et al., 2018) Data integration, analysis, and presentation (Moseneke & Brown, 2022; G. J. Watson et al., 2021; ZareRavasan, 2023) Data management and maintenance (Ashaari et al., 2021; Upadhyay & Kumar, 2020) Decision support systems (Ashaari et al., 2021; Upadhyay & Kumar, 2020) Database management (Mandal, 2018; Moseneke & Brown, 2022) Math and operation research (Akhtar et al., 2019; Kim et al., 2018) Machine learning/Big data (Kim et al., 2018; G. J. Watson et al., 2021) Optimization (Akhtar et al., 2019) Network analysis (Akhtar et al., 2019) Cloud computing (G. J. Watson et al., 2021) 5G (G. J. Watson et al., 2021) Blockchain (G. J. Watson et al., 2021) Robotics (G. J. Watson et al., 2021)
	Managerial knowledge and skills 7 components, 6 references	Being effective regarding managing projects (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020) Having communication capabilities (G. J. Watson et al., 2021; Zamora & Carlos Barahona, 2016) Understanding business needs of stakeholders such as functional managers, suppliers, and customers (Gupta & George, 2016; Zamora & Carlos Barahona, 2016) Collaborating with stakeholders to determine ADA opportunities (Gupta & George, 2016) Coordinating ADA-related activities among employees and stakeholders (Gupta & George, 2016) Having a good sense of where to apply big data (Gupta & George, 2016) Understanding and evaluating the output extracted from big data (Gupta & George, 2016)
	Leadership knowledge and skills 6 components, 1 reference	Having growth mindset toward continuous learning and self-improvement (G. J. Watson et al., 2021) Acting with agility and decisiveness (G. J. Watson et al., 2021) Engaging with ethical and moral issues (G. J. Watson et al., 2021) Building networking and relationship skills (G. J. Watson et al., 2021) Developing knowledge of data analytics (G. J. Watson et al., 2021) Updating capability for new technologies (G. J. Watson et al., 2021)
	Business/ Domain knowledge and skills 5 components, 4 references	Being capable of interpreting business problems and frame suitable analytical solutions based on available knowledge (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020; ZareRavasan, 2023)

		Being knowledgeable about business functions (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020) Being knowledgeable about the business environment (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020) Understanding business policies and plans (Mandal, 2018; Upadhyay & Kumar, 2020) Developing appropriate solutions (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020)
	Process knowledge and skills 4 components, 3 references	Managing project lifecycle (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020) Being knowledgeable about success factors that depend on data-driven decision making (DDDM) (Ashaari et al., 2021) Being knowledgeable about the role of business analytics processes in supporting DDDM (Ashaari et al., 2021) Understanding ADA process policies and plans (Ashaari et al., 2021)
	Technology management knowledge and skills 4 components, 3 references	Understanding technological trends (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020) Being aware of the critical success factors for the organization (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020) Being knowledgeable about the role of business analytics (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020) Being capable of learning new technologies (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020)
	Relation/ Communication knowledge and skills 3 components, 4 references	Being efficient of executing work in a collective environment (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020) Being capable of educating others (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020) Communicating effectively with stakeholders (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020; ZareRavasan, 2023)
	Data knowledge and skills 3 components, 7 references	Manipulating/ accessing data (Kim et al., 2018; Moseneke & Brown, 2022; Thajudeen, 2018; Vidgen, 2014; Vidgen et al., 2017; H. J. Watson, 2014; Zamora & Carlos Barahona, 2016) Having data visualization skills (H. J. Watson, 2014) Querying and manipulating structured data (Kim et al., 2018)
	Reasoning/Problem solving knowledge and skills (Moseneke & Brown, 2022) 0 components, 1 reference	
	Product development knowledge and skills (Kim et al., 2018) 0 components, 1 reference	
Team structure 4 capabilities, 5 references	Data scientist 2 components, 1 reference	Specialists (Kim et al., 2018) Manager (Kim et al., 2018)
	Analyst 2 components, 4 references	Business Analyst (Popovič et al., 2018; H. J. Watson, 2014; Zamora & Carlos Barahona, 2016) Business Intelligence Analyst (H. J. Watson, 2014)
	IT specialist (Popovič et al., 2018; Vidgen, 2014; Vidgen et al., 2017; H. J. Watson, 2014) 0 components, 4 references	
	End-users / Customers (H. J. Watson, 2014) 0 components, 1 reference	
Personality Traits 1 capability, 4 references	Personality traits 6 components, 4 references	Being curious (Vidgen, 2014; Vidgen et al., 2017; H. J. Watson, 2014; Zamora & Carlos Barahona, 2016) Being innovative/creative (Vidgen, 2014; Vidgen et al., 2017) Practical intelligence (Vidgen, 2014; Vidgen et al., 2017) Analytical thinking (Vidgen, 2014; Vidgen et al., 2017) Working independently (Vidgen, 2014; Vidgen et al., 2017) Having communication and empathy skills (Vidgen, 2014; Vidgen et al., 2017)

Table 5: Hierarchical list of people capabilities for ADA

4.2 Culture capabilities and components for ADA

For culture, we identified seven high-level capabilities. These are collectively addressed in 17 capabilities and 69 components, as depicted in Table 6

High-level Capabilities	Capabilities	Components
Human Strategy/ Skills and Talent management 4 capabilities, 21 references	Skilling the workforce (Gökalp et al., 2021; G. J. Watson et al., 2021) 16 components, 12 references	Providing coaching/incentives to enable DDDM (Gupta & George, 2016; Yu et al., 2021; Zamora & Carlos Barahona, 2016) Training by conducting case studies (Kim et al., 2018; Mandal, 2018) Using formal coursework (Kim et al., 2018; H. J. Watson, 2014) Evaluating the performance of candidates and employees against the defined performance criteria to meet organizational needs (Gökalp et al., 2021) Managers receiving the training first and training the employees soon after (Korherr & Kanbach, 2021) Conducting programming workshops (Mandal, 2018) Training by conducting demo sessions (Mandal, 2018) Using knowledge repository (Kim et al., 2018) Mentoring (Kim et al., 2018) Providing lessons- learned reports (Daradkeh, 2019) Providing off-site events and educational sessions (Storm & Borgman, 2020) Using educational platforms (Moseneke & Brown, 2022) Training on technical aspects like big data analytics (Gupta & George, 2016) Exploitation of existing competencies and new knowledge (Gupta & George, 2016) Making sure employees can make use of the new technologies being introduced (Moseneke & Brown, 2022) Developing decision makers skill set (ZareRavasan, 2023)
	Promoting continuous learning (Thajudeen, 2018; H. J. Watson, 2014) 5 components, 12 references	Engagement of participants in learning (Medeiros et al., 2020; Mikalef & Krogstie, 2020; Popovič et al., 2018; Upadhyay & Kumar, 2020) Creating collaborative settings for collective learning (Akhtar et al., 2019; Gökalp et al., 2021; Mandal, 2018; H. J. Watson, 2014) Helping employees easily investigate new and related knowledge (AL-Ma'aitah, 2020; Gupta & George, 2016) Helping employees attain new and applicable knowledge related to their jobs (AL-Ma'aitah, 2020; Gupta & George, 2016) Helping employees understand new and related knowledge attained from big data (AL-Ma'aitah, 2020; Gupta & George, 2016)
	Employee engagement 3 components, 4 references	Ensuring employee motivation and commitment (Bamel & Bamel, 2020; Upadhyay & Kumar, 2020; H. J. Watson, 2014) Providing constructive feedback/criticism to employees on job performance (Upadhyay & Kumar, 2020) Promoting employee empowerment (Popovič et al., 2018)
	Incentivization 1 component, 1 reference	Providing incentive mechanism for promoting fast learning and application of new technologies in the organization (Mandal, 2018)
Collaboration and Teamwork 4 capabilities,18 references	Cross-functional team building (Akhtar et al., 2019; Gökalp et al., 2021; Korherr & Kanbach, 2021; Vidgen, 2014; Yeoh et al., 2015) 3 components, 11 references	Acquisition of right people (AL-Ma'aitah, 2020; Daradkeh, 2019; Mesaros et al., 2016; Thajudeen, 2018; Vidgen et al., 2017) Partnering with data science experts and researchers (Vidgen, 2014; Vidgen et al., 2017) Recruiting and retaining highly skilled talent (H. J. Watson, 2014)
	Team collaboration and coordination (Arsal et al., 2022; Martinez et al., 2021) 2 components, 6 references	Facilitating horizontal collaboration (Akhtar et al., 2019; Thajudeen, 2018; Zamora & Carlos Barahona, 2016) Managing conflicts in teamwork and supporting collaboration in teams (Upadhyay & Kumar, 2020)
	Promotion of data sharing practices within the organization (Thajudeen, 2018; H. J. Watson, 2014; Zamora & Carlos Barahona, 2016) 2 components, 4 references	Building trust among stakeholders for data sharing (Thajudeen, 2018) Being able to explore and acquire new and relevant data related to projects (Upadhyay & Kumar, 2020)

	Effective convey of data-driven insights to stakeholders for decision-making (Kim et al., 2018; Martinez et al., 2021; Moseneke & Brown, 2022) 0 components, 3 references	
Top Management support /Commitment 3 capabilities, 16 references	Commitment to ADA culture and use 5 components, 8 references	Perceiving data as substantial asset (AL-Ma'aitah, 2020; Ashaari et al., 2021; Korherr & Kanbach, 2021; Upadhyay & Kumar, 2020; Yu et al., 2021) Using big data and predictive analytics for decision making (AL-Ma'aitah, 2020; Ashaari et al., 2021; Mesaros et al., 2016; ZareRavasan, 2023) Promoting a decision-making culture that combines analytical insights with human intuition (ZareRavasan, 2023) Supporting real-time decision making (Popovič et al., 2018) Using the support of BI tools for decision making (Korherr & Kanbach, 2021)
	Sponsorship and portfolio management (Bamel & Bamel, 2020; Daradkeh, 2019; Mesaros et al., 2016; Popovič et al., 2018; Vidgen et al., 2017; H. J. Watson, 2014) 4 components, 9 references	Time investment on ADA projects (Gupta & George, 2016; Moseneke & Brown, 2022) Investing in ADA in alignment with the organizational vision (Gökalp et al., 2021) Considering and evaluating different sponsorship and funding models and options (Gökalp et al., 2021) Monitoring, optimizing and evaluating projects in ongoing portfolios (Gökalp et al., 2021)
	Keeping employees informed on ADA 2 components, 5 references	Communication on benefits of ADA (AL-Ma'aitah, 2020; Daradkeh, 2019; Kim et al., 2018; H. J. Watson, 2014) Creating awareness in regard to vision, mission and goals on ADA (Mandal, 2018)
Organizational Change / Change Management 3 capabilities, 13 references	Leading change in organization (Medeiros et al., 2020; Thajudeen, 2018; Vidgen, 2014; Vidgen et al., 2017; Zamora & Carlos Barahona, 2016) 8 components, 9 references	Promoting change towards data-driven culture with clear communication and being an exemplary role (Korherr & Kanbach, 2021; Storm & Borgman, 2020) Adoption of effective change management towards a dynamic business management (Daradkeh, 2019) Understanding the scope and desire to changes (Gökalp et al., 2021) Assessing stakeholders' and employees' readiness and willingness for changes (Gökalp et al., 2021) Identifying and deploying action plans to motivate stakeholders and employees towards changes (Gökalp et al., 2021) Monitoring and sustaining organizational changes (Gökalp et al., 2021) Showing concrete results and relevance of DDDM for transformation (Storm & Borgman, 2020) Redefining organizational structure, leadership, and business processes for creating data-driven culture across the organization (Gökalp et al., 2021)
	Openness for experimentation and innovative initiatives (Korherr & Kanbach, 2021) 5 components, 6 references	Being open to override one's own intuition when data contradict one's viewpoints (Gupta & George, 2016; Yu et al., 2021) Creating an environment of curiosity (Thajudeen, 2018; Zamora & Carlos Barahona, 2016) Having an open mindset and the willingness to use unconventional approaches if required (Korherr & Kanbach, 2021) Being willing to break up existing hierarchies and structures to achieve a common goal (Korherr & Kanbach, 2021) Being open to new ideas and approaches that challenge current practices on the basis of new information (Ashaari et al., 2021)

	Leading People Towards Change 4 components, 2 references	Effective listening and persuasion of employees towards change (G. J. Watson et al., 2021) Being able to empathize with employees to understand their fears, contexts, and aspirations against change (G. J. Watson et al., 2021) Supporting workers through individualized training programs for role changes (G. J. Watson et al., 2021) Increasing stakeholders' participation in organizational changes (Gökalp et al., 2021)
Business Understanding 2 capabilities, 17 references	Having a clear business vision and need (Daradkeh, 2019; Kim et al., 2018; Korherr & Kanbach, 2021; Mikalef et al., 2019; H. J. Watson, 2014; Yeoh et al., 2015) 4 components, 11 references	Describing stakeholders' needs (Martinez et al., 2021; Storm & Borgman, 2020) Understanding the customers and user segments (Mesaros et al., 2016; H. J. Watson, 2014) Determining the project success metrics in the early phase (Kim et al., 2018; Mesaros et al., 2016) Choosing the right business goals, needs, and problems (Kim et al., 2018)
	Alignment of data strategy with business strategy (Abeza et al., 2023; Daradkeh, 2019; Kim et al., 2018; Korherr & Kanbach, 2021; Mesaros et al., 2016; Mikalef et al., 2019; Moseneke & Brown, 2022; Popovič et al., 2018; Vidgen et al., 2017; H. J. Watson, 2014; Yeoh et al., 2015; Zamora & Carlos Barahona, 2016) 3 components, 15 references	Utilizing data-based insights for the creation of new strategies to assess and improve business performance (Ashaari et al., 2021; Gupta & George, 2016) Determining the target data science capabilities in line with organizational vision (Gökalp et al., 2021) Establishing and maintaining a strategic plan and roadmap to drive alignment among business, data science, and IT units (Gökalp et al., 2021)
Ethics/Values 1 capability, 2 references	Ethics/Values 2 components, 1 reference	Awareness on ethical issues (G. J. Watson et al., 2021) Keeping up with changing social values and potential ethical pitfalls of new technologies (G. J. Watson et al., 2021)

Table 6: Hierarchical list of culture capabilities for ADA

5 Discussions

This study maps the people and culture related capabilities essential for ADA success into a hierarchical list which structure serves as a clear roadmap for building workforce competency and cultural adaptability. Embedding these capabilities into practice fosters a resilient, adaptable culture that sustains effective data analytics initiatives.

While technical skills remain fundamental, their effectiveness is significantly influenced by managerial, leadership, and communication capabilities (Mandal, 2018; G. J. Watson et al., 2021; Zamora & Carlos Barahona, 2016). Organizations should adopt an integrated approach to talent development, ensuring employees possess both technical expertise and the ability to collaborate and communicate effectively (Kim et al., 2018; G. J. Watson et al., 2021). Technical skills like programming (Ashaari et al., 2021; Kim et al., 2018; Mandal, 2018; Upadhyay & Kumar, 2020; Vidgen, 2014; Vidgen et al., 2017; H. J. Watson, 2014) and statistical analysis (Ashaari et al., 2021; Upadhyay & Kumar, 2020; H. J. Watson, 2014) are essential, but must be complemented by managerial knowledge such as effective project management (Ashaari et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020) and communication skills (G. J. Watson et al., 2021; Zamora & Carlos Barahona, 2016). Leadership qualities like a growth mindset and agility in decision-making (G. J. Watson et al., 2021) are also crucial for fostering innovation and adaptability in ADA projects.

Another important insight is the role of continuous learning in fostering a data-driven culture (G. J. Watson et al., 2021). Continuous learning and innovation must be integrated into ADA workflows to foster adaptability and efficiency (G. J. Watson et al., 2021). Structured frameworks for corrective actions, workflow optimization, and retrospective reviews refine processes and drive

sustained improvement. Automating repetitive tasks, leveraging team dashboards for shared data resources, and providing guidelines for asking clear, goal-oriented questions streamline decision-making and enhance operational efficiency (Akhtar et al., 2019; Gökalp et al., 2021; Mandal, 2018; Upadhyay & Kumar, 2020). The findings highlight that organizations which prioritize training, mentoring, and knowledge-sharing initiatives are better equipped to navigate the complexities of ADA implementation (Kim et al., 2018; Mandal, 2018; G. J. Watson et al., 2021). The structured processes identified in this study emphasize the need for formalized training programs, collaborative learning environments, and incentives for knowledge acquisition (Gupta & George, 2016; Zamora & Carlos Barahona, 2016). This aligns with previous research that suggests competency-based models enhance workforce adaptability and innovation (Moseneke & Brown, 2022).

Moreover, leadership's commitment to ADA adoption emerges as a crucial element of success. The study underscores that top management support is not merely about financial investment but also about fostering a culture where data-driven decision-making is valued (Korherr & Kanbach, 2021; Storm & Borgman, 2020). Leaders play a pivotal role in promoting an environment where data-driven insights are integrated into strategic decision-making, thereby reducing resistance to change and enhancing organizational agility (Gökalp et al., 2021; Storm & Borgman, 2020).

Furthermore, the findings suggest that collaboration and coordination across different teams remain a challenge in ADA projects. Organizations often struggle with aligning technical and business perspectives, leading to inefficiencies in data analytics adoption (Gökalp et al., 2021). Addressing this issue requires dedicated efforts to bridge the communication gaps between data scientists, business analysts, and executives (Gökalp et al., 2021). By fostering cross-functional collaboration and providing structured frameworks for communication, organizations can improve the effectiveness of their data-driven initiatives. Additionally, organizational change management plays a critical role in sustaining ADA initiatives. The study emphasizes that organizations that actively engage employees in the transition to data-driven decision-making experience higher levels of adoption and integration (Daradkeh, 2019). Resistance to change is an important barrier in analytics-driven transformations, and structured approaches such as training programs, leadership engagement, and change agents can mitigate these challenges (Gökalp et al., 2021).

Finally, ethical governance, including transparency, privacy, and accountability, must be emphasized in theoretical models to bridge organizational practices with societal expectations (G. J. Watson et al., 2021). The integration of sustainability goals, such as environmental responsibility in machine learning workflows, can expand the scope of ethical frameworks in ADA contexts. Models should also consider the dynamic nature of data strategy alignment, adapting to evolving priorities and resource constraints.

6 Conclusions

Organizations need to leverage data analytics capabilities to gain competitive advantages. These capabilities encompass both technical and non-technical areas. Non-technical capabilities, namely those related to people and culture, have been understudied in extant literature. This study aims to address this gap by conducting an SLR to synthesize people and culture capabilities for ADA in a way that can guide organizations in maturing these capabilities. By analyzing 29 papers published between 2010 and 2024, the study synthesized a hierarchical list, where capabilities are organized in two levels and each capability is represented in a set of components, which provides diverse

ways of achieving the identified capabilities. It captures knowledge not only from capability literature but also from success factor and maturity model literature. In this way, it provides a more detailed and structured view of capabilities in non-technical areas, unlike previous SLRs (e.g., Mikalef et al., 2018; Mikalef & Krogstie, 2020).

A key limitation of this study is the lack of empirical validation, as the framework is based solely on literature synthesis; future research could enhance our understanding of these capabilities by conducting empirical studies, such as interviews and case studies in organizations, particularly across different domains and organizational characteristics (e.g., size, sector). Future assessments could also focus on the implementation of these capabilities, evaluating processes that actualize them within varied organizational settings.

Acknowledgement

This study was supported by funding from the Research Universities Support Program (ADEP), provided through Middle East Technica University (METU) under project ADEP-704-2024-11481 by the Council of Higher Education (YÖK) in Türkiye.

References

Abeza, G., O'Reilly, N., Nadeau, J., & Abdourazakou, Y. (2023). Big data in professional sport: the perspective of practitioners in the NFL, MLB, NBA, and NHL. *Journal of Strategic Marketing*, *31*(8), 1413–1433.

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.

AL-Ma'aitah, M. A. (2020). Utilizing of big data and predictive analytics capability in crisis management. *Journal of Computer Science*, *16*(3), 295–304.

Al-Sai, Z. A., Abdullah, R., & Husin, M. H. (2019). Big Data Impacts and Challenges: A Review. *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology, JEEIT 2019* 150–155.

Al-Sai, Z. A., Abdullah, R., & Husin, M. H. (2020). Critical Success Factors for Big Data: A Systematic Literature Review. *IEEE Access*, *8*, 118940–118956.

Arsal, R. E., Durdu, H. T., & Tongarlak, M. H. (2022). Organizational Enablers of Data-Driven Digital Transformation: A Case Study from Banking Industry. *2022 IEEE Technology and Engineering Management Conference (TEMSCON EUROPE)*, 210–216.

Arunachalam, D., Kumar, N., & Kawalek, J. P. (2018). Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. *Transportation Research Part E: Logistics and Transportation Review*, *114*, 416–436.

Ashaari, M. A., Singh, K. S. D., Abbasi, G. A., Amran, A., & Liebana-Cabanillas, F. J. (2021). Big data analytics capability for improved performance of higher education institutions in the Era of IR 4.0: A multi-analytical SEM & ANN perspective. *Technological Forecasting and Social Change*, *173*(August), 121119.

Bamel, N., & Bamel, U. (2020). Big data analytics based enablers of supply chain capabilities and firm competitiveness: a fuzzy-TISM approach. *Journal of Enterprise Information Management*, *34*(1), 559–577.

Bose, R. (2009). Advanced analytics: opportunities and challenges. *Industrial Management & Data Systems*, *109*(2), 155–172.

Brinch, M., Gunasekaran, A., & Fosso Wamba, S. (2020). Firm-level capabilities towards big data value creation. *Journal of Business Research*, vol. 131, 539-548.

Daradkeh, M. (2019). Critical success factors of enterprise data analytics and visualization ecosystem: An interview study. *International Journal of Information Technology Project Management (IJITPM)*, *10*(3), 34–55.

Demir, N., Aysolmaz, B., & Özcan-Top, Ö. (2024). Critical Success Factors in Data Analytics Projects: Insights from a Systematic Literature Review. *Conference on E-Business, e-Services and e-Society*, 129–141.

Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2021). Artificial Intelligence and Business Value: a Literature Review. *Information Systems Frontiers*, *24*(August), 1709–1734. -w

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.

Gökalp, M. O., Kayabay, K., Gökalp, E., Koçyiğit, A., & Eren, P. E. (2021). Assessment of process capabilities in transition to a data-driven organisation: A multidisciplinary approach. *IET Software*, *15*(6), 376–390.

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information Systems*, *26*(3), 191–209.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information and Management*, *53*(8), 1049–1064.

Hariri, R. H., Fredericks, E. M., & Bowers, K. M. (2019). Uncertainty in big data analytics: survey, opportunities, and challenges. *Journal of Big Data*, *6*(1), 1–16.

Hattingh, M., Marshall, L., Holmner, M., & Naidoo, R. (2019). Data science competency in organisations: A systematic review and unified model. *ACM International Conference Proceeding Series*, 1-8

Kim, M., Zimmermann, T., Deline, R., & Begel, A. (2018). Data scientists in software teams: State of the art and challenges. *IEEE Transactions on Software Engineering*, *44*(11), 1024–1038.

Kitchenham, B., & Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*,October 2021.

Korherr, P., & Kanbach, D. (2021). Human-related capabilities in big data analytics: a taxonomy of human factors with impact on firm performance. *Review of Managerial Science*, *17*(6), 1943–1970.

Korsten, G., Aysolmaz, B., Ozkan, B., & Turetken, O. (2024). A Capability Maturity Model for Developing and Improving Advanced Data Analytics Capabilities. *Pacific Asia Journal of the Association for Information Systems*, *16*(4).

Korsten, G., Aysolmaz, B., Turetken, O., Edel, D., & Ozkan, B. (2022). ADA-CMM: A Capability Maturity Model for Advanced Data Analytics. In *55th Annual Hawaii International Conference on System Sciences, HICSS 2022,* 266-275.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.

Mandal, S. (2018). An examination of the importance of big data analytics in supply chain agility development: A dynamic capability perspective. *Management Research Review*, *41*(10), 1201–1219.

Martinez, I., Viles, E., & Olaizola, I. G. (2021). A survey study of success factors in data science projects. *2021 IEEE International Conference on Big Data (Big Data)*, 2313–2318.

Medeiros, M. M. de, Hoppen, N., & Maçada, A. C. G. (2020). Data science for business: benefits, challenges and opportunities. *Bottom Line*, *33*(2), 149–163.

Mesaros, P., Carnicky, S., Mandicak, T., Habinakova, M., Mackova, D., & Spisakova, M. (2016). Model of key success factors for Business Intelligence implementation. *Journal of Systems Integration*, 3–15.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big Data Analytics Capabilities and Innovation: The Mediating Role of Dynamic Capabilities and Moderating Effect of the Environment. *British Journal of Management*, *30*(2).

Mikalef, P., & Krogstie, J. (2020). Examining the interplay between big data analytics and contextual factors in driving process innovation capabilities. *European Journal of Information Systems*, *29*(3), 260–287.

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and E-Business Management*, *16*(3), 547–578.

Moseneke, F., & Brown, I. (2022). Big data analytics capabilities and value creation at the work-practice level: A South African case study. *EPiC Series in Computing*, *85*, 196–213.

Popovič, A., Hackney, R., Tassabehji, R., & Castelli, M. (2018). The impact of big data analytics on firms' high value business performance. *Information Systems Frontiers*, *20*(2), 209–222.

Ratna, D., Khodijah, A., Govindaraju, R., & Prihartono, B. (2023). The role of knowledge-oriented leadership in fostering innovation capabilities: The mediating role of data analytics maturity. *IEEE Access*, 129683 -129702.

Saltz, J. S., & Shamshurin, I. (2016). Big data team process methodologies: A literature review and the identification of key factors for a project's success. *Big Data*, 2872–2879.

Storm, M., & Borgman, H. P. (2020). Understanding challenges and success factors in creating a data-driven culture. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 5399–5408.

Surbakti, F. P. S., Wang, W., Indulska, M., & Sadiq, S. (2020). Factors influencing effective use of big data: A research framework. *Information and Management*, *57*(1), 103146.

Thajudeen, S. S. (2018). Advanced analytics solutions: Towards a data driven organisation. *Offshore Technology Conference Asia*, OTC-28591-MS

Upadhyay, P., & Kumar, A. (2020). The intermediating role of organizational culture and internal analytical knowledge between the capability of big data analytics and a firm' s performance. *International Journal of Information Management*, *52* (November 2019), 102100.

Vidgen, R. (2014). Creating business value from Big Data and business analytics: organizational, managerial and human resource implications. In *Hull University Business School*, vol. 94..

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, *261*(2), 626–639.

Voigt, T., et al. (2021). Advanced Data Analytics Platform for Manufacturing Companies. 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), 01-08.

Watson, G. J., Desouza, K. C., Ribiere, V. M., & Lindič, J. (2021). Will AI ever sit at the C-suite table? The future of senior leadership. *Business Horizons*, *64*(4), 465–474.

Watson, H. J. (2014). Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems*, *34*(1), 65.

Yeoh, W., Koronios, A., Yeoh, W., & Koronios, A. (2015). Critical Success Factors for Business Intelligence Systems. *Journal of Computer Information Systems*, *4417*(2010).

Yu, W., Wong, C. Y., Chavez, R., & Jacobs, M. A. (2021). Integrating big data analytics into supply chain finance: The roles of information processing and data-driven culture. *International Journal of Production Economics*, *236*(April), 108135.

Zamora, D., & Carlos Barahona, J. (2016). Data-driven innovation at the Grupo Pellas SER company. *Emerald Emerging Markets Case Studies*, *6*(2), 1–16.

ZareRavasan, A. (2023). Boosting innovation performance through big data analytics: An empirical investigation on the role of firm agility. *Journal of Information Science*, *49*(5), 1293–1308.

10. Mitigating the Risk of Fake User-Generated Content

Christopher Gusenbauer
Johannes Kepler University, Linz Austria
mail@christophergusenbauer.at

Michaela K. Trierweiler
Johannes Kepler University, Linz Austria
michaela.trierweiler@jku.at

Abstract
User-generated content (UGC) dominates the internet. Tech-companies and social-media platforms are used to moderate UGC and drive their business models with it. However, startups and traditional firms might lack knowledge in that area. Based on a literature review, we identified 74 risks and 111 corresponding safeguards and concatenated them into a 'fake information risk mitigation' catalog. A potential overreliance on detection and underutilization of preventative and recovery measures were discovered. A high reliance on algorithmic detection became apparent, while education was underrepresented. Thus, this study argues for an increased focus on education and training. The large amount and variety of discovered risks and measures suggests the need for an interdisciplinary, collaborative approach of companies and authorities to fight fake UGC, as previous studies discussed smaller sets of risks and measures.

Keywords
User-generated Content, Fake Content, Fake News, Risk Mitigation, Disinformation, Platforms.

1. Introduction
Altogether, Alphabet, Meta, Netflix, Microsoft, and TikTok generate more than half of all internet traffic (Sandvine, 2024). These enterprises are 'Big-Tech' players of the so-called platform- or internet-economy (e.g., Erkayhan, 2020; Obermaier & Mosch, 2019). These kind of business models depend on qualitative, engaging, and accurate content. Further characteristics are community and collaboration (Bodrožić & S. Adler, 2022), fueled by network-effects (e.g., Shapiro & Varian, 1998), and the intermediary function of the platform owner (e.g., Kirchner & Schüßler, 2019). Web content is increasingly dominated by user generated content (UGC) (Bond, 2015). However, only a small percentage of active users are the creators of half of UGC (Baeza-Yates & Saez-Trumper, 2015). Nowadays, traditional enterprises, but also start-ups must integrate increasingly more UGC into their systems and offerings - often for the first time.

Since OpenAI (2022) launched ChatGPT in November 2022, generative artificial intelligence (AI) is rapidly improving and had become an easy source for creating fake content (onlinesicherheit.at, 2024). This increases the risk for companies of distributing false information produced by their platform-users. Disinformation and fake content are spread to influence ordinary users, manipulate elections, create disruption, or are even used in cyberwar activities (Shu–Min Hou et al., 2023). As fake UGC spreads faster than truth (Vosoughi et al., 2018), it is quite difficult to identify and to correct it (Gupta, Kumar, et al., 2022). Big-Techs have developed guidelines and sophisticated algorithms to combat the risk of fake UGC. However, holistic strategies and frameworks for this issue which are suitable for smaller organizations seem not to exist. While some research (e.g., Farhoudinia et al., 2023; Kertysova, 2018; Zhang & Ghorbani, 2020) covered individual types of

UGC, they have not been structured around risks and mitigating measures. Research was mostly focused on providing a state-of-the-art view of the research field but has not been addressed to practitioners. This study aims to close this gap by providing a fake information risk mitigation (FIRM) catalogue: We analyzed different types of fake UGC, such as fake news (e.g., Zhang & Ghorbani, 2020), disinformation (e.g., Rubin, 2019; Shu et al., 2020) and fake reviews (e.g., Agnihotri et al., 2021) and identified a grand set of 111 safeguards to them proposed in literature. By this, we aim to extent the existing scientific knowledge base and further, give practical guidelines. We think managers and personnel responsible for overseeing UGC offerings need a compendium of existing risks and countermeasures when implementing and continuously facilitating UGC. We aim to support regional platform providers, which are often vital to their specific ecosystem, such as local newspapers with comment sections, forum websites, interactive government websites, and regional event websites. Because they often lack the scale, data, and expertise needed to effectively address the issue of fake UGC.

A mixed methods approach, consisting of a systematic literature review, qualitative content analysis, and taxonomy development, was conducted. We analyzed the importance of risks and measures (via a ranking) and assigned corresponding measures to these risks they mitigate. During this mapping-process, different risks and measures were grouped and we built a parent abstraction level for risks and measures acting as a taxonomy. Further, areas of cooperation between individual facilitators, society, and governments were identified and discussed. The remainder gives an introduction of different types of UGC, their common characteristics and vulnerabilities. It explains the research process including the selection of literature samples to build the FIRM catalog, which encompasses risks and measures, on a granular level, but also as groupings. After the discussion, this paper closes with a summary and future research paths.

2. (Fake) User Generated Content

To identify risks and measures in the context of fake UGC, it is necessary to define the term ‘user-generated content’. Santos (2021) systematically reviewed different definitions of UGC and synthetized it as “*any kind of text, data or action performed by online digital systems users, published and disseminated by the same user through independent channels”* (p. 108). This generic definition includes different kinds of user-generated content, such as reviews, ratings, blog posts, social media feeds, but also video, graphical or audio – and their fake siblings. Santos (2021) describes common vulnerabilities of UGC: the data itself, actions by users, digital systems, the users themselves, publishing/dissemination channels, and emergent effects through combined actions of (potentially malicious or tricked) users. For example, an individual user that upvotes soon after a content was posted has no effect, but multiple users quickly upvoting content can cause it to trend. However, other studies may define the term UGC more narrowly. For example, Naab and Sehl require content to have a “rationale” and, therefore, excluded “Applications that simply allow the expression of a judgment without any opportunity for reasoning” (2016, p.1258), i.e., a simple like-button.

The terms fake news, disinformation, and misinformation often overlap, are used interchangeably, or have unclear boundaries. According to Fahroudinia et al. (2023), misinformation stands for “false, inaccurate, or incomplete information” (p.7), while disinformation means “incorrect information shared intentionally” (p.7). In contrast, Rubin (2019) makes the separation more distinct, calling it unintentional in nature. Fake news can be defined as “intentionally false, fabricated news articles that can deceive the reader” (Farhoudinia et al., 2023, p. 7), while other

authors include low-quality or incomplete news (Ozbay & Alatas, 2020) or false stories (Zhang & Ghorbani, 2020) in their definitions. Reviews are fake if "the review was written for the purpose of promoting or downgrading a product, service or company" (Scherr et al., 2019, p. 454). As for, so many different types of fakes exist, we decided for a broader definition of UGC. We used the terms disinformation, fake news, misinformation, and false news interchangeably. And as a conclusion, this research developed the following definition for fake UGC: "*Fake user-generated content is user-generated content that was (a) created and/or shared with/without malicious intent, and (b) contains incorrect, distorted, half-true, manipulative or manipulated content to such an extent, that it could not be considered accidental, because it influences the entirety of its message (or its key points)*". The part (b) of this definition is to ensure that minor errors in, for example, a quality blog article, does not immediately classify it as fake UGC. By this new definition, we aim to compensate some limitations of previous research having a narrower scope.

3. Methodological Approach

For research rigor, we applied a mixed method approach in three stages: As the FIRM catalogue would be literature based, we decided to follow a double strategy starting with a scoping review (Grant & Booth, 2009), followed by applying steps of the systematic literature review (SLR) approach from Kitchenham et al. (2007). Due to the novelty of our research topic, we saw the need to include some grey literature (Garousi et al., 2019), as well. Further, we oriented on qualitative content analysis (QCA) techniques from Mayring (2000) to assess the found literature samples, and applied taxonomy-building strategies of Nickerson et al. (2013) to build the meta-groupings of UGC risks and their mitigation measures.

In concrete, in stage one, we applied different sets of search-terms in SCOPUS and Google-Scholar, because both act as meta-search engines across multiple disciplines and were assumed to provide an extensive sample baseline. After, in total, 97 search queries, some repetition of results applied, which suggested a kind of theoretical saturation. The most promising results came from the following search string: [(measures OR strategies OR actions) AND detection AND fake AND news]. With respect to the grey literature, we used strings like [How does facebook/google detect fake content?], but we excluded sources from smaller companies or so-called individual experts. In total, 219 sources of English and German speaking authors from the years 2015 to 2024 were assessed by including/excluding criteria, e.g., citation-rate, methodology used, credibility of authors respectively publishing institution, and relevance by content. Five main pillars were defined for relevancy with the research question: (i) mostly about fake UGC, (ii) includes textual and/or reaction-based UGC, (iii) only includes textual and/or reaction-based UGC, (iv) includes risks of UGC, and (v) includes protective measures against fake UGC. Each source was rated along these pillars and included in the sample when either (a) its content matched at least four of five pillars, (b) was a Big-Tech company resource, or (c) was classified as 'meta-paper', i.e., a tertiary study. We excluded all sources providing pure technical solutions for increasing detection scores of existing safeguards. By this process we narrowed down to the final sample of 48 sources containing multiple viewpoints from different domains, such as security, AI, human rights, information management, data mining, human-computer interaction, social network analysis, journalism, and administrative sciences.

In the second stage, this literature sample was evaluated by applying QCA techniques. Firstly, we color-coded in each source any mentioning or description of a specific fake UGC risk and their countermeasure. Secondly, we transferred these findings into an Excel spreadsheet for

paraphrasing and abstraction to build a general category structure. For this collection process, we asked the questions of ‘What risks were described?’ and ‘What preventative measures were mentioned?’. Every new risk or measure, that was not a clear duplicate, was added to the table. Otherwise, the source was associated with the already identified risk or measure category. By assessing the sources in detail one-by-one, we were able to apply an inductive category development, as described by Mayring (2000). Further, this incremental procedure resulted in a matrix-structure, that allowed us to map each mentioned risk to its mitigation strategy accordingly. As a result, 74 risks associated to fake UGC were identified and 111 countermeasures and safeguards for mitigation of these risks.

In the third stage, we created parent groupings of risk and mitigation strategies by using a taxonomy building method (Nickerson et al., 2013) supported by a mind mapping strategy (Eppler, 2006). We used mind-map visualizations, because this allowed us to see associations much quicker, spot doublets, or wrongly placed elements easier. For example, the risk of ‘influences politics and elections” and the risk of ‘cyberstalking, online harassment, and cyberbullying’ were subsumed under the risk category of ‘Societal Effects & Issues’. Building a kind of taxonomy allows us to show our results on an abstract level with the advantage for users of the FIRM catalogue to look for specific areas of risks related to fake UGC. This process resulted in 16 groups of risks connected to fake UGC and 15 associated mitigation strategies.

4. The Fake Information Risk Mitigation (FIRM) Catalog

Aim of this research was to answer the question of which risks concerning fake UGC might companies face when facilitating user generated content on their websites and apps and which mitigation-strategies can be applied. After the analysis and taxonomy-building process, 74 risks and 111 mitigation measures, were identified and compiled into the FIRM catalog. These items were grouped into 16 risks and 15 mitigation strategies to fight these risks. Our analysis showed that different types of mitigation strategies could be applied to different types of fake UGC risks, building an n:m relationship. We think a single chart would not be handy and probably confusing to a reader. Therefore, we present the FIRM catalog in two parts: firstly, the identified risks related to fake UGC and secondly, the mitigation measures. This split intends for a quick and easy use, also by practitioners. Due to the original length of the tables and the limited space in this paper, we decided to mention the sources just one-time per category, even if an author might have contributed to different risk-mentions or multiple mitigation strategies.

4.1. Risks Connected to Fake UGC

This chapter presents charts and rankings concerning risks and the risk groupings of fake UGC, as well as the risk-part of the FIRM catalogue in detail. Figure 1 shows the identified 16 types of risks as our parent-grouping, the risk-taxonomy. The four major risk categories consist of societal implications, risks based on human nature, risks based on the nature of fake UGC, and algorithmic quality issues as a detection method.

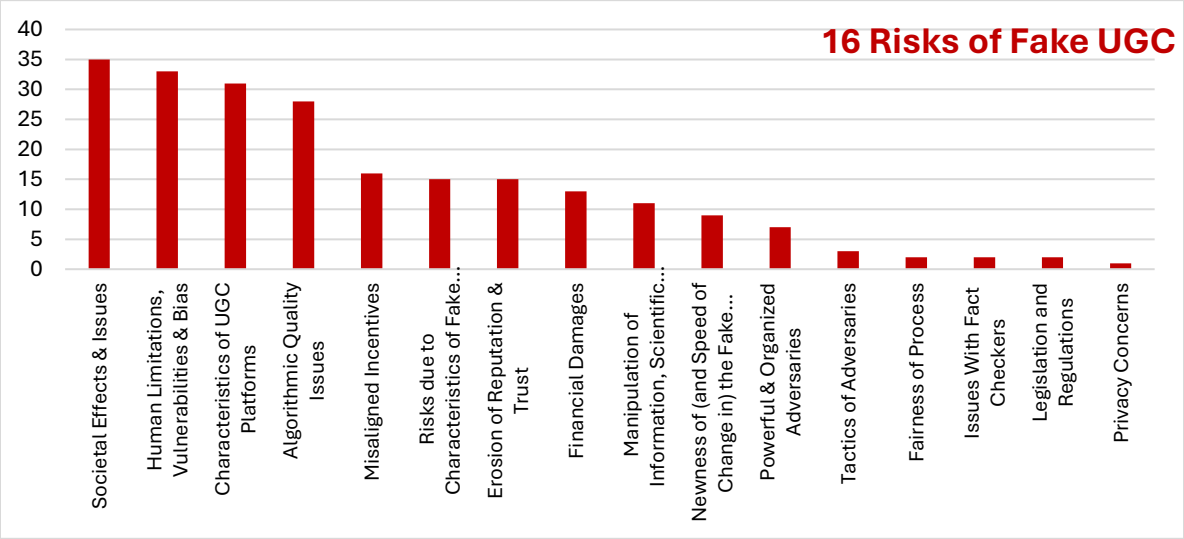

Fig 1: Taxonomy of Risks related to Fake UGC (ranked according to mentions in sources)

To detail these fake UGC risk categories, Table 1 shows the FIRM catalogue along the ranking of Fig. 1 as our risk-taxonomy (red areas), while the subordinated risks are added below as a kind of array, because they do not follow a specific order as we counted them when appearing in a source. The associated mitigation strategies are in blue color writing to the right. For simplicity reasons, we present the Top 10 of risks with their mitigation strategies.

Societal Effects & Issues	**Education**
Influences politics and elections | Influencing the public opinion | Under- or over-reporting of positive or negative opinions | Destabilization (upheaval, incitement, violence, and lack of order) | Polarization and degradation of social dialog | Cyberstalking, online harassment, and cyberbullying | Lack of public awareness | not taking serious and important	
(Brundage et al., 2018; Carter et al., 2021; Chen et al., 2023; Crowell, 2017; Dabbous et al., 2021; Donaker et al., 2019; Einwiller & Kim, 2020; Google, 2019; Gupta, Kumar, et al., 2022; Gupta, Li, et al., 2022; Islam et al., 2020; Lunga & Mthembu, 2019; Mut Camacho & Rueda Lozano, 2021; Papanastasiou, 2018; Pherson et al., 2020; Rubin, 2019; Sharma et al., 2019; Shu–Min Hou et al., 2023; Zhang & Ghorbani, 2020; Zhou & Zafarani, 2020)	
Human Limitations, Vulnerabilities & Bias	**Special Protections for Certain Groups | Supporting Human Fake Detection | Develop Alternative Systems | Platforms or Networks**
Low cognition decreases the detection abilities of individuals as well as their adjustment ability | conservatives might believe fakes more || users lack time and interest in the detection | users lack time and interest in improving their detection skills | humans are liable to a wide range of biases | organizations, news outlets, politicians, and influencers/celebrities accidentally sharing fake UGC | low media and digital literacy | human mental limits for content moderation	
(Brundage et al., 2018; Chen et al., 2023; Dabbous et al., 2021; Farhoudinia et al., 2023; Google, 2019; Gupta, Kumar, et al., 2022; Gupta, Li, et al., 2022; Kertysova, 2018; Lim, 2020; Lunga & Mthembu, 2019; Pherson et al., 2020; Rubin, 2019; Sharma et al., 2019; Shu et al., 2020; Zhang & Ghorbani, 2020; Zhou & Zafarani, 2020)	
Characteristics of User Generated Content Platforms	**Improve Platform Characteristics | Develop Alternative Systems | Platforms or Networks | Regulation and Legislations**
Echo chambers or filter bubbles | micro-targeting | unverified online ads | high speed of sharing and propagation in UGC platforms | viral breakthrough risk | the large number of bots operating | smoke-screen attacks to divert attention away | risk increases during major events or catastrophes | end-to-end encrypted, closed-off nature	
(Carter et al., 2021; Chen et al., 2023; Coutinho & José, 2019; Dabbous et al., 2021; Farhoudinia et al., 2023; Fredheim & Pamment, 2024; Google, 2019; Gupta, Kumar, et al., 2022; Gupta, Li, et al., 2022; Islam et al., 2020; Kertysova, 2018; Lim, 2020; Lunga & Mthembu, 2019; Papanastasiou, 2018; Rubin, 2019; Sharma et al., 2019; Shu et al., 2020; Zhang & Ghorbani, 2020; Zhou & Zafarani, 2020)	
Inherent Risks of Fake User Generated Content	**Improve Platform Characteristics**

Truth being on a spectrum – an objective truth is not always available | fake content spreads faster than truthful content | clickbait | a correct information consumed after the consumption of fake UGC does not offset fake information | more exposure to fake UGC leads to increased trust in that fake UGC | fake content also influences subsequent posters	
(Brundage et al., 2018; Chen et al., 2023; Farhoudinia et al., 2023; Google, 2019; Gupta, Kumar, et al., 2022; Gupta, Li, et al., 2022; Lim, 2020; Lunga & Mthembu, 2019; Papanastasiou, 2018; Pherson et al., 2020; Sharma et al., 2019; Zhang & Ghorbani, 2020; Zhou et al., 2020; Zhou & Zafarani, 2020)	
Algorithmic Quality Issues	**Algorithmic and Artificial Intelligence Supported Measures | Manual Fact-Checking | Peergroup Control**
Bias in algorithms and AI | quality issues with datasets for training and evaluation | Training process attacks (data poisoning/exploratory/black box probing/adversarial inputs) | manipulation of recommendation algorithms | misuse of defense mechanisms, such as crowd flagging | abuse of filtering by governments (e.g., dictatorships) | difficulty of generalizing detection | regional and geographic cultural barriers for algorithms and AI | Linguistic barriers for detection algorithms/AI | lack of explainability of AI/algorithms | Balancing timeliness with detection accuracy | maintenance of knowledge bases and fact-check databases | multilingual capabilities of fake UGC creators | parody-content	
(Brundage et al., 2018; Carter et al., 2021; Chen et al., 2023; Einwiller & Kim, 2020; Farhoudinia et al., 2023; Fredheim & Pamment, 2024; Giełczyk et al., 2019; Google, 2019; Gupta, Kumar, et al., 2022; Kertysova, 2018; Pherson et al., 2020; Sharma et al., 2019; Zhang & Ghorbani, 2020)	
Erosion of Reputation & Trust	**Education, Improve Transparency and Accountability**
…in public authorities | … in journalism and infomediaries, reducing their credibility | … in, the perception of, and the reputation of a company	
(Brundage et al., 2018; Chen et al., 2023; Farhoudinia et al., 2023; Gupta, Kumar, et al., 2022; Islam et al., 2020; Lunga & Mthembu, 2019; Mut Camacho & Rueda Lozano, 2021; Zhou & Zafarani, 2020)	
Financial Damages	**all others indirectly**
stock price and financial losses | decreases the value and revenue of good content through a perceived lack of quality and is lost in the noise of fake content	
(Brundage et al., 2018; Chen et al., 2023; Coutinho & José, 2019; Farhoudinia et al., 2023; Gupta, Kumar, et al., 2022; Gupta, Li, et al., 2022; Islam et al., 2020; Lunga & Mthembu, 2019; Mut Camacho & Rueda Lozano, 2021; Papanastasiou, 2018; Sharma et al., 2019; Zelenka et al., 2021; Zhou & Zafarani, 2020)	
Manipulation of Information, Scientific Facts and Perception	**Education, Supporting Human Fake Detection | Peergroup Control**
altering the perception of scientific and historical facts | public health impacts | pseudo-truth, i.e. omitting, and selectively using facts | pseudo-science	
(Chen et al., 2023; Farhoudinia et al., 2023; Gupta, Kumar, et al., 2022; Gupta, Li, et al., 2022; Islam et al., 2020; Melki et al., 2021; Roth & Pickles, 2020)	
Newness of (and Speed of Change in) the Fake User Generated Content Domain	**Strategic Measures**
No pre-existing experience | field and technology are evolving rapidly | no accepted standard definition of relevant terms	
(Chen et al., 2023; Google, 2019; Gupta, Kumar, et al., 2022; Kertysova, 2018; Lim, 2020; Mut Camacho & Rueda Lozano, 2021; Pherson et al., 2020; Sharma et al., 2019)	
Powerful & Organized Adversaries	**Punish and Limit Bad Actors**
influence-operations all around the globe | use of fake UGC in cyberwar | (powerful) players and stakeholders involved | Organized fake attacks on specific individuals	
(Crowell, 2017; Google, 2019; Pherson et al., 2020; Rubin, 2019; Sharma et al., 2019; Shu–Min Hou et al., 2023)	

Table 1: Top 10 Categories of Risks Connected to Fake UGC

4.2. Fake UGC Mitigation Strategies

With respect to the found fake UGC mitigation measures, we classified them as prevention [P], detection [D], and recovery [R] measures, or a combination. This basic structure follows an approach applied by Girgenti and Hedley (2011) on fraud risk mitigation strategies. Fake UGC can be considered a kind of fraud, because fake UGC is often spread by false or stolen user-identities; and identity theft is a typical fraud type. Figure 2 presents the set of 15 mitigation categories as our taxonomy, showing their rank and proportions of measure types as counted across

all sources. Even after reducing duplets, with overall 48 mentions, detection mitigation strategies were the most recommended by the sources in our sample, followed with 34 mentions of prevention measures. A combination of both, achieved with 13 mentions the third place. Recovery measures received limited attention in our literature sample.

The by far most mentioned detection strategy was the usage of algorithmic detection-methods, however education, or legal regulation was just of minor interest in the assessed literature sample. Therefore, we had a more detailed look into the area of algorithmic and AI supported measures: The biggest mentioned group consisted of neural networks, in particular recurrent neural networks (RNNs) (e.g., Islam et al., 2020; Meta, 2020) and classical machine learning algorithms, like tree-based models, e.g., random forest (e.g., Sharma et al., 2019; Shetty et al., 2023; Zhou et al., 2020). In contrast, linguistic models were only mentioned by three authors (Conroy et al., 2015; Sharma et al., 2019; Zhou & Zafarani, 2020).

Table 2 shares the details for the Top 10 of the identified fake UGC mitigation strategies. The sequence orientates along Figure 2. We present the mitigation measures in blue color and the corresponding risks in red color by using the same manner like for Table 1.

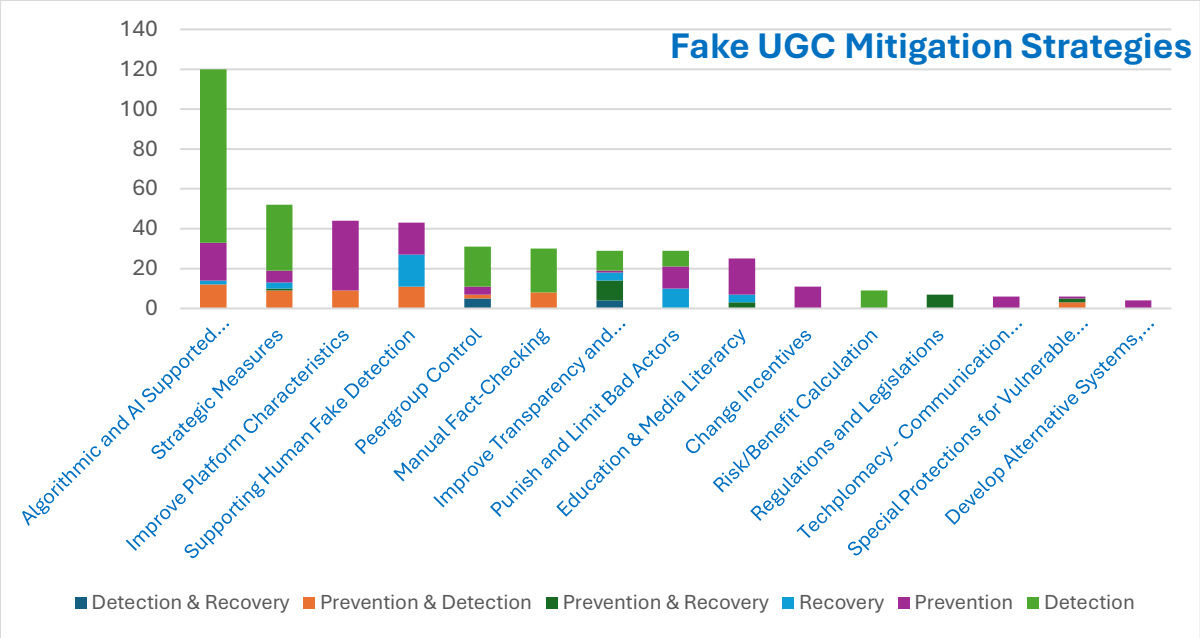

Fig 2: Taxonomy of Risk Mitigation Strategies associated with Fake UGC

5. Discussion

This study analyzed the risks and mitigation measures concerning fake UGC using a mixed-methods approach. The FIRM catalog was developed to provide a comprehensive overview of existing risks and mitigation strategies. A SLR and qualitative content analysis of 48 sources identified 74 risks and 111 mitigation measures: The four most mentioned risk groups were (i) societal effects, (ii) human limitations, (iii) vulnerabilities, characteristics of user-generated

content platforms, and (iv) algorithmic quality issues. Our study found detection measures mentioned the most, followed by prevention measures; recovery measures were only slightly represented. The most mentioned detection measures were 'algorithmic and artificial intelligence supported measures' suggesting that algorithmic measures are used unproportionally more often as a defense mechanism compared to other strategies. We suggest that further investment in prevention measures could significantly help in the fight against fake UGC, especially considering most detection measures were found in 'algorithmic and artificial intelligence supported measures' which might not be in the skillset of smaller regional platform providers or traditional companies.

Algorithmic & AI Supported Measures	**Algorithmic Quality Issues**
analysis is multimodal and includes images, comments, reactions, user profiles, and other relevant data associated with UGC-like metadata [D] | linguistic and semantic analysis [D] | pattern-recognition or deviation analysis [D] | detection techniques appropriate for different types of fake content [D] | generalize detection algorithms and AI to be event-independent [D] | analyze sharing and propagation patterns [D] | user behavior analysis [P+D] | assess and highlight the credibility and trustworthiness of an information's source [P] | checking if external online-sources are of low quality, fake or "trash" sites [D] | feedback-based identification by using a user's responses and comments [D] | train generative AI with user responses to a specific content better early mitigation and detection [D] | consider a user's viewpoint and ideology [D] | prioritizing newer content over older content [P] | limiting the number and frequency if posting activities [P] | enable effective early detection and intervention [D] | remove fake content [R] | remove duplicates of previously identified fake UGC and identify their common sources [D] | identify and eliminate influential fake spreaders [D] | check content automatically for objective facts [D] | check verified sources using structured data, knowledge graphs, and linked data [D] | verify by third parties for truthfulness and authenticity [P+D] | pay to small niche areas or communities [P] | preferring usefulness, trustworthiness, and authoritativeness over determining the objective truth [D]	
(Brundage et al., 2018; Carter et al., 2021; Chen et al., 2023; Conroy et al., 2015; Crowell, 2017; Fredheim & Pamment, 2024; Glassdoor, 2023b, 2023a, 2024; Google, 2019; Gupta, 2023; Gupta, Kumar, et al., 2022; Gupta, Li, et al., 2022; Islam et al., 2020; Kertysova, 2018; Li et al., 2018; Meta, 2020, 2024; Mut Camacho & Rueda Lozano, 2021; Papanastasiou, 2018; Pherson et al., 2020; Pritchett, 2024; Rosen, 2021; Rubin, 2019; Scherr et al., 2019; Sharma et al., 2019; Shetty et al., 2023; Shu et al., 2020; Twitter, 2024; Yang et al., 2022; Zelenka et al., 2021; Zhang & Ghorbani, 2020; Zhou et al., 2019, 2020; Zhou & Zafarani, 2020)	
Strategic Measures	**Newness of (and Speed of Change in) the Fake User Generated Content Domain**
set up a mitigation-strategy, response plan, and identify risks [P+R] | screen for influence operations by governments or other large entities and combat them [P+D] | Real-time visualization of data for analysis and pattern spotting [D] | measure the effectiveness of interventions and protective measures [D] | constantly evolve the strategy and used technology [P+D] | establishing a crisis or a task-force team [R] | increase R&D in relevant areas [P] | hybrid AI-human approach [D] | partnerships between UGC service providers [D] | establish universal standards and global screening protocols [D] | consider multiple roles [P] | consider societal values, i.e., collectivistic/individualistic characteristics of a country's population, when designing measures [P] | use of federated governance models for content moderation [P]	
(Carter et al., 2021; Chen et al., 2023; Conroy et al., 2015; Coutinho & José, 2019; Dabbous et al., 2021; Donaker et al., 2019; Einwiller & Kim, 2020; Google, 2019; Gupta, Li, et al., 2022; Kertysova, 2018; Li et al., 2018; Mut Camacho & Rueda Lozano, 2021; Pherson et al., 2020; Pritchett, 2024; Rosen, 2021; Sharma et al., 2019; Shu–Min Hou et al., 2023; Twitter, 2024; Zelenka et al., 2021; Zhang & Ghorbani, 2020)	
Improve Platform Characteristics	**Characteristics of User Generated Content Platforms | Inherent Risks Because of Characteristics of Fake User Generated Content**
slow down and limit the spread of misinformation [P] | detecting and limiting clickbait [P+D] | elevated protections during major events [P] | do not personalize content and recommendations in critical situations [P] | hide an article's past sharing statistics [P] | limiting political ads to own-country creators [P] | authentication and verification of users [P] | support of quality journalism by providing additional tools, software, and analytics, especially for them [P] | proactively posting quality news content in closed platforms, e.g. in a trustable daily bulletin style [P] | reduce and pay attention to echo chambers [P] | digital watermarking [P+D] | whitelisting [P]	
(Brundage et al., 2018; Chen et al., 2023; Coutinho & José, 2019; Crowell, 2017; Donaker et al., 2019; Facebook, 2024; Glassdoor, 2023a; Google, 2019; Gupta, 2023; Gupta, Kumar, et al., 2022; Islam et al., 2020; Kertysova, 2018; Meta, 2020, 2024; Mut Camacho & Rueda Lozano, 2021; Papanastasiou, 2018; Pherson et al., 2020; Pritchett, 2024; Rosen, 2021; Scherr et al., 2019; Sharma et al., 2019; Shu et al., 2020; Twitter, 2024; Yang et al., 2022; Zelenka et al., 2021; Zhou et al., 2020)	

Supporting Human Fake Detection	**Human Limitations, Vulnerabilities & Bias | Manipulation of Information | Scientific Facts and Perception**
Offering verified real-time facts, such as voting numbers and dates [P] | Providing reliable information about a topic next to (potential) fake UGC [R] | Highlighting a special status of a "credible source" and removing it if "bad" content is identified [P] | Allowing and promoting users explaining why their content is credible [P] | mandating that users cite multiple sources [P] | Corrective communication by platforms providing alternative truth and addressing fake claims [R] | disclose positive inspection outcomes to users [P+D] | correct false content [R] | show warning notices or labels to users when content is identified as potentially misleading or false [P] | Actions should vary based on the current type of the detected problem – is the UGC "misleading, disputed, or unverified [P+D] | implementing in-app/in-service fact-checking to reduce the effort of the checking process [P+D] | Understanding the psychological aspects of fake UGC [P+D] | Inoculating users with fake UGC [P] | Forcing users who have shown poor performance in detecting fake UGC to (re)earn platform abilities by completing educational content [P]	
(Chen et al., 2023; Facebook, 2024; Google, 2019; Gupta, Kumar, et al., 2022; Gupta, Li, et al., 2022; Kertysova, 2018; Lim, 2020; Lunga & Mthembu, 2019; Meta, 2020, 2024; Papanastasiou, 2018; Pherson et al., 2020; Pritchett, 2024; Rosen, 2021; Roth & Pickles, 2020; Rubin, 2019; Shu et al., 2020; Twitter, 2024; Zhou & Zafarani, 2020)	
Peergroup Control	**Characteristics of User Generated Content Platforms | Manipulation of Information | Scientific Facts and Perception | Fairness of Process | Issues with Fact Checkers | Tactics of Adversaries | Algorithmic Quality Issues**
enable the collaborative creation of UGC [P] | allow ratings and feedback on fact-checking (the process and individual fact-checkers) [D] | content flagging, crowd sourcing, or voting processes [D] | social accountability measures [P] | allow users to add info and data confirming truth (evidence) [D+R] | require users to proofread other UGC [P]	
(Brundage et al., 2018; Chen et al., 2023; Coutinho & José, 2019; Donaker et al., 2019; Facebook, 2024; Glassdoor, 2023b, 2024; Google, 2019; Gupta, Kumar, et al., 2022; Islam et al., 2020; Meta, 2024; Pherson et al., 2020; Pritchett, 2024; Sharma et al., 2019; Shu et al., 2020; Twitter, 2024, 2022; Zelenka et al., 2021; Zhou & Zafarani, 2020)	
Manual Fact-Checking	**Algorithmic Quality Issues | Issues with Fact Checkers | Tactics of Adversaries**
a user's self-assessment of whether something is fake [P+D] | Forensic tools [D] | browser extension fake detectors [D] | Fake Content Review Sites (for example, Fact Checking Portals [D] | An independent news verification service should be established [D]	
(Chen et al., 2023; Dabbous et al., 2021; Gupta, Kumar, et al., 2022; Islam et al., 2020; Kertysova, 2018; Lim, 2020; Lunga & Mthembu, 2019; Papanastasiou, 2018; Pherson et al., 2020; Rubin, 2019; Zhang & Ghorbani, 2020; Zhou & Zafarani, 2020)	
Improve Transparency and Accountability	**Fairness of Process | Privacy Concerns | Issues with Fact Checkers | Erosion of Reputation & Trust**
establish something akin to an appeals court that provides the right to appeal decisions [D+R] | usage of explainable AI [D] | users should get explanations of why they see an ad [P] | Users should be able to anonymously provide feedback [R] | private feedback directly to the UGC provider (e.g. the company) should be enabled [D]	
(Chen et al., 2023; Donaker et al., 2019; Glassdoor, 2023a; Google, 2019; Gupta, 2023; Islam et al., 2020; Kertysova, 2018; Meta, 2024; Pherson et al., 2020; Sharma et al., 2019; Shu–Min Hou et al., 2023; Twitter, 2024; Zelenka et al., 2021; Zhou & Zafarani, 2020)	
Punish & Limit Bad Actor	**Powerful & Organized Adversaries | Tactics of Adversaries**
Fake accounts need to be closed or suspended [R] | bots and bot networks must be detected [D] | demote or pre-hide potential fake UGC and fake accounts [P] | sue, combat, and denounce malicious players [P+D] | Access should be blocked for malicious parties (potentially their devices, networks, companies, and even countries) [P] | punishments being administered should be increased with each offense [P] | restrict or even disallow fake UGC creators from the ability to place ads on UGC platforms [P]	
(Crowell, 2017; Facebook, 2024; Glassdoor, 2023b, 2023a; Google, 2019; Gupta, 2023; Gupta, Kumar, et al., 2022; Kertysova, 2018; Meta, 2024; Mut Camacho & Rueda Lozano, 2021; Pherson et al., 2020; Pritchett, 2024; Rosen, 2021; Roth & Pickles, 2020; Shu et al., 2020; Twitter, 2024; Zelenka et al., 2021; Zhou & Zafarani, 2020)	
Education	**Erosion of Reputation & Trust | Societal Effects & Issues | Manipulation of Information | Scientific Facts and Perception**
increasing media and digital literacy [P] | onboarding systems that educate users about fake UGC, spotting bad sources, and so on [P] | minimum educational requirements for influential posters and journalists [P] | improving their self-efficacy [P] | Increasing users' trust in authorities, moderators, and credible sources [P+R] | improving the reputation of high-quality news outlets and journalists [P+R] | Gamification [P] | decontaminate affected users [R]	
(Chen et al., 2023; Dabbous et al., 2021; Facebook, 2024; Google, 2019; Gupta, Kumar, et al., 2022; Kertysova, 2018; Lim, 2020; Lunga & Mthembu, 2019; Melki et al., 2021; Pherson et al., 2020; Rosen, 2021; Rubin, 2019; Sharma et al., 2019; Shu et al., 2020; Shu–Min Hou et al., 2023; Twitter, 2024; Zhang & Ghorbani, 2020)	
Change Incentives	**Misaligned Incentives**

incentives and the business model of UGC providers need to be aligned with wanting to provide users with truthful and high-quality content [P] | increase cost incurred, be it financials or increased needed effort, when posting fake UGC [P]	
(Brundage et al., 2018; Chen et al., 2023; Donaker et al., 2019; Fredheim & Pamment, 2024; Google, 2019; Gupta, Li, et al., 2022; Papanastasiou, 2018; Rosen, 2021; Zhou & Zafarani, 2020)	

Table 2: Top10 Risks Mitigation Strategies to Fight Fake UGC

Regulation and legislation present a dual challenge: they establish frameworks for action and penalties against fake content providers, while loopholes can create risks (Gupta, Kumar, et al., 2022). This study advocates for legislation that aligns corporate incentives with societal needs. Although, education is mentioned in some literature (e.g., Chen et al., 2023; Dabbous et al., 2021), other strategies to address negative societal impacts remain limited. However, we think reducing fake UGC benefits society and vulnerable groups like children. Media literacy, critical thinking, and education foster informed consumers (Pfadenhauer & Grenz, 2012) and reduce sharing of fake news (Dame Adjin-Tettey, 2022). Media literacy was underrepresented in our literature sample with only 18 mentions. However, education programs like in Finland (Kertysova, 2018; OSIS, 2018) demonstrate its value. We conclude that educational measures remain underfunded and should get greater investments.

Current literature seems to be limited in the number of discovered risks and measures. The most matching risks of this paper with a single source were 32 (of 74) found in Gupta, Li et al. (2022), while the most matching measures were 45 (of 111) found in a publication from Google (2019) provided for the 2019 Munich Security Conference. This suggests that the reviewed literature only covered the spectrum of risks and measures in its totality, but individual studies did not. Perhaps, because individual studies focus more on providing easily accessible entry points into research in subareas of UGC, but did not focus on a complete defense strategy. This non-holistic view of previous research, we aimed to overcome with our FIRM catalogue.

Our study reveals a wide range of risks and measures against fake UGC, emphasizing the need for a multi-faceted approach to defense against fake information. However, companies need resources for UGC moderation and small teams may struggle to develop these measures themselves. Hence, the FIRM catalogue with its taxonomy of 16 risks and 15 mitigation strategies, shall allow a focused starting point and assist regional or local platform providers in combating disinformation. Furthermore, we think that most measures of our FIRM catalogue could be applied to various levels, including governments and NGOs. Measures like increasing media literacy should be implemented at both company level and government, i.e. citizen education. Techplomacy or collaboration amongst industrial associations would be another approach to share resources and knowhow about patterns of fake UGC. From our perspective, the top five measures should be: (i) increasing media and digital literacy, (ii) establishing internal fact checkers, (iii) use technical solutions for linguistic analysis, (iv) support for content flagging, and (v) assessing information credibility and trustworthiness. It was found that neither companies nor governments can solve the issue alone, so restricting measures to just one perspective may diminish the overall results.

6. Conclusion, Limitations and Further Research

The study examines risks posed by fake UGC and potential protective measures. Based on a systematic literature review, 74 risks and 111 protective measures associated with fake UGC were identified and documented. The presented FIRM catalog consists of a taxonomy of 16 risk categories and 15 mitigation strategies and encompasses in detail the subordinate risks and

safeguards. This grouping allows practitioners to look for specific risks and their mitigation strategies and to keep oversight. Frequently cited risks include societal impact, human weaknesses, platform characteristics, and algorithmic quality issues. Mitigation strategies focus on AI-powered approaches, strategic actions, platform improvements. An imbalance between preventive and detective measures was found, with preventive measures underrepresented. Furthermore, educational measures seem to be insufficiently researched. We recommend focusing on education and media literacy as this countermeasure suggest being an effective tool against negative societal impacts caused by fake UGC. Limitations to this study might occur by applying just two scientific data bases or potentially outdated information from the grey literature. Future research could focus on different types of UGC, e.g. graphical content, follow up on updates to mitigation strategies, and create consistent standards for technical solutions. Our FIRM catalog aimed to be a starting point for research that could lead to a framework to help companies integrate UGC and protection strategies at the same time.

References

Agnihotri, D., Ahmed, S., Darekar, D., & Gadkari, C. (2021). *Examination of Fake Reviews of Product and Approach to Filter them from Legitimate Reviews*. 2021 IEEE Pune Section International Conference, PuneCon 2021. Scopus.

Baeza-Yates, R., & Saez-Trumper, D. (2015). Wisdom of the Crowd or Wisdom of a Few? An Analysis of Users' Content Generation. *Proceedings of the 26th ACM Conference on Hypertext & Social Media*, 69–74.

Bodrožić, Z., & S. Adler, P. (2022). Alternative Futures for the Digital Transformation: A Macro-Level Schumpeterian Perspective. *Organization Science*, *33*(1), 105–125.

Bond. (2015). *Internet Trends 2015*. https://www.bondcap.com/report/it15

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G. C., Steinhardt, J., Flynn, C., hÉigeartaigh, S. Ó., Beard, S., Belfield, H., Farquhar, S., … Amodei, D. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation* (No. arXiv:1802.07228). arXiv.

Carter, M., Tsikerdekis, M., & Zeadally, S. (2021). Approaches for Fake Content Detection: Strengths and Weaknesses to Adversarial Attacks. *IEEE Internet Computing*, *25*(2), 73–83. Scopus.

Chen, S., Xiao, L., & Kumar, A. (2023). Spread of misinformation on social media: What contributes to it and how to combat it. *Computers in Human Behavior*, *141*, 107643.

Conroy, N. K., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, *52*(1), 1–4. Scopus.

Coutinho, P., & José, R. (2019). A Risk Management Framework for User-Generated Content on Public Display Systems. *Advances in Human-Computer Interaction*, *2019*, 1–18. https://doi.org/10.1155/2019/9769246

Crowell, C. (2017). *Our approach to bots and misinformation*. https://blog.x.com/en_us/topics/company/2017/Our-Approach-Bots-Misinformation

Dabbous, A., Aoun Barakat, K., & de Quero Navarro, B. (2021). Fake news detection and social media trust: A cross-cultural perspective. *Behaviour & Information Technology*, *41*(14), 2953–2972.

Dame Adjin-Tettey, T. (2022). Combating fake news, disinformation, and misinformation: Experimental evidence for media literacy education. *Cogent Arts & Humanities*, *9*(1), 2037229.

Donaker, G., Kim, H., & Luca, M. (2019). Designing better online review systems: How to create ratings that buyers and sellers can trust. *Harvard Business Review*, *2019*(November-December), 1–9. Scopus.

Einwiller, S., & Kim, S. (2020). How Online Content Providers Moderate User-Generated Content to Prevent Harmful Online Communication: An Analysis of Policies and Their Implementation. *Policy & Internet*, *12*.

Eppler, M. J. (2006). A Comparison between Concept Maps, Mind Maps, Conceptual Diagrams, and Visual Metaphors as Complementary Tools for Knowledge Construction and Sharing. *Information Visualization*, *5*(3), 202–210.

Erkayhan, D. S. (2020, July 28). Digitale Plattformen und Plattformökonomie. *SEEBURGER Blog*. https://blog.seeburger.com/de/digitalisierung-und-plattformoekonomie-im-maschinen-und-anlagenbau/

Facebook. (2024). *Wie geht Facebook mithilfe von unabhängigen Faktenprüfern mit Falschmeldungen um? | Facebook-Hilfebereich*. https://www.facebook.com/help/1952307158131536

Farhoudinia, B., Ozturkcan, S., & Kasap, N. (2023). Fake news in business and management literature: A systematic review of definitions, theories, methods and implications. *Aslib Journal of Information Management*, *ahead-of-print*(ahead-of-print).

Fredheim, R., & Pamment, J. (2024). Assessing the risks and opportunities posed by AI-enhanced influence operations on social media. *Place Branding and Public Diplomacy*. Scopus.

Garousi, V., Felderer, M., & Mäntylä, M. V. (2019). Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and Software Technology*, *106*, 101–121.

Giełczyk, A., Wawrzyniak, R., & Choraś, M. (2019). *Evaluation of the existing tools for fake news detection: Vol. 11703 LNCS* (p. 151). Scopus.

Girgenti, R. H., & Hedley, T. P. (Eds.). (2011). *Managing the risk of fraud and misconduct: Meeting the challenges of a global, regulated, and digital environment*. McGraw-Hill.

Glassdoor. (2023a). *Review Integrity & Fraud Detection | Glassdoor Help Center*. https://help.glassdoor.com/s/article/Review-Integrity-Fraud-Detection?language=en_US

Glassdoor. (2023b). *Review Integrity and Anonymity Submission | Glassdoor Help Center*. https://help.glassdoor.com/s/article/Review-Integrity-and-Anonymity-Submission?language=en_US

Glassdoor. (2024). *Fraudulent reviews | Glassdoor Help Center*. https://help.glassdoor.com/s/article/Fraudulent-reviews?language=en_US

Google. (2019, February 16). *Fighting disinformation across our products—How google fights disinformation*. Google. https://blog.google/around-the-globe/google-europe/fighting-disinformation-across-our-products/

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, *26*(2), 91–108.

Gupta, A. (2023, March 31). *New ways we tackle fake contributions on Google Maps*. Google. https://blog.google/products/maps/google-maps-fake-contributions-ai-machine-learning/

Gupta, A., Kumar, N., Prabhat, P., Gupta, R., Tanwar, S., Sharma, G., Bokoro, P. N., & Sharma, R. (2022). Combating Fake News: Stakeholder Interventions and Potential Solutions. *IEEE Access*, *10*, 78268–78289.

Gupta, A., Li, H., Farnoush, A., & Jiang, W. (2022). Understanding patterns of COVID infodemic: A systematic and pragmatic approach to curb fake news. *Journal of Business Research*, *140*, 670–683. Scopus.

Islam, M. R., Liu, S., Wang, X., & Xu, G. (2020). Deep learning for misinformation detection on online social networks: A survey and new perspectives. *Social Network Analysis and Mining*, *10*(1). Scopus.

Kertysova, K. (2018). Artificial Intelligence and Disinformation. *Security and Human Rights*, *29*(1–4), 55–81.

KI-gesteuerte Bots und die Verbreitung von Desinformation. (2024, June 4). onlinesicherheit.at., https://www.onlinesicherheit.gv.at/Services/News/KI-Bots-Desinformation.html

Kirchner, S., & Schüßler, E. (2019). The Organization of Digital Marketplaces: Unmasking the Role of Internet Platforms in the Sharing Economy. In G. Ahrne & N. Brunsson (Eds.), *Organization outside Organizations* (1st ed., pp. 131–154). Cambridge University Press.

Kitchenham, B., & Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering*. *2*.

Li, C.-S., Xiong, G., & Tapia, E. M. (2018). New frontiers in cognitive content curation and moderation. *APSIPA Transactions on Signal and Information Processing*, *7*(1). Scopus.

Lim, S. (2020). Academic library guides for tackling fake news: A content analysis. *The Journal of Academic Librarianship*, *46*(5), 102195. Scopus.

Lunga, C. M., & Mthembu, M. V. (2019). Investigating the Source and Strategies Adopted by Mainstream Media in Combating Fake News in the Kingdom of Eswatini. *African Journalism Studies*, *40*(4), 96–111. Scopus.

Mayring, P. (2000). Qualitative Content Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [On-Line Journal], Http://Qualitative-Research.Net/Fqs/Fqs-e/2-00inhalt-e.Htm*, *1*.

Melki, J., Tamim, H., Hadid, D., Makki, M., El Amine, J., & Hitti, E. (2021). Mitigating infodemics: The relationship between news exposure and trust and belief in COVID-19 fake news and social media spreading. *PLOS ONE*, *16*(6), e0252830. Scopus.

Meta. (2020). *Here's how we're using AI to help detect misinformation*. https://ai.meta.com/blog/heres-how-were-using-ai-to-help-detect-misinformation/

Meta. (2024). *Faktenprüfung auf Facebook, Instagram und Threads*. Meta-Hilfebereich für Unternehmen. https://de-de.facebook.com/business/help/2593586717571940

Mut Camacho, M., & Rueda Lozano, A. (2021). DISINFORMATION AS A CORPORATE RISK: THE NEED FOR A NEW METHODOLOGICAL APPROACH. *Vivat Academia. Revista de Comunicación*, *155*, 113–129.

Naab, T. K., & Sehl, A. (2016). Studies of user-generated content: A systematic review. *Journalism*, *18*(10), 1256–1273.

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. *European Journal of Information Systems*, *22*(3), 336–359.

Obermaier, R., & Mosch, P. (2019). Digitale Plattformen – Klassifizierung, ökonomische Wirkungslogik und Anwendungsfälle in einer Industrie 4.0. In R. Obermaier (Ed.),

Handbuch Industrie 4.0 und Digitale Transformation: Betriebswirtschaftliche, technische und rechtliche Herausforderungen (pp. 379–417). Springer Fachmedien.

OpenAI. (2022, November 30). *Introducing ChatGPT*. https://openai.com/index/chatgpt/

Ozbay, F. A., & Alatas, B. (2020). Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics and Its Applications*, *540*, 123174.

Papanastasiou, Y. (2018). *Fake News Propagation and Detection: A Sequential Model* (SSRN Scholarly Paper No. 3028354).

Pfadenhauer, M., & Grenz, T. (2012). Anzeichen einer neuen Partizipations- kultur? *Medien +Erziehung*, *Jg. 56*(6), 17–32.

Pherson, R. H., Mort Ranta, P., & Cannon, C. (2020). Strategies for Combating the Scourge of Digital Disinformation. *International Journal of Intelligence and CounterIntelligence*, *34*(2), 316–341.

Pritchett, D. (2024, February 13). *How machine learning keeps contributed content helpful*. Google. https://blog.google/products/maps/how-machine-learning-keeps-contributed-content-helpful/

Rosen, G. (2021, March 22). How We're Tackling Misinformation Across Our Apps. *Meta*. https://about.fb.com/news/2021/03/how-were-tackling-misinformation-across-our-apps/

Roth, Y., & Pickles, N. (2020). *Updating our approach to misleading information*. https://blog.x.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information

Rubin, V. L. (2019). Disinformation and misinformation triangle. *Journal of Documentation*, *75*(5), 1013–1034. Scopus.

Sandvine. (2024, March). *Global Internet Phenomena*. https://www.sandvine.com/phenomena

Santos, M. L. B. dos. (2021). The "so-called" UGC: An updated definition of user-generated content in the age of social media. *Online Information Review*, *46*(1), 95–113.

Scherr, S. A., Polst, S., & Elberzhager, F. (2019). *Beware of the Fakes – Overview of Fake Detection Methods for Online Product Reviews: Vol. 11578 LNCS* (p. 467). Scopus.

Shapiro, C., & Varian, H. R. (1998). *Information rules: A strategic guide to the network economy*. Harvard Business Review Press.

Sharma, K., Qian, F., Jiang, H., Ruchansky, N., Zhang, M., & Liu, Y. (2019). Combating Fake News. *ACM Transactions on Intelligent Systems and Technology*, *10*(3), 1–42. Scopus.

Shetty, M., Mahesh, Shetty, S. B., & Shetty, A. (2023). *Automated Method for Fake News Detection Using Machine Learning*. 2023 International Conference on Network, Multimedia and Information Technology, NMITCON 2023. Scopus.

Shu, K., Bhattacharjee, A., Alatawi, F., Nazer, T. H., Ding, K., Karami, M., & Liu, H. (2020). Combating disinformation in a social media age. *WIREs Data Mining and Knowledge Discovery*, *10*(6), e1385.

Shu–Min Hou, Shao–Yi Lai, & Wen-Cheng Fu. (2023). Exploring Information Warfare Strategies during the Russia–Ukraine War on Twitter. *The Korean Journal of Defense Analysis*, *35*(1), 19–44. Scopus.

Twitter. (2024). *Twitter Community Notes Challenges*. https://communitynotes.x.com/guide/en/about/challenges

Twitter, P. (2022). *An update on reporting potential misinformation on Twitter*. https://blog.x.com/en_sea/topics/company/2022/update-on-reporting-potential-misinformation-on-twitter

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151.

Yang, X., Zhang, J., Chen, K., Zhang, W., Ma, Z., Wang, F., & Yu, N. (2022). *Tracing Text Provenance via Context-Aware Lexical Substitution*. *36*, 11613–11621. Scopus.

Zelenka, J., Azubuike, T., & Pásková, M. (2021). Trust Model for Online Reviews of Tourism Services and Evaluation of Destinations. *Administrative Sciences*, *11*(2), 34. Scopus.

Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, *57*(2), 102025. Scopus.

Zhou, X., Jain, A., Phoha, V. V., & Zafarani, R. (2020). Fake News Early Detection. *Digital Threats: Research and Practice*, *1*(2), 1–25. Scopus.

Zhou, X., & Zafarani, R. (2020). A Survey of Fake News. *ACM Computing Surveys*, *53*(5), 1–40. Scopus.

Zhou, X., Zafarani, R., Shu, K., & Liu, H. (2019). *Fake News: Fundamental theories, detection strategies and challenges*. 836–837. Scopus.

11. Perceptions and Governance of Emerging Technologies in New Zealand: Preparing the Next Generation for an IT-Driven Future

Shafiq Alam
Massey University
s.alam1@massey.ac.nz

Kasuni Weerasinghe
Auckland University of Technology
kasuni.weerasinghe@aut.ac.nz

David J. Pauleen
National Chung Cheng University
d.pauleen@protonmail.com

Hamed Jafarzadeh
MACQUARIE University
hamed.jafarzadeh@mq.edu.au

Nazim Taskin
Boğaziçi University
nazim.taskin@bogazici.edu.tr

Ji Yu
Massey University
j.yu3@massey.ac.nz

Abstract
In the era of Industry 4.0, the world is rapidly evolving with emerging technologies like Artificial Intelligence and the Internet of Things. These technologies bring various benefits to sectors, such as healthcare, education, governance, and transportation, but they also raise concerns about privacy, security, and usability. To maximise the advantages and minimise negative effects, appropriate policies and regulations are crucial. Understanding public perceptions of emerging technologies is vital for their acceptance and effective governance. This research focuses on New Zealand, aiming to comprehend public perceptions through a survey conducted in three stages: literature review, survey instrument development, and empirical data collection. Initial findings from the survey of 450 responses indicate that the public shows moderate support for emerging technologies in New Zealand. However, addressing concerns related to data privacy and accountability is essential. Future work involves testing correlations between knowledge levels and perceptions/concerns, revising the survey instrument, and conducting a nationwide survey with a demographic and ethnicity-based approach to gauge public sentiment.

Keywords
Emerging Technologies, Artificial Intelligence, Facial Recognition, Virtual Reality

1 Introduction
The pursuit of technological advancement has driven human development, opening up new opportunities and transforming society. The Industry 4.0 revolution, characterised by digital transformation and innovative technologies like Artificial Intelligence (AI), Virtual Personal Assistants (VPAs), and immersive technologies, has significantly enhanced organisational productivity and addressed commercial and societal challenges (Castello & van der Meulen, 2018). A recent study from New Zealand found that more than 50% of the organisations polled thought AI would significantly affect individuals, society, and the industry or that it has already had this effect (Robb et al., 2020). New Zealand government, commercial, and educational sectors have

embraced emerging technologies to increase productivity and efficiency at work (Chen & Chang-Richards, 2022). However, for these emerging technologies to make a significant impact, they must first emerge, grow, and be adopted by society.

Although the word 'emerging technologies' is typically used to describe a new technology, it can also apply to an existing technology that is still being developed (Rotolo et al., 2015). The term is commonly reserved for technologies that have, or are anticipated to have, substantial societal or economic implications. In Aotearoa New Zealand (NZ), emerging technologies, such as the Internet of Things (IoT), blockchain, big data, cybersecurity, robotics, and Augmented Virtual Reality (AVR) have been steadily integrated into everyday activities. AI is being used, for instance, in the healthcare industry to facilitate the diagnosis and treatment of chronic illnesses (Jiang et al., 2017). 3D printing can significantly cut carbon emissions and aid in recycling used materials (Ikram, 2022). The NZ government leverages big data to develop precise and efficient climate action strategies (Yukich Clendon et al., 2023).

Recent implementations of facial recognition by NZ police is an example that highlights the importance of public perceptions of emerging technologies in NZ (Block, 2019). Facial recognition technology offers benefits, but also raises privacy concerns over data collection for individuals (Shufeng, 2021). For use in NZ, facial recognition technology must undergo a high level of scrutiny and testing on its accuracy as mistakes in identifying a person could be damaging to individuals (Lynch et al., 2022). Other significant risks include: (1) AI having embedded ethnic, religious, or social biases that disadvantage a particular person or group (Gavighan et al., 2019; Soleimani et al., 2021); (2) risks to privacy inherent in AI systems as they collect enormous volumes of data, which can extract patterns and information from the data (Bulchand-Gidumal, 2020); (3) automated analysis utilising machine learning (ML) algorithms that access consumer data and which can increase the risks to individual privacy (Kopalle et al., 2022; Kathriarachchi et al., 2024) by connecting diverse data sources; and (4) automated and augmented decision making by AI and ML, which raises issues of accountability and transparency over how and why decisions are made (Mittelstadt et al., 2016).

As a result, researchers, scientists, and decision-makers are seriously debating the societal ramifications of these emerging technologies (Zhang & Dafoe, 2020). Numerous studies, including that of West et al. (2019), argue that the general public should play a significant role in influencing the development and use of such technologies, given the risk factors. It is crucial to understand how the general public, including stakeholders (i.e., corporations, governments, and researchers), view these technologies and their effects to allow relevant agencies to inform and engage the public about emerging technologies' features, prospects, and pitfalls and develop and implement appropriate policies (Zhang & Dafoe, 2020; Kathriarachchi et al., 2024).

This study aims to promote a comprehensive understanding of emerging technologies (AI, AVR, facial and voice recognition, IoT, and robotics) in NZ, focusing on how the public perceives their impact, including benefits and risks. The paper is organised as follows: the next section will provide an overview of related regulations, followed by section 3, the literature review which discusses the development of the survey. Section 4 presents the results of the survey followed by conclusions and recommendations for future research.

2 The New Zealand Context

The government of Aotearoa New Zealand (NZ) is tightening its restrictions on using emerging technologies. Proper governance is crucial to maximising the advantages of employing emerging technologies while minimising their drawbacks. Instead of developing specific regulations to restrict the use of AI, the NZ government, for instance, is looking at progressively including AI controls into existing policies and laws as they are revised and updated (Stuart, 2020). The following are three examples of current and possible regulatory responses to the employment of algorithms in emerging technologies that worry the NZ government[1]. Firstly, the Opaque Nature of AI Decision-making: "If we cannot understand how a decision has been made, we cannot check whether it is accurate or biased" (New Zealand Legislation, 1982). According to Section 23 of NZ's Official Information Act 1982, when a public service agency, Minister of the Crown, or organisation makes a decision or recommendation in respect of any person, the person has the right to and shall be given a written statement of the findings on material issues of fact, a reference to the information on which the findings were based, and the reasons for the decision or recommendation (New Zealand Legislation, 1982). If so, the justifications must be provided in writing with any conclusions and a mention of the data that served as the foundation for the results. Secondly, Privacy: As revealed by the New Zealand Human Rights Commission (2018), privacy concerns are common in the information age. AI-driven decision-making systems gather data from diverse sources, including non-personal and personally identifiable information in the AI algorithm context. The New Zealand Privacy Act 2020 (New Zealand Legislation, 2020) can be the starting point for data collection and use regulations. Finally, Biases in Algorithms: for instance, because of the lack of female voices in the training datasets, some speech recognition algorithms may handle female sounds less accurately than male voices (Tatman, 2016). Current regulations in NZ do not cover algorithmic decision-making. However, the concerns about algorithmic biases are not limited to this situation and may involve more broad and existing NZ legal obligations (Gavighan et al., 2019).

Public opinion is crucial for the NZ government to adopt or revise policies related to emerging technologies. Ignoring public views, including those of researchers, technologists, corporations, and governments, can lead to social, governance, and technology-related problems. Understanding the public's perception of emerging technologies and their governance is essential for creating informed policies and promoting public awareness about the opportunities and challenges they present (Zhang & Dafoe, 2020). With this in mind, the current research aims to explore the public perceptions of emerging technologies in NZ, and how they vary in terms of knowledge level, support, concerns, and governance for emerging technologies. The research study is divided into two distinct phases, with this paper focusing on Phase 1. In Phase 1, the research team undertook three crucial stages: a comprehensive literature review, the development of a survey instrument, and a pilot study. The literature review provided essential insights into emerging technologies and public perceptions in NZ. Building on this knowledge, the researchers crafted a well-structured survey instrument to capture the public's perspectives on these technologies. To test the effectiveness of the survey, a pilot study was conducted, gathering 450 responses for analysis. This paper presents a descriptive analysis of the pilot survey findings, shedding light on the public's knowledge, attitudes, concerns, and governance preferences regarding emerging technologies.

[1] https://www.mbie.govt.nz/dmsdocument/5754-artificial-intelligence-shaping-a-future-new-zealand-pdf

Phase 2 will involve a full-scale survey to delve deeper into the data collected, enabling policymakers and researchers to make informed decisions and establish appropriate policies and regulations for the optimal utilisation of emerging technologies in NZ.

3 Literature Review

To conduct a structured literature review, relevant literature was identified using three academic databases (Scopus, Web of Science, and Discover) to search for peer-reviewed journal articles published from 2020 to 2022. The chosen time frame aimed to capture the rapid developments and trends in public perceptions of emerging technologies. After eliminating duplicates and excluding irrelevant articles based on abstracts, 148 related papers were identified. Out of these, 59 papers were considered relevant after examination of their full texts. Of the relevant studies, 32 incorporated surveys to gauge public perceptions (Table 1), with 25 of them utilising survey questionnaires. Among the technologies studied, Artificial Intelligence (AI) emerged as a prominent topic, with a focus on healthcare, and predominantly from the United States.

In addition to academic sources, NZ-specific information was gathered from various sources, including government, university, and organisation websites, such as Digital.govt.nz and Data.govt.nz. Notably, AI was a dominant theme in NZ sources, with articles discussing its current use, benefits, regulatory issues, and future prospects, underscoring the significant role and potential of AI in the country's context[2,3]. Moreover, the researchers closely monitored the usage of emerging technologies in NZ, particularly by the government, by referencing recent mass media reports[4] (e.g., Radio New Zealand).

Before designing the final questionnaire, a group of government agencies were surveyed with the cooperation of the Multi Agency Research Network (MARN) [5]. MARN members provided their views on: 1) Preferred emerging technologies for the survey; 2) Important aspects of public perceptions towards emerging technologies; 3) Recommended sources for the project; and 4) Awareness of similar publicly available questionnaires. In the initial survey, we received 14 valid responses from different NZ government agencies (see Fig 1). They expressed interest in various emerging technologies, with the top eight being AI, facial recognition, robotics, IoT, voice recognition, the metaverse, AVR, and 5G. They recommended asking the public about their knowledge level, information sources, opinions on technology use, concerns, satisfaction with regulations, and suggestions for government regulations. Additionally, participants provided relevant sources for the research.

Based on the review of the related literature, and the results of the initial survey from the 14 respondents, we designed a questionnaire with 36 questions to capture the public's perceptions of emerging technologies . We divided the questionnaire into four sections. The first section collects the demographic information about participants. Everyone living in NZ above the age of 18 was eligible to fill out the questionnaire.

[2] AI Forum NZ, "Artificial Intelligence: Shaping a Future New Zealand." Available: https://aiforum.org.nz/reports/artificial-intelligence-shaping-a-future-new-zealand/

[3] C. Gavighan, A. Knott, J. Maclaurin, J. Zerilli, and J. Liddicoat, "Government use of artificial intelligence in New Zealand." Available: https://ourarchive.otago.ac.nz/handle/10523/9372.

[4] We searched government websites such as https://www.digital.govt.nz/, academic websites such as https://www.otago.ac.nz/caipp/index.html, and organisational websites such as https://aiforum.org.nz/.

[5] The Multi Agency Research Network (MARN) brings together New Zealand government agencies who are interested in research collaboration between each other and NZ universities.

#	Citation	Description	Technology	Country	Area	Questionnaire available? (Y/N)
1	(Chatterjee & Sreenivasulu, 2021)	Impact of AI regulation and governance	AI	India	On the Internet	Y
2	(Banerjee et al., 2021)	Impact of AI on clinical education	AI	The UK	Healthcare	Y
3	(Abouzeid et al., 2021)	Role of Robotics and AI in Oral Health	Robotics and AI	Saudi Arabia	Healthcare	Y
4	(Crockett et al., 2020)	Public perceptions of AI applications	AI	The UK	No specific area	Y
5	(Kankanamge et al., 2021)	AI-driven disaster management	AI	Australia	Natural disaster	Y
6	(Yeh et al., 2021)	AI and Sustainable development	AI	Taiwan	ustainable developme-r	Y
7	(Lund et al., 2020)	Academic librarians' perceptions of AI	AI	International	Academic library	Y
8	(Yoon et al., 2021)	AI adoption in libraries	AI	The US and Canada	Academic library	Y
9	(Yüzbaşıoğlu, 2021)	Attitudes of dental students towards AI	AI	Turkey	Healthcare	Y
10	(Keser & PEKİNER, 2021)	Future of AI in Oral Radiology	AI	Turkey	Healthcare	Y
11	(Tang et al., 2021)	IoT for smart parcel locker logistics	IoT	China	Logistic	Y
12	(Jain, 2021)	IoT Medical Devices for Anemic Pregnant Women	IoT	India	Healthcare	Y
13	(Abdullah & Fakieh, 2020)	AI applications in healthcare employee	AI	Saudi Arabia	Healthcare	Y
14	(Khanagar et al., 2021)	AI in dental students' perceptions	AI	Saudi Arabia	Healthcare	Y
15	(Aggarwal et al., 2021)	Patient perceptions on AI in Health Care Data	AI	The UK	Healthcare	Y
16	(Hervieux & Wheatley, 2021)	Academic librarians' perceptions of AI	AI	Canada and US	Library	Y
17	(Leenhardt et al., 2021)	Perception and Expectations in Capsule Endoscopy	AI	Europe	Healthcare	Y
18	(Ganji & Parimi, 2021)	User perception of IoT healthcare devices	IoT	India	Healthcare	Y
19	(Lozano et al., 2021)	Perception of AI in Spain	Robots and AI	Spain	No specific area	Y
20	(Lai & Rau, 2021)	Public perception model of facial recognition	Facial recognition	China	No specific area	Y
21	(Chen et al., 2020)	Perception of Automated Vehicles	Automated vehicle	Taiwan	Transportation	Y
22	(Jiang & Cheng, 2021)	Robotic applications in public health	Robotics	China	Healthcare	Y
23	(Zhang & Dafoe, 2020)	US public opinion on AI governance	AI	The US	No specific area	Y
24	(Seng et al., 2021)	User perceptions of facial recognition	Facial recognition	The US and Canada	Not specific area	Y
25	(Cui & Wu, 2021)	Public perceptions of AI in China	AI	China	N specific area	Y
26	(Robb et al., 2020)	Public perception of robotics in danger zone	Robotics	The UK	In danger activities	N
27	(Kim et al., 2020)	Students' perceptions of AI teaching assistants	AI	The US	Education	N
28	(Kashive et al., 2020)	User perception of AI-enabled e-learning	AI	India	Education	N
29	(Kassens-Noor et al., 2021)	Public perceptions of an AI-mediated future	AI	The US	No specific area	N
30	(Antes et al., 2021)	Perceptions of healthcare technologies with AI	AI	The US	Healthcare	N
31	(Zhang & Yencha, 2022)	Perceptions towards hiring algorithms	AI	The US	Hiring	N
32	(Esmaeilzadeh et al., 2021)	Patients' perceptions toward AI in healthcare	AI	The US	Healthcare	N

Table 7: Papers using surveys as an instrument

Section 2 examines participants' knowledge of selected emerging technologies (AI, AVR, facial and voice recognition, IoT, and robotics) and their trusted information sources. Section 3 focuses on participants' perceptions of these technologies, rating benefits, challenges, and potential usage in various contexts. Section 4 explores participants' concerns regarding privacy, security, and data sharing with government agencies. In the final section, participants' awareness and satisfaction with regulations on these technologies are assessed. Overall, the questionnaire aims to understand public knowledge, perceptions, concerns, and governance preferences for emerging technologies.

To assess the questionnaire's effectiveness and gain initial insights into New Zealanders' perceptions of emerging technologies, we conducted a pilot survey, receiving 450 valid responses. The survey explored demographic information, participants' knowledge of the technologies, their attitudes toward the usage, and opinions on government regulations. The diverse participant pool provided valuable insights into prevailing sentiments and perceptions of these innovations.

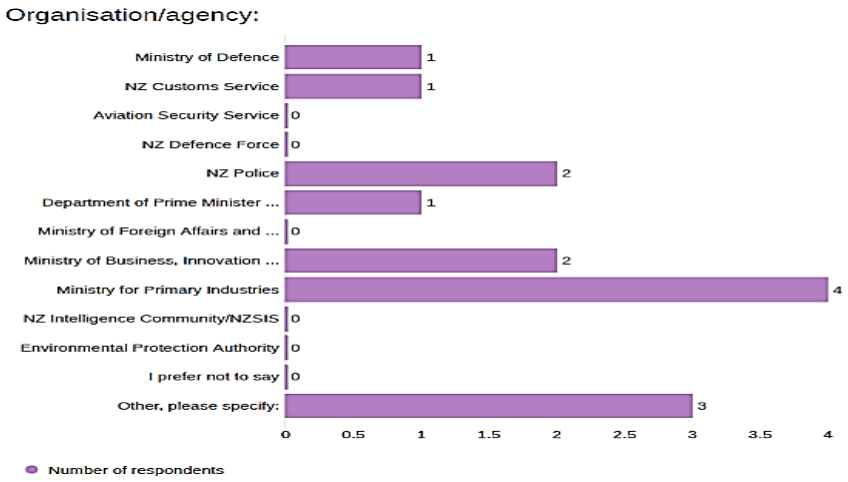

Fig 8: Number of respondents in the initial survey

4 Results and Discussion

In this section, we present an analysis of the survey responses gathered from the respondents. Before delving into the specific findings, it is important to understand the demographic background of the participants. We explore their age, education level, geographical distribution, ethnicity, and gender to gain insights into the representation of the NZ population in the survey. Additionally, we assess the respondents' knowledge about the surveyed emerging technologies and their perceptions regarding the use of these technologies in the country. The section aims to provide a comprehensive overview of the public's attitudes and awareness concerning emerging technologies in NZ, with a focus on data privacy, willingness to share data with government agencies, and awareness of relevant regulations. By understanding these aspects, we draw valuable conclusions and recommendations for future decision-making and policy development in this domain.

4.1 Demographic information

The survey respondents consisted 10.22% of 18-24 year olds, 18% of 25-34 year olds, 15.11% of 35-44 year olds, 19.11% of 45-54 year olds, 15.56% of 55-64 year olds, and a 22%, and the highest count of 65 years and older. The higher proportion of 65 years and olds may be due to the wider age range covered (up to 90 years and over) and potential survey reach restrictions. Moreover, the survey's composition aligns well with NZ's demographic makeup[6], suggesting no evident age bias found in the survey.

In terms of **educational background**, the highest number of people have a high school qualification (N=128, 28.44%), followed by a bachelor's degree (N=127, 28.22%). The lowest number of people have a PhD (N=10, 2.22%). Among the respondents, 96.65% possessed a high school education or higher, while 47.77% percent held a bachelor's degree or advanced academic qualifications. According to Education Counts (2021), the statistical results for 2021 indicated that 35% of NZ's entire population possessed a bachelor's degree or higher. The respondents in this survey displayed a slightly higher education level. However, considering that individuals with higher academic qualifications often have a better understanding of emerging technologies and are

[6] https://ecoprofile.infometrics.co.nz/new%20zealand/Population/AgeComposition

generally more willing to participate in such a survey, the educational distribution of respondents can be considered acceptable.

Regarding geographical distribution, the number of respondents from the Auckland region is 30.89% (N=139), from the Canterbury region 13.78% (N=62), and from the Wellington region 12.00% (N=54) respondents. The rest of the questionnaire respondents come from other smaller regions around NZ. According to stats.govt.nz, the population distribution in the survey aligns closely with the geographic distribution of NZ's population. Based on the survey responses the New Zealand European population constitutes the largest ethnic group (N=289), accounting for 64.22% of the total population. The Māori (N=52) and Asian (N=43) communities follow closely as the second and third-largest ethnic groups, respectively. According to stats.govt.nz, the population distribution in the survey aligns closely with the ethnic distribution of NZ's population. In terms of gender among the respondents, 51.33% identified as male (N=231), 48.22% as female (N=217), and there was an additional 0.44% who chose not to disclose their gender. This indicates a relatively even gender distribution, suggesting the survey results are unlikely to be biased due to gender-related factors.

4.2 Knowledge about surveyed emerging technologies

Regarding the overall knowledge level of emerging technologies (Fig 2), most respondents chose "moderately familiar" (N=692), followed by a similar proportion selecting "slightly familiar" (N=681). Over 80% of participants expressed at least "slight familiarity," while over 50% reported being "moderately familiar." Notably, 5% of respondents (N=118) stated they were "extremely familiar" with these technologies. When examining specific aspects, facial/voice recognition technologies were the most familiar, with over 90% of respondents (N=413) having at least "slight familiarity." AI followed closely, with 86% of respondents (N=389) reporting the same level of familiarity. Conversely, IoT and AVR were less known, with 123 and 121 respondents, respectively, admitting no understanding of these fields, accounting for over 25% of total respondents for each technology.

Regarding the familiarity of respondents with emerging technologies, the findings align with the previous question. In total, 72.98% of respondents are at least "slightly familiar," with 43.51% being "moderately familiar" or more. However, a notable number of individuals are entirely unfamiliar with IoT and AVR technologies, with 183 and 165 people, respecti
vely, falling into this category.

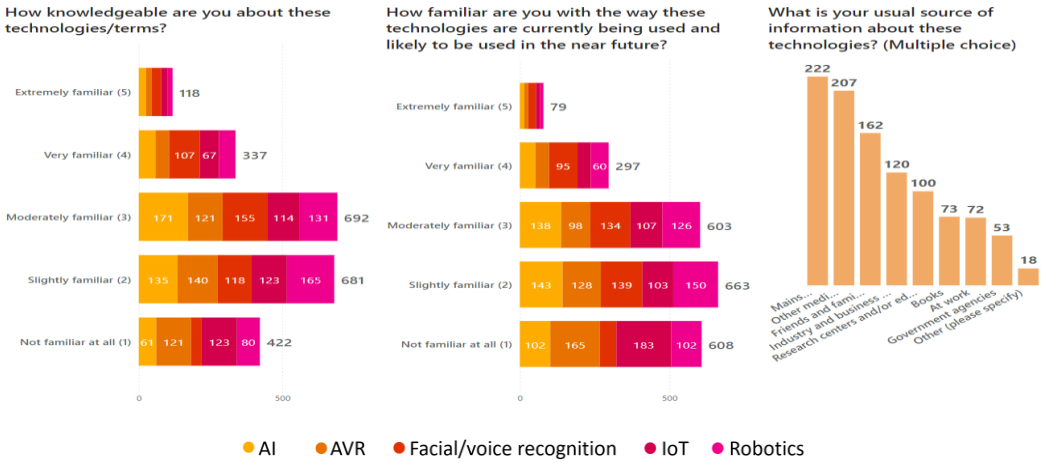

Fig 2: Left to right — How knowledgeable are you about these technologies? | How familiar are you with these technologies? | What is your usual source of information about these technologies?

Regarding the sources from which people obtain information about emerging technologies, the top two sources are mainstream media (N=222) and other media (N=207), encompassing newspapers, magazines, and social media, accounting for 41.77% of the total. Friends and family members, as well as industry and business websites, are also significant channels of information. In contrast, government agencies and other sources are less utilised, suggesting the need to enhance the government's role in information dissemination about emerging technologies.

4.3 Perceptions of the use of the surveyed technologies

4.3.1 Support Level

Regarding the support for the development and use of emerging technologies (Fig 3), combining five technologies, respondents showed the highest preference for "somewhat support" (N=712, 31.64%), while "strongly oppose" (N=142, 6.31%) received the least support. Meanwhile, over a quarter of respondents fell into the category of "neither support nor oppose" (N=606, 26.93%). In the breakdown of individual technologies, the results were relatively even, except for AI, which stood out with a higher level of support compared to other technologies. Approximately 231 individuals chose either "somewhat support" or "strongly support" for AI, accounting for over 50% of the respondents.

4.3.2 Concerns (Beneficial or Harm)

The survey also explored the perceptions of whether emerging technologies are beneficial or harmful to the public. Unlike the previous question, this one included 7 gradients ranging from "extremely harmful" to "extremely beneficial." Similarly, the results show a strong consistency with the previous question about support for these technologies. In the aggregate for all five technologies, the majority of respondents (N=590, 26.22%) chose "somewhat beneficial" as their perception. Following closely were those who maintained a neutral stance, considering the technologies neither beneficial nor harmful (N=513, 22.8%). The next group comprised those who believed the technologies to be beneficial (N=318, 14.13%). On the other hand, a total of 376 respondents (16.71%) expressed some degree of concern, indicating they see the technologies as somewhat harmful, harmful, or extremely harmful, with an additional 68 respondents (3.02%)

considering them to be extremely harmful. Regarding individual technologies, the percentages were relatively even across the board.

4.3.3 Areas to Develop and Use

The survey further inquired about the areas in which respondents believe the government can increase the use of technology. In this multiple-choice question, the top three areas selected by the respondents were healthcare (N=264), national security (N=254), and education (N=243). The concentration of responses to this question reflects the consistent expectations of the public for using technology to bring about changes in the fields of healthcare, national security, and education.

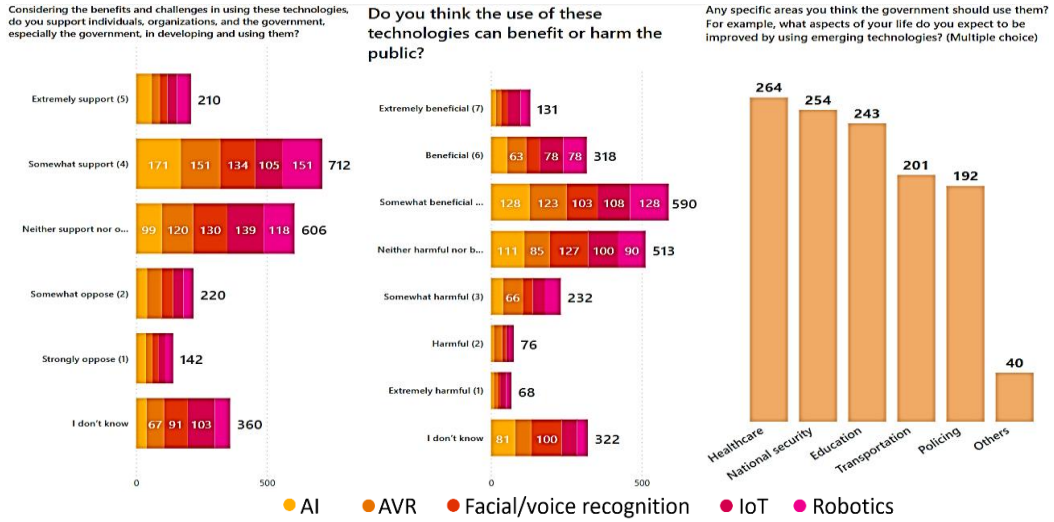

Fig 3: Left to right — Do you support individuals, organisations, and the government in developing and using them? | Do you think the use of these technologies can benefit or harm the public? | Any specific areas you think the government should use technologies?

4.4 Governance of technologies

The survey also asked respondents about the governance of emerging technologies in NZ. Fig 4 shows box and whisker plots to illustrate the results. The bottom and top edges of the dark and light blue box represent the first quartile (Q1) and third quartile (Q3), respectively. The line inside the box represents the median or the second quartile (Q2). The whiskers extend to the minimum and maximum values of the data, representing the range. In the middle of each box, the pink circles with numbers represent the mean values. The box and whisker plot provides a visual representation of the distribution of the data, indicating the spread, central tendency, and potential outliers.

4.4.1 Data Privacy

The survey on the importance of data privacy in the usage of five emerging technologies (Fig 4) reveals that people highly value data privacy, with average scores exceeding 75% and median scores above 80%. Furthermore, the third quartile for each technology is at 100%, indicating that at least 25% of the respondents selected the option of 100%, signifying extreme importance. Notably, facial/voice recognition and AI technologies stand out, with median scores around 90%. These findings indicate that the public places significant importance on information privacy and

security. As these emerging technologies become more prevalent, it is crucial to prioritise data privacy concerns. Failure to address these concerns adequately could lead to a loss of public support and trust in these technologies.

4.4.2 Willingness to Share Data

The next survey question asked respondents about their willingness to share data with government agencies, where 100% indicated being extremely willing. The results show that 25% of the respondents gave a highly positive response, with their choices falling between 80% and 100%. More than 50% of the respondents selected values above 60%, i.e. somewhat willing. Only less than 25% of the respondents' choices were below 50%, with a tendency to be unwilling to share data with government agencies. This statistical analysis suggests that although data privacy is a widespread concern among the public, government agencies are still generally trusted and accepted by the majority of people. To some extent, they are willing to disclose their personal information to the government.

4.4.3 Awareness of Regulations or Laws

The next question was used to assess the public's awareness of the regulations or laws regarding to emerging technologies. It is widely recognised that robust and well-defined legal regulations are crucial to ensuring that emerging technologies operate within reasonable and beneficial boundaries, benefiting humanity and future generations. However, the statistics reveal that 50% of the respondents have very little awareness about such legal regulations, less than 20%. Moreover, 25% among them fall below 5%, indicating almost complete lack of awareness. Only 25% of the respondents indicated an awareness level exceeding 50%.

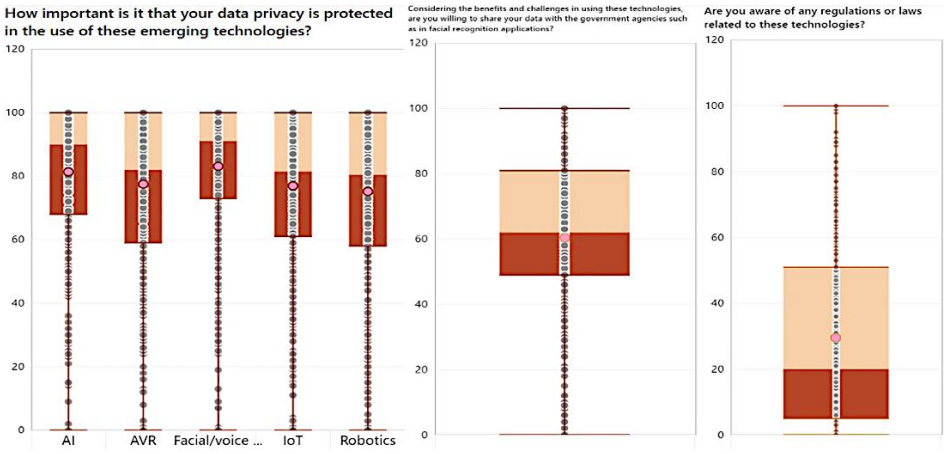

Fig 4: Perceptions of Data Privacy, Data Sharing, and Awareness of Regulations regarding Emerging Technologies.

This could be attributed to the incomplete nature of legal regulations concerning emerging technologies, or the limited public awareness of these regulations even if they exist. Establishing regulations or laws for emerging technologies is the responsibility of government agencies, and there is a long road ahead. It is essential for government agencies to work towards comprehensive legal frameworks and public awareness campaigns to bridge this gap and ensure that the public is well-informed about the legal aspects of emerging technologies. The actions respondents expect the government to take to regulate the use of these technologies

include assuring security, transparency, and privacy, and preventing misuse of such technologies.

4.5 Correlation Analysis of Knowledge with Support and Concerns

Fig 5 shows a positive correlation between knowledge of emerging technologies and the level of support. For instance, the more knowledge the public has about robotics, the higher their level of support for this technology (correlation coefficient=0.27). Additionally, higher knowledge of AI is associated with higher knowledge of AVR (correlation coefficient=0.72), and this is reflected in the increased level of acceptance for AVR (correlation coefficient=0.23). From these findings, we can conclude that individuals with higher knowledge of one technology tend to have increased knowledge of other technologies, leading to greater support for their development. Public education and awareness programmes are essential strategies to garner support for emerging technologies.

Similarly, Fig 6 shows a positive correlation between knowledge of emerging technologies and concerns. The more knowledge the public has about robotics, the more they perceive this technology as beneficial (correlation coefficient=0.18). This trend is observed for AI, AVR, facial/voice recognition, and IoT as well. Although the correlation coefficient for concerns is slightly lower than for support (0.27), strengthening public education remains crucial to addressing concerns and fostering an informed and supportive society.

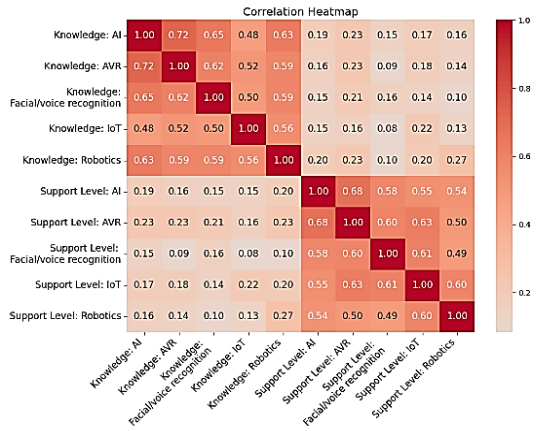

Fig 5: Correlation between knowledge of technologies and support level

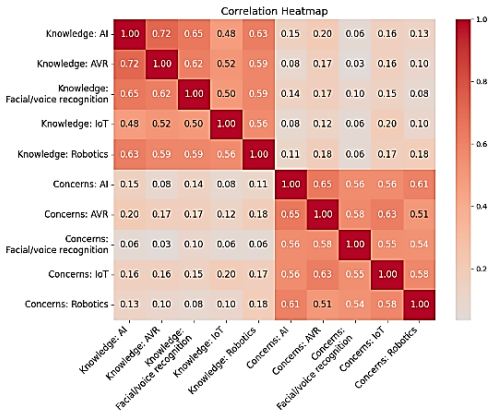

Fig 6: Correlation between knowledge of technologies and concerns (benefit or harm)

6 Conclusion and Recommendations for Future Work

This paper highlights the significance of understanding public perceptions of emerging technologies for their successful adoption and governance in New Zealand. The study includes a systematic literature review and the development of a comprehensive survey questionnaire to capture public viewpoints. Initial results from the pilot survey are also reported. The key findings of the survey concluded that emerging technologies, such as AI, VPAs, and immersive technologies, have the potential to significantly affect society and various sectors. To ensure responsible development and use of these technologies, it is crucial to consider public perceptions and concerns. The pilot survey indicates that the public generally supports the development of emerging technologies, but data privacy and biases are important issues that need attention.

The New Zealand government's approach of incorporating AI controls into existing policies to regulate technology usage is a step in the right direction. However, more comprehensive legal frameworks and public awareness campaigns are needed to bridge the gap in public awareness of regulations related to emerging technologies.

To further enhance understanding, future research should explore demographic and ethnicity-based reactions to emerging technologies. Additionally, aligning stakeholder perceptions with government strategies is vital for effective governance. Delving into various related issues, such as Industry 4.0, civic participation, and national defence, will provide valuable insights for shaping policies and addressing societal challenges.

In conclusion, a well-informed and supportive public is crucial for maximising the benefits and minimising the drawbacks of emerging technologies in New Zealand. By prioritising data privacy, transparency, and accountability, the country can harness the potential of these technologies to drive positive societal changes. Continued research and engagement with stakeholders will play a pivotal role in shaping a sustainable and inclusive future for emerging technologies in New Zealand.

Funding information

This project was funded by the Multi Agency Research Network (MARN), with additional support by the Massey University Research Fund (MURF). The study was conducted by the Management, Analytics, and Decision-making (MAD) research group

References

Block, G. (2019, December 5). Privacy concerns over police's new "state of the art" facial recognition system. Stuff. https://www.stuff.co.nz/national/117957684/privacy-concerns-over-polices-new-state-of-the-art-facial-recognition-system

Bulchand-Gidumal, J. (2020). Impact of Artificial Intelligence in Travel, Tourism, and Hospitality. Handbook of E-Tourism, 1–20. https://doi.org/10.1007/978-3-030-05324-6_110-1

Castello, K., & van der Meulen, R. (2018). Gartner identifies five emerging technology trends that will blur the lines between human and machine. Gartner Inc. https://www.gartner.com/en/newsroom/press-releases/2018-08-20-gartner-identifies-five-emerging-technology-trends-that-will-blur-the-lines-between-human-and-machine

Chatterjee, S., & Sreenivasulu, N. (2021). Impact of AI regulation and governance on online personal data sharing: from sociolegal, technology and policy perspective. Journal of Science and Technology Policy Management.

Chen, S.-Y., Kuo, H.-Y., & Lee, C. (2020). Preparing society for automated vehicles: Perceptions of the importance and urgency of emerging issues of governance, regulations, and wider impacts. Sustainability, 12(19), 7844.

Chen, X., & Chang-Richards, A. (2022). Technology implementation status and perceived benefits: A study of New Zealand construction organisations. IOP Conference Series: Earth and Environmental Science, 1–9.

Education Counts, E. (2021). Ministry of Education—Education Counts. Ministry of Education. https://www.educationcounts.govt.nz/home

Gavighan, C., Knott, A., Maclaurin, J., Zerilli, J., & Liddicoat, J. (2019). Government Use of

Artificial Intelligence in New Zealand (Commissioned Report for External Body). https://ourarchive.otago.ac.nz/bitstream/handle/10523/9372/NZLF report.pdf?sequence=1&isAllowed=y

Ikram, M. (2022). Transition toward green economy: Technological innovation's role in the fashion industry. Current Opinion in Green and Sustainable Chemistry, 37. https://doi.org/10.1016/j.cogsc.2022.100657

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. Stroke and Vascular Neurology, 2(4), 230–243.

Jiang, H., & Cheng, L. (2021). Public Perception and Reception of Robotic Applications in Public Health Emergencies Based on a Questionnaire Survey Conducted during COVID-19. International journal of environmental research and public health, 18(20), 10908.

Kashive, N., Powale, L., & Kashive, K. (2020). Understanding user perception toward artificial intelligence (AI) enabled e-learning. The International Journal of Information and Learning Technology.

Kathriarachchi , E., Alam, S., Weerasinghe, K., & Pauleen, D. (2024). Risks of e-commerce Recommender Systems: A Scoping Review. Australasian Journal of Information Systems, 28, https://doi.org/10.3127/ajis.v28.4869

Kim, J., Merrill, K., Xu, K., & 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), 1902-1911.

Kopalle, P. K., Gangwar, M., Kaplan, A., Ramachandran, D., Reinartz, W., & Rindfleisch, A. (2022). Examining artificial intelligence (AI) technologies in marketing via a global lens: Current trends and future research opportunities. International Journal of Research in Marketing, 39(2), 522–540.

Lai, X., & Rau, P.-L. P. (2021). Has facial recognition technology been misused? A public perception model of facial recognition scenarios. Computers in Human Behavior, 124, 106894.

Lynch, N., Campbell, L., Purshouse, J., & Betkier, M. (2022). Facial recognition technology in New Zealand: Towards a legal and ethical framework. IOP Conference Series: Earth and Environmental Science, Volume 1101, Technology and Innovation. /articles/report/Facial_Recognition_Technology_in_New_Zealand_Towards_a_Legal_and_Ethical_Framework/17204078/1

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data and Society, 3(2). https://doi.org/10.1177/2053951716679679

New Zealand Legislation. (1982). Official Information Act 1982. https://www.legislation.govt.nz/act/public/1982/0156/latest/DLM65628.html

New Zealand Legislation. (2020). Privacy Act 2020. https://www.legislation.govt.nz/act/public/2020/0031/latest/LMS23223.html

NZ Human Rights Commission. (2018). Privacy, Data and Technology: Human Rights Challenges in the Digital Age (Issue May). https://www.hrc.co.nz/files/5715/2575/3415/Privacy_Data_Technology_-_Human_Rights_Challenges_in_the_Digital_Age_FINAL.pdf

Robb, D. A., Ahmad, M. I., Tiseo, C., Aracri, S., McConnell, A. C., Page, V., Dondrup, C., Garcia, F. J. C., Nguyen, H. N., Pairet, E., Ramirez, P. A., Semwal, T., Taylor, H. M.,

Wilson, L. J., Lane, D., Hastie, H., & Lohan, K. (2020). Robots in the danger zone: Exploring public perception through engagement. ACM/IEEE International Conference on Human-Robot Interaction, 93–102. https://doi.org/10.1145/3319502.3374789

Rotolo, D., Hicks, D., & Martin, B. (2015). What is an emerging technology? SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2564094

Seng, S., Al-Ameen, M. N., & Wright, M. (2021). A first look into users' perceptions of facial recognition in the physical world. Computers & Security, 105, 102227.

Soleimani, M., Intezari, A. Pauleen, D. (2021). Mitigating cognitive biases in developing AI-assisted recruitment systems: A knowledge sharing approach. International Journal of Knowledge Management 18 (1)

Shufeng, Z. (2021). The application of facial recognition in China and India: Potential risks and regulation suggestions. Journal of US-China Public Administration, 18(3). https://doi.org/10.17265/1548-6591/2021.03.003

Stuart, C. (2020). How the NZ government will regulate AI. Computerworld. https://www.computerworld.com/article/3532505/how-the-nz-government-will-regulate-ai.html

Tatman, R. (2016). Google's speech recognition has a gender bias. Making Noise & Hearing Things. https://makingnoiseandhearingthings.com/2016/07/12/googles-speech-recognition-has-a-gender-bias/

West, S. M., Whittaker, M., & Crawford, K. (2019). Discriminating Systems: Gender, Race and Power in Artificial Intelligence. 1–33.

Yukich Clendon, O. M. M., Carpenter, J. K., Kelly, D., Timoti, P., Burns, B. R., Boswijk, G., & Monks, A. (2023). Global change explains reduced seeding in a widespread New Zealand tree: Indigenous Tūhoe knowledge informs mechanistic analysis. Frontiers in Forests and Global Change, 6. https://www.frontiersin.org/articles/10.3389/ffgc.2023.1172326

Zhang, B., & Dafoe, A. (2020). U.S. Public opinion on the governance of artificial intelligence. AIES 2020 - Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 187–193.

Zhang, L., & Yencha, C. (2022). Examining perceptions towards hiring algorithms. Technology in Society, 68, 101848.

12. PREDICTING LEARNING STYLES WITH AI: TOWARD ADAPTIVE AND PERSONALIZED EDUCATION

Gobinda Chandra Sarker
University of Dhaka
gcsarker001@gmail.com

Md Mehedi Hasan
University of Dhaka
mmhasan@du.ac.bd

Md. Rakibul Hoque
University of Dhaka
rakibul@du.ac.bd

M. Helal Uddin Ahmed
University of Dhaka
helal@du.ac.bd

Abstract

Artificial intelligence holds significant potential for enhancing adaptive learning environments. However, effective personalization requires a deep understanding of individual learner characteristics, particularly their preferred learning styles. This study presents an Artificial Neural Network (ANN) - based model, aligned with the VARK framework (Visual, Auditory, Reading/Writing, Kinesthetic), to identify student learning preferences using survey data collected from 700 students across schools, colleges, and universities in Bangladesh. A hybrid architecture combining multi-label classification and multi-output regression was employed to predict both the dominant learning styles and the degree of preference for each. The ANN outperformed traditional machine learning algorithms - including Support Vector Machine, Random Forest, Decision Tree, and K-Nearest Neighbors - achieving an F1-score of 0.92 and R^2 score of 0.96. Performance further improved with the integration of K-Means clustering, boosting the F1-score to 0.96. The regression component of the model provides a percentage-based prediction of how strongly a student prefers each learning style, offering a more granular and nuanced understanding of individual preferences. Compared to conventional approaches, this multiheaded approach is more flexible and informative, enabling the early identification of learning styles and facilitating the development of personalized educational content prior to course delivery.

Keywords
Learning Style; VARK; Clustering; Deep Learning; Artificial Neural Network

1. Introduction

In recent years, the education sector has seen a major shift with the integration of technology and data-driven practices. A key advancement is the use of ML techniques to predict students' learning styles (LS) (Essa et al., 2023). LS defines how individuals absorb and process information: some prefer visuals, others auditory input, reading, or hands-on activities (El-Saftawy et al., 2024). Accurate identification of these styles enhances learning outcomes. Machine Learning (ML) models can assess preferences from surveys, behavior, quiz scores, participation, and academic history (Ezzaim et al., 2024). This helps educators customize teaching methods and materials to each learner, improving engagement, comprehension, and retention (Raleiras et al., 2022). It also

supports adaptive systems that tailor lessons and resources to student needs, aiding struggling learners and accelerating advanced ones. As online and remote learning expands (Pardamean et al., 2022), accurate prediction of LS becomes even more critical to improving the effectiveness of digital education.

ML-driven LS prediction has become a key research area, with frameworks such as Fleming's VARK model (Fleming, 2001), which categorizes learners as Visual, Auditory, Read / Write or Kinesthetic, widely adopted for its simplicity and adaptability. The Honey and Mumford model (Honey and Mumford, 1986), common in corporate training, classifies learners as activists, reflectors, theorists, or practicists, although it lacks VARK's sensory focus. The Felder-Silverman Learning Styles Model (FSLSM) (Felder, 2002) adds depth with four dimensions: active-reflective, sensing-intuitive, visual-verbal, and sequential-global. VARK's emphasis on sensory modalities (Visual, Auditory, Reading/Writing, and Kinesthetic) aligns well with quantifiable behavioral indicators that can be captured through survey data or digital interaction logs, making it more compatible with ML models, particularly in classification tasks. Unlike the other complex cognitive or personality-based models, VARK provides a clear and intuitive structure that educators and learners can readily understand and apply, enhancing practical usability (Lee, 2025).

Recent studies have used ML models such as decision tree (DT), support vector machine (SVM), k-nearest neighbors (KNN), artificial neural network (ANN), and deep learning (Rasheed and Wahid, 2021; Rivas et al., 2021; Iatrellis et al., 2021) to predict LS from various data sources. These include learning management system (LMS) interactions, quiz scores, eye tracking data (Luo, 2021), keystroke patterns, and physiological signals (Dutsinma and Temdee, 2020). The extracted features provide objective, scalable insight into learning preferences. A popular method involves analyzing students' online behavior, clickstream data, content interaction time, and forum participation to infer LS (Sayed et al., 2024). NLP has also been applied to written responses and forum posts for cognitive analysis (Alqahtani et al., 2023). Another approach is survey-based prediction using models such as FSLSM, Kolb's theory (Kolb, 2014), or VARK, analyzed with factor analysis, clustering, or correlation. The survey approaches can be further enhanced by integrating behavioral analytics, such as interaction patterns and performance metrics from LMS, into ML pipelines to improve the accuracy of learning style predictions.

This study develops a predictive model to identify learning styles using survey data based on the VARK framework. VARK offers a clear and accessible classification suitable for both students and educators. Using responses from 700 students, we apply a multi-label classification model to identify each student's preferred learning styles, and a multi-output regression model to estimate the degree of their preference for each style. This hybrid modeling approach, combining classification and regression, overcomes the limitations of single-label classification by accommodating mixed learning style preferences commonly observed among students. It allows for a more nuanced and personalized educational experience. The results aim to inform adaptive teaching strategies, helping educators align instruction with individual learning needs. To support further research and development, the dataset, questionnaire, and source code have been made publicly available at: https://github.com/gcsarker/learning_style_prediction.git.

The key contributions of this research are as follows.

- This study integrates multilabel classification and regression techniques to provide a comprehensive understanding of learning preferences, allowing for the capture of multi-modal learning styles, which is a significant improvement over traditional methods.
- We developed the dataset by collecting survey responses from 700 college and university students in Bangladesh, ensuring the reliability and validity of the results.

The proposed system can be implemented at the beginning of a course to assist in curriculum design, without the need for course completion, in contrast to behavioral approaches commonly discussed in the literature.

2. Literature Review

Predicting students' learning styles (LS) is vital to improving education by enabling personalized, engaging, and effective learning experiences (Sayed et al., 2024). Since students process information differently, for example, visually, auditorily, kinesthetically, or through reading/writing, adapting instruction to these styles improves understanding and retention (Yotta, 2023). This not only improves academic results and motivation but also reduces frustration. It helps educators optimize teaching strategies, support diverse learners, and promote self-directed learning. In general, LS prediction fosters adaptive environments that support academic and lifelong success (Essa et al., 2023).

Several LS models have been proposed, including Kolb's experiential model (Sanjabi and Montazer, 2020), Felder-Silverman (FSLSM) (El-Bishouty et al., 2019), Honey and Mumford's approach (Lang, 2023), and the VARK model (El-Saftawy et al., 2024). Kolb's model classifies learners as Diverging, Assimilating, Converging, or Accommodating, although a lack of empirical validation limits it (Liu et al., 2024). FSLSM categorizes learners into four dimensions, such as Active or Reflective and Visual or Verbal (Masegosa et al., 2024). The VARK model, which divides learners into visual, auditory, read / write, and kinesthetic types, is favored for its simplicity and broad applicability (Jebbari et al., 2023). Therefore, we adopt the VARK model in this research.

Many studies have explored predicting students' learning preferences, often using data from digital learning platforms. Massive Open Online Courses (MOOCs) like Coursera and edX track metrics such as course completion, video watch time, and quiz scores, while LMS platforms like Moodle and Blackboard collect data on assignments, discussions, and resource usage. Agarwal (Agarwal et al., 2022) proposed a hybrid recommendation system in MOOCs that combines cluster-based collaborative filtering with Semantic Web Rule Language (SWRL), a language used to write logical rules that enhance the reasoning capabilities of ontologies. Learners are grouped based on FSLSM-derived preferences using behavioral data, and improvements in user experience were confirmed through logs, feedback, and analysis. Such platforms also gather data like video watch duration, quiz performance, forum activity, and clickstreams. User feedback, including surveys and content ratings, further refines predictions. ML models trained on these datasets can classify students into VARK categories, enabling tailored learning. Sayed (Sayed et al., 2024) used clickstream data and ML algorithms (KNN, RF, SVM, LR) with semantic associations to predict preferences, showing high accuracy on the OULAD (Open University Learning Analytics Dataset).

In another approach, Dutsinma and Temdee (Dutsinma and Temdee, 2020) integrated physiological data - such as heart rate and blood pressure - with the VARK model to predict learning styles (LS), achieving 85% - 90% accuracy using decision tree algorithms. Enhancing prediction further, Rao and Arunachalam (Rao et al., 2021) utilized ensemble techniques combining SVM, Naive Bayes, and Random Forest. Gambo (Gambo et al., 2022) explored deep learning by comparing CNN and MCNN architecture on facial image data, with MCNN proving more effective for VARK classification. Hasibuan (Hasibuan et al., 2019) employed Latent Semantic Indexing (LSI) alongside Artificial Neural Networks (ANN), emphasizing thesignificance of internal cognitive attributes over observable behaviors. Collectively, these studies highlight the recent research trend of using machine learning for advancing personalized education. Survey-based approaches to LS prediction, particularly those using the VARK framework, offer notable benefits. They overcome the need for personal interaction data and can be applied early in a course to facilitate instructional planning. For example, Gayathri (Gayathri et al., 2023) adopted a questionnaire-driven method using the FSLSM model, while Jebbari (Jebbari et al., 2023) assessed several machine learning algorithms - including NN, RF, DT, BN, ANN, and SVM - for VARK classification, identifying Neural Networks as the top performer with 98.7% accuracy in e-learning environments. These findings affirm the utility of VARK-based surveys in educational settings.

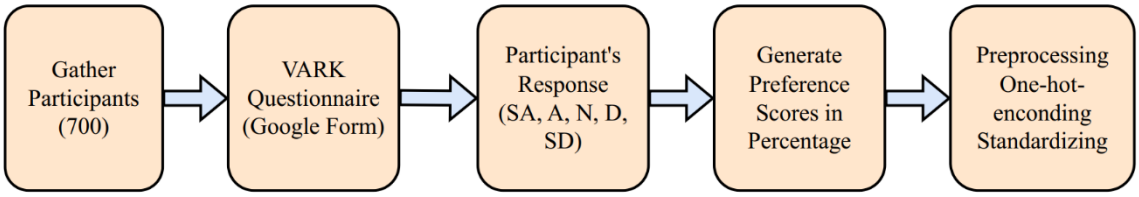

Fig. 1: Data Preprocessing steps.

While most LS studies rely on classification, it's essential to recognize that learners often exhibit multiple style preferences. To address this, Altamimi (Altamimi et al., 2022) introduced a regression-based model that assigns probabilistic scores to each learning style, offering a more nuanced view of learner tendencies. Building on this idea, our research proposes a hybrid strategy: employing multilabel classification to identify VARK styles while integrating regression analysis to quantify the intensity of each preference. As our model is survey-driven, it can be deployed before the semester begins, supporting the development of tailored learning resources.

3. Methodology

This section provides a comprehensive overview of the proposed student LS identification system. Fig. 1 illustrates the workflow diagram of this research.

3.1. Data Preparation

We collected data through a six months long survey (from February 2024 to July 2024) from 700 students (388 males, 312 females) across various schools and colleges in Bangladesh. Most of the students are from undergraduate and college level. The distribution of the students in terms of education level is provided in Fig. 2. A 28-question survey was used to identify students' learning preferences—Visual, Auditory, Reading/Writing, or Kinesthetic. The survey, distributed via Google Forms, asked students to rate their agreement with each statement on a five-point scale

(Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree). Each question corresponded to a specific learning style. Responses were scored as follows: Strongly Agree = 1.0, Agree = 0.75, Neutral = 0.5, Disagree = 0.25, and Strongly Disagree = 0.0. A student's preference for a learning style was determined by averaging the scores of the related questions:

$$P_{\text{style}} = \frac{\sum_{i=1}^{n} S_i}{n} \tag{1}$$

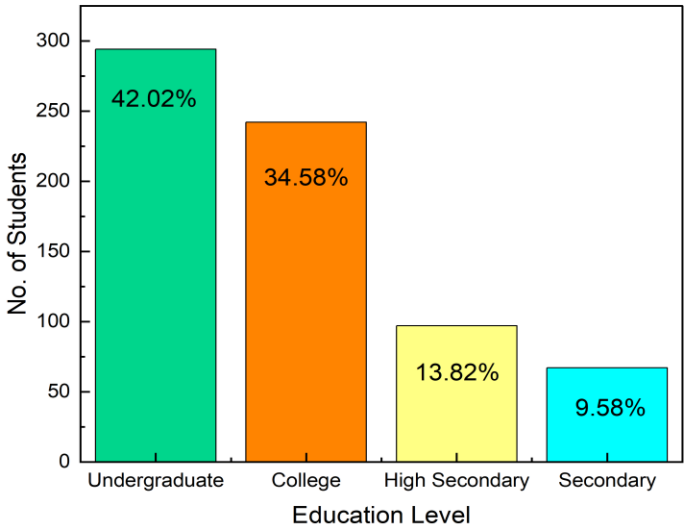

Fig. 2: Distribution of students from the survey in terms of education level.

Visual learners made up the largest group (280 students, 40%), followed by Reading/Writing (165 students, 23.5%) and Auditory (150 students, 21.4%). Kinesthetic learners were the smallest group (105 students, 15%). Finally, the score dataset is standardized using a mean and standard deviation of 1. Fig. 1 illustrates the steps for data preprocessing in our study.

3.2. K-means Clustering

The normalized score data is first processed using the K-means clustering algorithm, an unsupervised method that groups data into *k* clusters by minimizing within-cluster variance. Each point x_j is assigned to the nearest centroid μ_i based on Euclidean distance C_i as shown in Eq. 2. Centroids are then updated as the mean of all points in each cluster using Eq. 3. This process is repeated until the centroids no longer change, aiming to minimize the objective function as described in Eq. 4

$$C_i = \{x_j \mid \|x_j - \mu_i\| \leq \|x_j - \mu_m\|, \ \forall m \neq i\} \tag{2}$$

$$\mu_i = \frac{1}{|C_i|} \sum_{x_j \in C_i} x_j \tag{3}$$

$$J = \sum_{i=1}^{k} \sum_{x_j \in C_i} \|x_j - \mu_i\|^2 \tag{4}$$

To choose the optimal *k*, we use the elbow method in Eq. 5, which evaluates the within-cluster sum

of squares (WCSS). The point where WCSS reduction slows significantly indicates the optimal *k*. In our study the optimal number of K is 4. Once clustering is complete, the resulting cluster labels are appended to the original data and passed to the ANN model, enhancing its predictive performance.

$$WCSS(k) = \sum_{i=1}^{k} \sum_{x_j \in C_i} \|x_j - \mu_i\|^2 \tag{5}$$

3.3. Artificial Neural Network (ANN)

ANNs are ML models inspired by the human brain, commonly used for tasks like classification, regression, and pattern recognition. An ANN typically includes an input layer, one or more hidden layers, and an output layer. Each neuron computes a weighted sum of its inputs, adds a bias, and applies an activation function (f) as shown in Eq. 6.

$$y_j = f\left(\sum_{i=1}^{n} w_{ij} x_i + b_j\right) \tag{6}$$

Here, x_i are input features, w_{ij} are weights, b_j is the bias, and $f(\cdot)$ is a nonlinear activation function such as ReLU, sigmoid, or tanh. The network is trained using backpropagation, which adjusts the weights to minimize the error E. The update rule is given by Eq. 7. where η is the learning rate. This allows the network to learn from data and improve its predictions over time.

$$w_{ij} \leftarrow w_{ij} - \eta \frac{\partial E}{\partial w_{ij}} \tag{7}$$

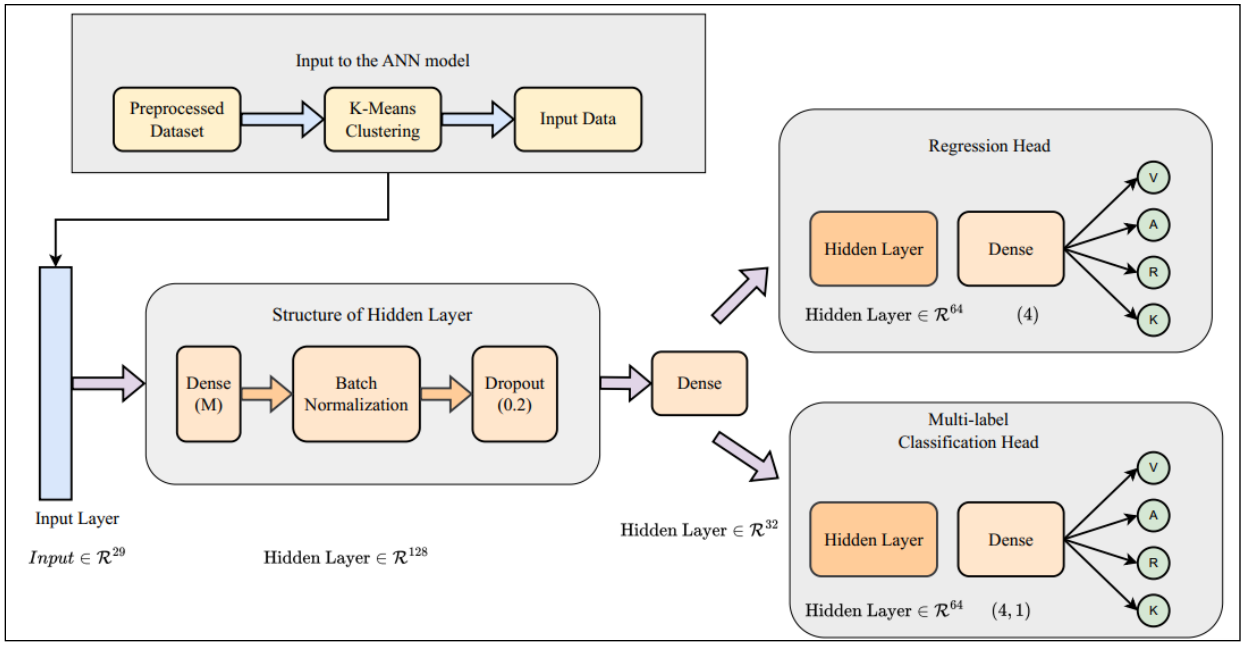

Fig. 3: Model architecture of the proposed system

3.4. The Proposed Model

The architecture of the proposed model is illustrated in Fig. 3. The process begins with a preprocessed dataset where each sample is standardized. K-means clustering is then applied to group the data into four clusters. The resulting cluster labels are appended to the original input features, forming a combined input vector of dimension R^{29}.

This combined input is passed to a neural network consisting of a shared backbone and two output heads. The input first goes through a dense layer of size *M*, followed by batch normalization and a dropout layer with a rate of 0.2. A subsequent dense layer reduces the feature space to R^{32}, which is then fed into two separate branches:

- **Regression Head:** This branch includes a hidden layer of size R^{64} followed by a dense layer that outputs four continuous values. These values represent the predicted preference scores for each learning style: Visual (V), Auditory (A), Reading/Writing (R), and Kinesthetic (K).
- **Multi-label Classification Head:** This branch also includes a hidden layer of size R^{64} followed by a dense layer with four sigmoid outputs, indicating the presence or absence of each learning style category in a response.

By combining both clustering and dual-task learning (classification and regression), the model effectively captures both categorical and continuous aspects of individual learning preferences.

4. Results and Analysis

In this study, a student's preference to learning style is identified into four categories from the VARK learning style theory. First, the k-means algorithm is used to cluster the survey response. We have applied an elbow approach to determine the best number of clusters. From Fig. 4, we can conclude that the number of optimal clusters is four. Then the K-means algorithm is applied. The visualize the result, the dimensions are reduced to 3 dimensions only using Principal component analysis, which is shown in Fig. 5. It can be noticed that the algorithm clearly separates the four groups. Although we can see some overlaps between the groups. It can be attributed to the fact that students sometimes prefer multiple learning styles. Fig. 6 shows the number of different categories in each cluster, highlighting its performance. The output cluster labels are then combined with the questionnaire response and sent to the ANN.

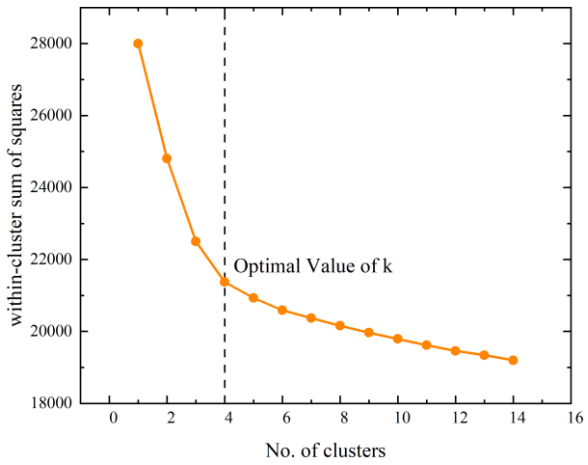

Fig. 4: Elbow plot to determine the optimal number of clusters.

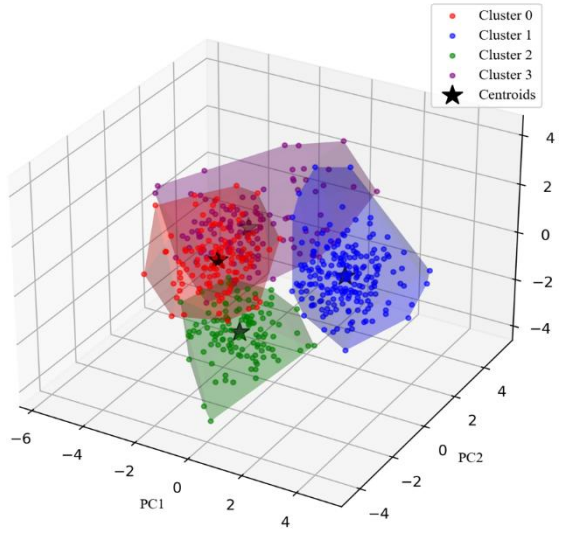

Fig. 5: Visualization of 4 clusters after dimensionality reduction using PCA.

Model	Precision	Recall	F1-Score	Hamming Loss
Support Vector Machine (SVM)	0.85	0.80	0.82	0.013
Decision Tree (DT)	0.80	0.78	0.79	0.018
Random Forest (RF)	0.88	0.84	0.86	0.010
K-nearest neighbor (KNN)	0.83	0.79	0.81	0.015
ANN	0.92	0.90	0.92	0.007
Kmeans + ANN	**0.96**	**0.93**	**0.95**	**0.005**

Table 1: Performance Comparison of Classification Approach

Four machine learning models (SVM, DT, RF, and KNN) were compared with ANN. The findings show that ANN outperforms the other models both in terms of classification and regression. The performance of the classification models is summarized in Table 1 based on four key evaluation metrics: Precision, Recall, F1-Score, and Hamming Loss. Precision measures the proportion of correctly predicted positive cases out of all predicted positive cases. A higher precision value indicates fewer false positives. Recall represents the proportion of correctly identified positive cases out of all actual positive cases. Higher recall suggests fewer false negatives. F1-Score is the harmonic mean of precision and recall values, providing a balanced measure of a model's accuracy. A higher F1-score indicates a better trade-off between precision and recall. Since we have applied multilabel classification, Hamming Loss quantifies the fraction of incorrectly classified instances. Lower values indicate fewer misclassifications and, thus, better model performance. Among the models, ANN demonstrated superior performance with precision (0.92), recall (0.90), and F1-score (0.91), achieving the lowest Hamming Loss (0.007). This indicates that ANN provides more accurate predictions with minimal errors compared to the other ML classifiers. On the other hand, Random Forest also performed well with an F1- score of 0.86, followed by SVM (0.82) and KNN (0.81). The Decision Tree model yielded the lowest classification performance, with an F1-score of 0.79 and the highest Hamming Loss (0.018).

Moreover, integrating cluster labels to the ANN input increases the precision (0.96), recall (0.93) and f1-score (0.94). On average the performance is increased by 4% after using KMeans, since the model receives some insights on the different categories from the labels. These findings reinforce the effectiveness of clustering with ANN in classifying students' learning preferences with higher accuracy, making it a promising approach for personalized learning recommendations.

Models	**R^2**				**MAPE**			
	V	A	R	K	V	A	R	K
SVR	0.955	0.944	0.937	0.943	0.047	0.047	0.070	0.038
DT	0.921	0.913	0.902	0.910	0.065	0.069	0.085	0.058
RF	0.962	0.954	0.946	0.951	0.042	0.045	0.062	0.035
KNN	0.940	0.930	0.925	0.933	0.055	0.058	0.073	0.045
ANN	0.975	0.968	0.960	0.965	0.035	0.038	0.050	0.030
Kmeans + ANN	**0.985**	**0.976**	**0.970**	**0.975**	**0.030**	**0.032**	**0.051**	**0.029**

Table 2: Performance Comparison of Regression Approaches

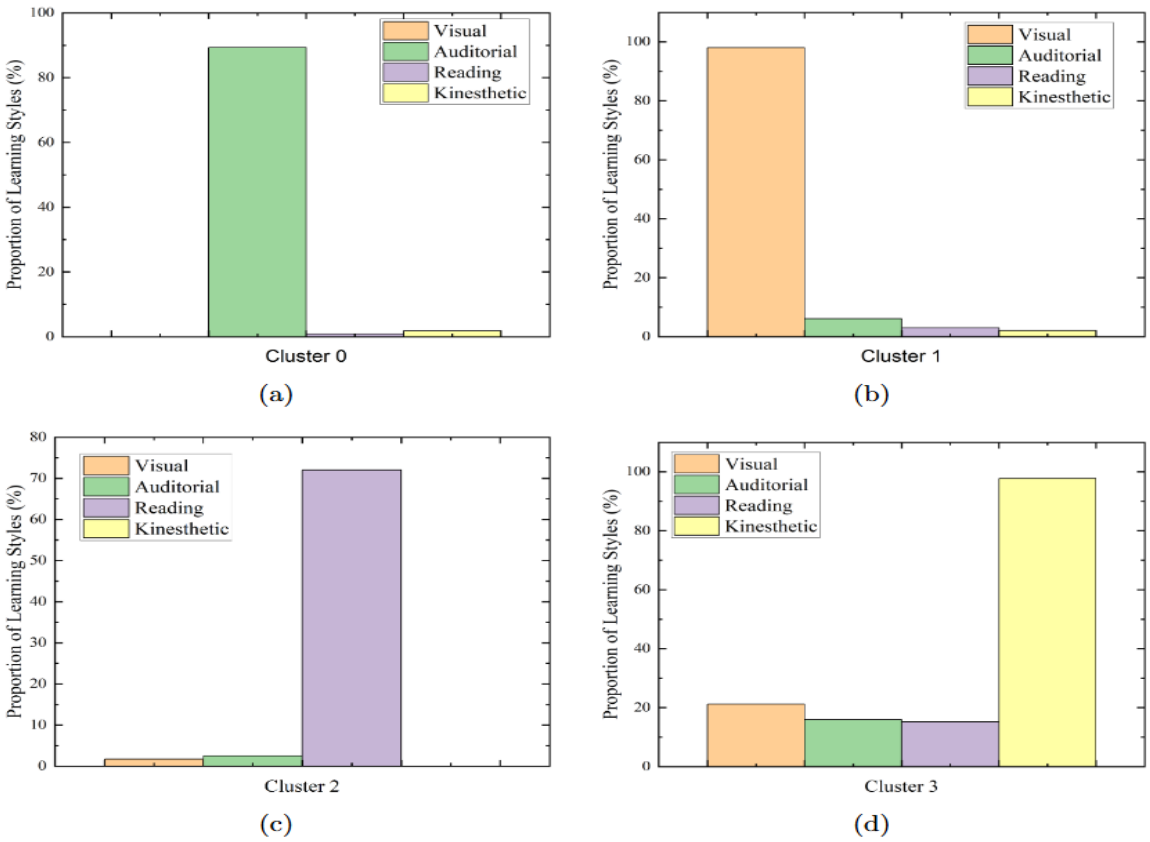

Fig. 6: Bar charts for learning styles percentages in each cluster.

The confusion matrix for the proposed model is provided in Fig. 7. Since we are doing multilabel classification, each category in (V, A, R and K) is compared with rest of categories in one vs. rest manner. The figure shows the model achieves very high true positives (TN) and true negatives (TN) on test set.

The performance comparison of regression analysis is presented in Table 2. This table evaluates five ML models (SVR, DT, RF, KNN, and ANN)—using two key regression metrics: R^2 (coefficient of determination) and Mean Absolute Percentage Error (MAPE). R^2 measures how well the model explains the variance in the target variable. A value closer to 1 indicates a better fit, meaning the model can accurately predict student learning preferences. MAPE quantifies the average percentage error between predicted and actual values. A lower MAPE indicates higher accuracy, as it means the model makes smaller relative errors. Among the models, the regression output head from the proposed model achieved the highest R^2 values across all test cases (0.975 to 0.965), indicating its strong ability to explain variance in student learning preferences. It also recorded the lowest RMSE (0.030 to 0.038) and MAPE (0.018 to 0.027), demonstrating its superior predictive accuracy with minimal error rates. In contrast, Random Forest (RF) performed well with an R^2 close to 0.95 and relatively low RMSE and MAPE values, making it the second-best model. Meanwhile, SVR and KNN exhibited moderate performance, whereas DT had the lowest R^2 scores and highest error rates, suggesting weaker predictive accuracy

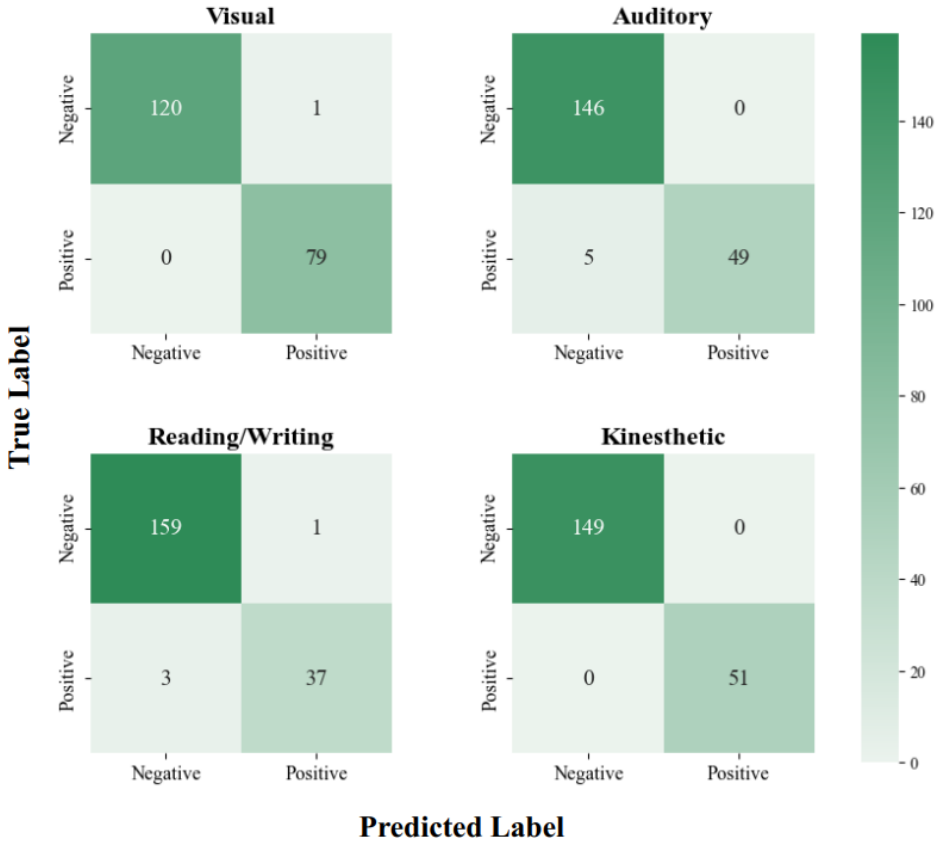

Fig. 7: Confusion matrix of the proposed ANN model

.

5 Conclusion

Understanding individual learning styles is essential for delivering personalized education. This study introduces a novel learning style prediction model that combines K-Means clustering with a multi-headed Artificial Neural Network, leveraging responses from a VARK-based questionnaire. Initially, K-Means clustering groups individuals based on their learning behavior, and the resulting cluster labels are subsequently fed into the ANN. The ANN performs multi-label classification to predict learning styles and simultaneously estimates a percentage-based distribution of learning preferences. The proposed model significantly outperforms traditional machine learning approaches such as Decision Trees, Support Vector Machines, and Random Forests, achieving an F1-score of 0.95 for classification and an R^2 score of 0.98 for regression. Incorporating cluster labels enhances the backbone ANN's performance by 4%, demonstrating the effectiveness of integrating unsupervised and supervised learning. Unlike conventional single-label classifiers, the proposed AI solution provides more detailed insights into individual learning preferences, offering a richer and more accurate representation of how learners engage with different educational modalities. This enhanced level of detail makes the model particularly applicable in real-life educational settings, such as personalized e-learning platforms, adaptive tutoring systems, and curriculum design, where understanding the unique combination of learning styles can lead to more effective content delivery, increased student engagement, and improved learning outcomes. **a** key limitation of this study is that the survey was primarily conducted among students from urban areas, particularly Dhaka and other major cities, due to the accessibility of online data collection platforms. As a result, the sample may over-represent students with better internet access and digital literacy. In future, we aim to include students from rural and underserved regions and examine learning style variations across different socioeconomic and geographic groups. Future work can also focus on integrating Explainable AI (XAI) techniques to improve model interpretability, developing real-time adaptive learning systems, and incorporating Natural Language Processing (NLP) to analyze open-ended questionnaire responses. Expanding the framework to support multimodal learning environments could further enhance its ability to tailor education to individual needs.

Acknowledgements

This research is supported by the Bangladesh Bureau of Educational Information and Statistics (BANBEIS), Ministry of Education, Bangladesh (Grant No. 37.20.0000.004.033.020.2016.2233). The authors, therefore, gratefully acknowledge the BANBEIS technical and financial support.

References

Agarwal, A., Mishra, D. S., and Kolekar, S. V. (2022). Knowledge-based recommendation system using semantic web rules based on learning styles for moocs. Cogent Engineering, 9(1):2022568.

Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., bin Saleh, K., Alowais, S. A., Alshaya, O. A., Rahman, I., Al Yami, M. S., et al. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. Research in Social and Administrative Pharmacy, 19(8):1236–1242.

Altamimi, A. M., Azzeh, M., and Albashayreh, M. (2022). Predicting students' learning styles using regression techniques. arXiv preprint arXiv:2209.12691.

Dutsinma, L. I. F. and Temdee, P. (2020). Vark learning style classification using decision tree with physiological signals. Wireless Personal Communications, 115(4):2875–2896.

El-Bishouty, M. M., Aldraiweesh, A., Alturki, U., Tortorella, R., Yang, J., Chang, T.-W., Graf, S., et al. (2019). Use of felder and silverman learning style model for online course design. Educational Technology Research and Development, 67(1):161–177.

El-Saftawy, E., Latif, A. A. A., ShamsEldeen, A. M., Alghamdi, M. A., Mahfoz, A. M., and Aboulhoda, B. E. (2024). Influence of applying vark learning styles on enhancing teaching skills: application of learning theories. BMC Medical Education, 24(1):1034.

Essa, S. G., Celik, T., and Human-Hendricks, N. E. (2023). Personalized adaptive learning technologies based on machine learning techniques to identify learning styles: A systematic literature review. IEEE Access, 11:48392–48409.

Ezzaim, A., Dahbi, A., Aqqal, A., and Haidine, A. (2024). Ai-based learning style detection inbadaptive learning systems: a systematic literature review. Journal of Computers in Education, pages 1–39.

Felder, R. M. (2002). Learning and teaching styles in engineering education. Fleming, N. (2001). Teaching and learning styles: Vark strategies. ND Fleming.

Gambo, F., Wajiga, G. M., Shuib, L., Garba, E. J., Abdullahi, A. A., and Bisandu, D. B. (2022). Performance comparison of convolutional and multiclass neural network for learning style detection from facial images. EAI Endorsed Trans. Scalable Inf. Syst., 9(35):e1.

Gayathri, G., Begam, M. F., and Aashika, G. (2023). Machine learning based learning style detection using felder-silverman framework. In 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), pages 1–7. IEEE.

Hasibuan, M. S., Nugroho, L. E., and Santosa, P. I. (2019). Model detecting learning styles with artificial neural network. Journal of Technology and Science Education, 9(1):85–95.

Honey, P. and Mumford, A. (1986). The manual of learning styles. Peter Honey.

Iatrellis, O., Savvas, I. K., Fitsilis, P., and Gerogiannis, V. C. (2021). A two-phase machine learning approach for predicting student outcomes. Education and Information Technologies, 26:69–88.

Jebbari, M., Cherradi, B., Moutaouakil, W., El Gannour, O., Hamida, S., and Raihani, A. (2023). Prediction and classification of learning styles using machine learning approach. In 2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), pages 1–7. IEEE.

Kolb, D. A. (2014). Experiential learning: Experience as the source of learning and development. FT press.

Lang, M. (2023). Learning styles and on-line learning analytics: An analysis of student behaviour based on the honey and mumford model. In International Conference on Human-Computer Interaction, pages 154–166. Springer.

Luo, Z. (2021). Using eye-tracking technology to identify learning styles: Behaviour patterns and identification accuracy. Education and information technologies, 26(4):4457–4485.

Pardamean, B., Suparyanto, T., Cenggoro, T. W., Sudigyo, D., and Anugrahana, A. (2022). Ai-based learning style prediction in online learning for primary education. IEEE Access, 10:35725–35735.

Raleiras, M., Nabizadeh, A. H., and Costa, F. A. (2022). Automatic learning styles prediction: A survey of the state-of-the-art (2006–2021). Journal of Computers in Education, 9(4):587–679.

Rao, C. S. et al. (2021). Personalization of learning management system using vark. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(7):1538–1545.

Rasheed, F. and Wahid, A. (2021). Learning style detection in e-learning systems using machine learning techniques. Expert Systems with Applications, 174:114774.

Rivas, A., Gonzalez-Briones, A., Hernandez, G., Prieto, J., and Chamoso, P. (2021). Artificial neural network analysis of the academic performance of students in virtual learning environ- ments. Neurocomputing, 423:713–720.

Sanjabi, T. and Montazer, G. A. (2020). Personalization of e-learning environment using the kolb's learning style model. In 2020 6th International conference on web research (ICWR), pages 89–92. IEEE.

Sayed, A. R., Khafagy, M. H., Ali, M., and Mohamed, M. H. (2024). Predict student learning styles and suitable assessment methods using click stream. Egyptian Informatics Journal, 26:100469.

Yotta, E. G. (2023). Accommodating students' learning styles differences in english language classroom. Heliyon, 9(6).

Lee, S. (2025). Educational Implications of VARK Learning Styles: Academic Performance and Pedagogical Preferences among Korean Pharmacy Students. Ind. J. Pharm. Edu. Res, 59(2), 512-517.

Liu, P., Yang, Z., Huang, J., & Wang, T. K. (2024). The effect of augmented reality applied to learning process with different learning styles in structural engineering education. Engineering, Construction and Architectural Management.

Masegosa, A. R., Cabañas, R., Maldonado, A. D., & Morales, M. (2024). Learning styles impact students' perceptions on active learning methodologies: a case study on the use of live coding and short programming exercises. Education Sciences, 14(3), 250.

13. RECOMMENDATIONS FOR TELECONSULTATION IMPLEMENTATION FOR HEALTHCARE PROVIDERS: A SYSTEMATIC LITERATURE REVIEW

Denise Selina Y. Tan
Auckland University of Technology
denise.tan@aut.ac.nz

Felix B. Tan
Auckland University of Technology
felix.tan@aut.ac.nz

Abstract

The COVID-19 pandemic has significantly accelerated the adoption of teleconsultation. This study aims to conduct a systematic review of recent research to distil recommendations for implementing teleconsultation service from the perspective of healthcare providers. Utilizing the Technology-Organization-Environment (TOE) framework, the research identifies technological, organizational, and environmental factors that are critical to its successful implementation. Key findings emphasize the importance of reliable technical infrastructure, comprehensive IT support, strategic partnerships, and clearly defined guidelines. By addressing these dimensions, healthcare providers can have a foundation to integrate teleconsultation into routine practices and ensure its sustainability beyond the pandemic.

Keynotes

teleconsultation, telemedicine, healthcare provider, TOE framework

1 Introduction and Background

The COVID-19 pandemic has ushered the digital transformation of healthcare delivery systems, with telemedicine emerging as a pivotal innovation in ensuring continuity of care during times of restricted physical interaction (Barbosa da Silva et al., 2024). In the post-COVID landscape, telemedicine's significance has only deepened, offering solutions that enhance accessibility, reduce healthcare costs, and improve patient outcomes (Nittas & von Wyl, 2020)

Telemedicine, as defined by the Federal Communications Commission (2024), encompasses a broad spectrum of healthcare interventions facilitated by telecommunications technologies including telesurgery, telemonitoring, and teleconsultation. Among these, teleconsultation—the remote interaction between healthcare providers and patients via information and communication technology—has evolved from being a complementary service to a fundamental component of modern healthcare (Iliyasu et al., 2024).

Despite teleconsultation's increasing significance, the implementation of teleconsultation remains uneven (Fernández Coves et al., 2022; Hassan et al., 2020). Certain populations and healthcare contexts, such as chronic care management, have embraced teleconsultation more readily due to clear utility and benefits (Dhunnoo et al., 2024). Conversely, its uptake in general clinical practice has been met with resistance (Nagarathinam et al., 2024).

While numerous reviews have explored barriers and enablers to telemedicine implementation, many of these studies aggregate findings across multiple forms of telemedicine, obscuring the unique challenges and opportunities associated with teleconsultation (Aldawoud et al., 2024; Oudbier et al., 2024). Among studies focusing on teleconsultation, the implementation factors are often narrowly tailored to specific contexts, such as chronic disease management (Senyel et al., 2024), or asynchronous teleconsultations (Osman et al., 2019). This narrow focus limits the generalizability of findings to broader healthcare settings.

Furthermore, some studies posit that the perspective of healthcare providers remain underrepresented (Grata-Borkowska et al., 2022; Oudbier et al., 2024). Healthcare providers play a pivotal role in ensuring the relevance, effectiveness, and sustainability of teleconsultation services, particularly in the post-COVID era (Garavand et al., 2022). Their acceptance and engagement of teleconsultation are essential for its long-term success and its integration into routine clinical practices.

Diel et al. (2023) observed that practical recommendations offered by many existing studies on teleconsultation adoption are mostly centered on the patient's perspectives, highlighting the need for further studies that examine physician's acceptance variables in an organizational context. Although some studies examining technology implementation for healthcare professionals exist, most were conducted before 2019 (Gagnon et al., 2012; Ross et al., 2016), potentially limiting their applicability in the post-pandemic healthcare environment.

This study is a step towards addressing this research gap by conducting a systematic review of teleconsultation implementation factors from the healthcare provider's perspective. The central research question guiding this inquiry is: *What are the current recommendations for healthcare providers in implementing teleconsultation services based on contemporary research?*

This paper focuses specifically on implementation. Drawing from the field of implementation science, adoption represents the decision to use a service or practice, while implementation encompasses the deliberate actions, strategies, and processes required to integrate the service—in this case, teleconsultation—effectively (Allen et al., 2017; Hausman & Stock, 2003). By examining implementation factors rather than mere adoption decisions, this research addresses the critical gap between intending to use teleconsultation and successfully operationalizing it in clinical practice. It aims to provide a starting point for healthcare providers wanting to adopt teleconsultation to understand and navigate teleconsultation implementation effectively.

The next section will describe the methods used to conduct the systematic literature review. Section 3 will describe the data analysis approach using the Technology-Organization-Environment framework (Tornatzky et al., 1990) and the study findings. Section 4 will discuss the results and recommendations, and Section 6 will conclude the paper.

2 Research Design

To address the research question, a systematic literature review was conducted, incorporating select elements of the PRISMA 2020 search strategy. While PRISMA is widely regarded as the gold standard for systematic reviews (Pussegoda et al., 2017), according to Teixeira da Silva and Daly (2024), strict adherence to its protocol can sometimes result in the over-exclusion of potentially relevant publications. Consequently, this study adapted the screening process of the PRISMA 2020 framework, to maintain rigor while ensuring study inclusivity.

2.1 Study Identification

Although there is no consensus on the optimal number of databases required for systematic reviews, multiple studies suggest that searching at least two large, well-curated databases is essential for a reliable literature review (Ewald et al., 2022). Furthermore, the suitability of selected databases often holds greater significance than mere database coverage (Ewald et al., 2022; Harari et al., 2020).

This study conducted a literature search across three electronic databases: Medline, CINAHL Complete, and ScienceDirect. Database searches on MEDLINE and CINAHL complete were done through EBSCOhost while ScienceDirect database search was done on their website. Keywords used to gather relevant literature are listed on Table 8.

Database	Query
MEDLINE and CINAHL Complete Via EBSCOhost	In Title: Telemedicine or telehealth or teleconsultation or teleconsulting or "remote consultation" or "video consultation" or "online consulting" AND In Abstract: recommendation or adoption or implementation or determinants or enablers AND In Abstract: facilitators or barriers or obstacles or challenges AND In Abstract: hospital or healthcare provider or clinics or institution or physician or doctor or clinician
ScienceDirect	Telemedicine or telehealth or teleconsultation or remote consult or video consult AND adoption or implementation or barrier or challenge AND provider or physician or hospital or clinic

Table 8: Keywords used for study identification

These databases were chosen based on their relevance. Medline is frequently cited as a primary database for healthcare research and is known for its wide coverage and ability to yield the most relevant results in this domain (Hartling et al., 2016; Mathew, 2024). CINAHL Complete was selected due to its extensive collection of qualitative healthcare research, particularly studies focusing on healthcare professionals and their practices rather than medical interventions (Wright et al., 2015). ScienceDirect was chosen as a supplementary database due to van Dinter et al. (2021) study that found it to be a rich source of primary studies on topics related to system innovation, such as telemedicine, aligning closely with this study's focus.

2.2 Study Selection and Screening

To ensure relevance and timeliness, the inclusion criteria presented in Table 9 were applied during the initial database search.

Inclusion Criteria	Definition
C1: Publication date	Only studies published between 2020 and 2024 were included to reflect the influence of COVID-19 on telemedicine's advancement
C2: Publication type	Only peer-reviewed studies published in academic journals or conferences were included. Publications in books, reports, or theses were excluded.
C3: Language	Only studies published in English were included

Table 9: Initial database search criteria

The initial database search returned 538 studies: 376 from Medline, 122 from CINAHL Complete, and 40 from ScienceDirect. These RIS files were imported into EndNote for further filtering using the exclusion criteria listed in Table 10.

Exclusion Criteria	Definition
C4: Duplicates removal	Records that point to the same studies as previous record are eliminated
C5: Non-empirical work	Literature and systematic review papers will be discarded to avoid accumulated interpretation bias (Baldwin et al., 2022)

Table 10: Study dataset clean-up exclusion criteria

A total of 96 duplicate records were found primarily by using EndNote's "Find Duplicate" function. A few additional duplicates were manually identified and removed leaving behind 442 unique publications. Review articles were excluded by searching the titles and abstracts for terms such as "systematic review," "PRISMA," and "scoping." A few review articles that were caught after this stage were also manually excluded and added to the total count.

This step eliminated 64 review articles, leaving 378 empirical papers for further screening. The titles and abstracts were manually read to further weed out irrelevant studies. Table 11 lists the criteria applied to publication titles and the abstract screening process.

Inclusion Criteria	Definition
C6: Focus on Healthcare Providers	Only studies that centered on the perspectives of healthcare providers (e.g., clinicians, nurses, hospital managers) were included. Studies that primarily focused on patient perspectives were excluded.
C7: Scope	Only studies that primarily refer to teleconsultation were included. Publications on other telemedicine applications (e.g., remote prescriptions, medical training, mobile apps) were excluded.
C8: Implementation	Studies had to address telemedicine implementation. Publications that are focused on telemedicine acceptance, utilization and effectiveness were removed.

Table 11: Titles and abstract screening criteria

It was observed during this stage that terms like telemedicine, and telehealth are often used interchangeably to refer to teleconsultation. To ensure inclusivity, criteria C7 screening allowed publication titles and abstracts that contain telemedicine and telehealth as long as the abstract also contains the keywords "consultation," "conference," and "e-consult". Criteria C8 ensures that the

included studies involved the adoption or implementation of teleconsultation into practice. Thus, studies that did not provide any inputs on adoption measures such as those that focused on measuring teleconsultation acceptance and effectiveness were removed. Through this process, 327 publications were eliminated, leaving 51 studies for full-text retrieval.

2.3 Full-Text Review and Final Selection

Study retrieval stage checks whether the full-text copies are available and retrievable. This stage identified 3 studies that were not fully accessible to the authors, leaving 48 studies for eligibility assessment. The eligibility assessment consists of reading the remaining papers in its entirety to determine its relevance to this research. Table 12 lists the inclusion criteria that guided the eligibility screening.

Due to the less stringent application of criteria C7, an additional 8 papers were removed during this stage. Overall, this process excluded 27 studies, resulting in 21 papers for review.

Inclusion Criteria	Definition
C9: Context	Studies had to focus on teleconsultation between healthcare providers and patients. Physician-to-physician teleconsultation studies were excluded.
C10: Recommendations	Publications had to provide recommendations for teleconsultation implementation. Studies that merely identified barriers or challenges without offering solutions were excluded. Recommendations that are only applicable to specific contextual settings were also removed

Table 12: Study eligibility assessment criteria

2.4 Characteristics of included studies

The selected studies spanned several geographic regions. North America contributed the largest number of studies a total of 10 papers. Australia followed with 5 studies. Asia was represented by studies from Hong Kong, Pakistan, Iran, and Kuwait, while individual studies from Europe (Italy), South America (Brazil) and Africa (Morocco) were also included. Looking at the publication year, the majority of studies were published in 2023 (8 studies), followed by 2024 (6 studies) and 2022 (5 studies). Fewer studies were published in 2020 (2 studies). This could indicate a temporary spike in telemedicine research due to the COVID-19 pandemic.

Physicians and clinicians are the most represented group (16 studies). Healthcare providers such as nurses, physiotherapists and staff were included in 6 studies. Only 3 studies included healthcare administrators, such as C-suite officers, practice managers and policymakers. Majority of the included studies employed qualitative methodologies, with semi-structured interviews used in 12 studies. Qualitative surveys were conducted in 8 studies. Other methods were used less frequently.

3 Results

3.1 Data Analysis using the TOE Framework

TOE framework is selected to structure the data analysis. Developed by Tornatzky et al. (1990), the Technology-Organization-Environment (TOE) framework is a well-established innovation adoption model that highlights three interrelated adoption dimensions - the technological context,

the organizational context, and the environmental context. Technological context refers to the tools and innovations that are relevant and accessible to the firm; organizational context points to the attributes of the firm such as the size, structure, and management while the environmental context looks at the external factors that nevertheless influence the firm's capability to adopt an innovation (Baker, 2012). The TOE framework is particularly effective in analyzing intra-firm innovation adoption, especially in cases where external influences play a pivotal role (Oliveira & Martins, 2011). Given that teleconsultation implementation is highly influenced by external factors such as government legislation (Al-Alawy & Moonesar, 2023), the inclusion of the environment dimension provides a broader and more realistic understanding of the forces driving teleconsultation and its sustainability particularly in the healthcare provider's perspective.

3.2 Coding Process and Category Development

A systematic coding approach was used to identify and organize implementation factors from the reviewed papers. Relevant text segments from each article were coded to capture both pain points (barriers and challenges) and enablers (facilitating factors) related to teleconsultation implementation. These were then clustered to form broader categories and thereafter structured according to TOE dimensions. A high-level summary of the categories is presented in Table 13.

TOE Dimension	Sub-categories
Technology	T1: Technology and infrastructure challenges
	T2: Technical literacy and competence
	T3: IT support
Organization	O1: Workflow and process integration
	O2: Management and leadership engagement
	O3: Guidelines and best practices
Environment	E1: Policy and regulatory support
	E2: External collaborations and partnerships

Table 13: TOE dimension and subcategories

Some recommendations span multiple categories due to the interrelated nature of these factors and thus the categories should be applied flexibly. While this study focuses on teleconsultation, the terms telemedicine and telehealth are often used interchangeably, as many referenced studies do not exclusively address teleconsultation. A summary of the categories that the authors address can be found in Table 14. The subsequent sections will discuss each dimension in detail.

3.3 Technology Dimension

3.3.1 Technical and infrastructure challenges

Technical challenges emerged as a critical barrier to the effective use of teleconsultation, as highlighted by all the reviewed studies. One of the most frequently cited issues is unreliable internet connectivity, which often results in interruptions and communication difficulties during consultations. Additionally, concerns regarding data privacy and security represent a significant obstacle, given that clinical practice frequently involves the transfer of private and sensitive patient information (Abdelghany et al., 2024; Antonacci et al., 2023; Barney et al., 2020; Jallal et al., 2023; Masood et al., 2024).

Another prominent challenge is the lack of integration and interoperability with other telemedicine support systems. This limitation increases the workload for both physicians and healthcare staff, as they must navigate disparate systems compared to traditional sessions (Antonacci et al., 2023; Jallal et al., 2023; Regragui et al., 2023; Theis et al., 2024). Furthermore, technical infrastructure problems, including hardware and software limitations, are particularly prevalent in studies conducted in less developed regions (Brown et al., 2023; Hosseini et al., 2024; Laub et al., 2022).

3.3.2 Technical literacy and competence

Across studies, participants consistently emphasized the necessity of training as a pre-requisite for effective teleconsultation use, identifying the lack of training as a significant barrier to adoption. Clinicians often reported feeling uncertain and anxious during the early stages of implementation, particularly regarding how to deliver care remotely and appropriately (Allison et al., 2023; Brown et al., 2023; Siddiqui et al., 2023).

The anxiety associated with teleconsultation use often stems from unfamiliarity with telehealth systems (Savira et al., 2024). For instance, rural clinicians highlighted their limited knowledge and confidence as a substantial inhibitor, particularly when conducting specialized sessions (Brown et al., 2023). However, some studies revealed that this resistance diminishes with increased exposure and familiarity (Haines et al., 2023; Savira et al., 2024). Clinicians gradually embraced teleconsultation as they observed improvements in patient care and outcomes, boosting their confidence in teleconsultation use (Allison et al., 2023; Brown et al., 2023). Moreover, having regular opportunities to practice skills was seen as vital for maintaining competence and fostering gradual acceptance (Allison et al., 2023; Haines et al., 2023).

Authors	Technology			Organization			Environment	
	T1	T2	T3	O1	O2	O3	E1	E2
Abdelghany et al. (2024)	x	x	x	x	x	x	x	
Agarwal et al. (2020)	x	x				x	x	
Allison et al. (2023)	x	x		x	x	x		x
Antonacci et al. (2023)	x	x		x		x	x	
Barney et al. (2020)	x	x	x	x	x	x	x	x
Bramble et al. (2024)	x	x	x			x	x	
Brown et al. (2023)	x	x			x	x	x	x
Butt et al. (2022)	x	x		x		x		
Etz et al. (2023)	x	x		x	x	x	x	
Fernández Coves et al. (2022)	x						x	
Haines et al. (2023)	x	x			x	x		
Hosseini et al. (2024)	x						x	x
Jallal et al. (2023)	x	x		x	x			
Laub et al. (2022)	x	x	x	x		x		
Masood et al. (2024)	x	x	x	x				
Nataliansyah et al. (2022)	x	x		x			x	x
Regragui et al. (2023)	x	x	x	x	x	x		
Savira et al. (2024)	x	x		x	x	x	x	
Siddiqui et al. (2023)	x	x						

Authors	Technology			Organization			Environment	
	T1	T2	T3	O1	O2	O3	E1	E2
Theis et al. (2024)	x	x		x	x		x	
Turner et al. (2022)	x	x		x		x	x	x

Table 14: Summary of categories covered by included papers

Despite the high demand, there is a dearth of research describing best practices for implementing teleconsultation programs. Existing knowledge is largely fragmented, leaving healthcare systems without well-defined guidelines for clinician training or standardized telemedicine education (Allison et al., 2023; Barney et al., 2020; Theis et al., 2024).

3.3.3 IT support

The availability of IT support has been repeatedly mentioned as integral for ensuring service continuity, reliability, and sustainability in telemedicine adoption. Abdelghany et al. (2024) stressed that dedicated IT support team are vital for ensuring continuity in care delivery, particularly during technical disruptions. Similarly, Turner et al. (2022) highlighted the importance of providing IT assistance, noting that some patients required additional support to successfully navigate telemedicine platforms. IT support also plays a critical role in building confidence among healthcare providers as well.

3.4 Organization Dimension

3.4.1 *Workflow and process integration*

A common view amongst study participants is the lack of workflow integration. While some participants view it as a tool that can save time and reduce their burden (Abdelghany et al., 2024), many find it increased their overall workload. The lack of integration led to poor coordination and task duplication. The additional administrative responsibilities exacerbate the burden on clinicians (Masood et al., 2024; Savira et al., 2024). Such inefficiencies often necessitate additional time for virtual consultations, further straining clinicians (Antonacci et al., 2023).

3.4.2 Management and leadership engagement

Some studies commented on the role of leadership support in creating a supportive environment for telemedicine. For instance, Barney et al. (2020) found that institutional backing, including the provision of necessary resources and infrastructure, is essential to establishing teleconsultation programs. Leadership support also contributed to clinician acceptance by fostering a culture of trust and innovation. Championing the adoption such as sharing success stories and emphasizing improvements in care quality and outcomes are effective strategies for encouraging “buy-in” among clinicians and healthcare staff (Brown et al., 2023).

3.4.3 Guidelines and best practices

Several studies have pointed out the need for protocols and guidelines to address telemedicine-specific challenges and support its integration into clinical workflows (Abdelghany et al., 2024; Barney et al., 2020; Jallal et al., 2023; Laub et al., 2022; Siddiqui et al., 2023). One of the most notable impacts of this gap is the difficulty clinicians face in adapting traditional practices to virtual care environments (Antonacci et al., 2023; Barney et al., 2020). Without such guidance, clinicians may feel unprepared to provide teleconsultation care optimally. For example, Brown et al. (2023) study found the need to establish "telehealth etiquette" as virtual sessions introduced problems such as patients participating in consultations while driving, performing household chores, or in public spaces. Such behaviors can compromise the effectiveness of care, highlighting the need for

clear patient-focused guidelines on how to prepare for and conduct teleconsultation appointments (Butt et al., 2022).

3.5 Environment Dimension

3.5.1 Policy and regulatory support

Effective policy and regulatory support are crucial for the successful and continued adoption of teleconsultation. Restrictive regulatory frameworks hinder the expansion of telemedicine services. Hosseini et al. (2024) noted the lack of a dedicated authority for telemedicine policymaking in Iran has resulted in uncoordinated projects leading to hesitancy among healthcare organizations to implement telemedicine services due to chronic uncertainties. Similar barriers exist in Hong Kong, where teleconsultation service is selectively permitted (Fernández Coves et al., 2022). Antonacci et al. (2023) cites poor reimbursement policies, complex procurement processes, and inadequate data governance frameworks have collectively limited teleconsultation uptake in Italy.

In contrast, supportive legislation and policies have been shown to significantly contribute to teleconsultation adoption. This is in part due to the impact of the COVID-19 pandemic with several studies acknowledging it as a big factor for its uptake (Antonacci et al., 2023; Brown et al., 2023; Etz et al., 2023; Jallal et al., 2023). For example, in Italy, simplifying compliance requirements for data governance, privacy, and reimbursement was instrumental in telemedicine uptake during the pandemic (Antonacci et al., 2023). However, as many of these measures were put into place in reaction to a temporary pandemic, a number of participants (Barney et al., 2020; Theis et al., 2024; Turner et al., 2022) have expressed their doubts on its sustainability particularly if these measures are suddenly retracted. Nevertheless, these measures demonstrated the power of regulatory support to enhance access to teleconsultation services and could serve as a foundation for long-term policies.

3.5.2 External collaborations and partnerships

A few studies also cited collaborations and external support as a factor in teleconsultation use. Allison et al. (2023) highlight how partnerships with several related institutions as well as IT service support from a university has helped mitigate technical and system-level challenges, enabling more seamless telemedicine operations. These collaborations demonstrate that external support not only facilitates immediate implementation but also provides the infrastructure for sustained usage.

Many studies have also suggested educational partnerships as a way to bolster telemedicine's effectiveness and sustainability. Integrating telemedicine competencies into healthcare education curricula equips future clinicians with the skills needed to navigate virtual care environments effectively (Antonacci et al., 2023).

Long-term sustainability of teleconsultation also depends on fostering trust, communication, and meaningful collaboration between institutions, such as metropolitan and rural healthcare providers. Brown et al. (2023) notes that strong relationships between clinicians in urban hubs and rural spokes improve referral quality and volume, creating a network that benefits both providers and patients. Together, these studies illustrate that institutional and educational collaborations are foundational to teleconsultation's sustained success.

4 Discussion

4.1 Recommendations

A set of recommendations were extrapolated based on the findings in Section 0. The technology domain is the most consistently raised dimension across the included studies and thus suggests that it is the key success driver for teleconsultation implementation. Therefore, this dimension should be prioritized over organizational and environmental factors during the initial stages of teleconsultation implementation. As technological systems evolve and stabilize, the organizational and environmental dimensions are expected to assume greater importance, addressing broader systemic and operational challenges to implementation

4.1.1 Recommendation 1: Invest in reliable technical infrastructure and IT Support

Across the studies reviewed, inadequate technology reliability and the lack of adequate support consistently emerged as significant sources of frustration for end-users. Challenges such as unstable internet connections, lack of platform integration, and inadequate hardware significantly disrupt telemedicine workflows and discourage both providers and patients from its use. These issues have an immediate and detrimental impact on the perceived value and effectiveness of teleconsultation platforms, a result that is consistent with Anthony Jr (2024) findings. A reliable and interoperable technical system is therefore foundational in ensuring teleconsultation's continued value and experience.

While technical training for clinicians and staff is important, the benefits of such training often take time to materialize. In contrast, having a robust IT support system provides an immediate safety net, offering immediate solutions to technical issues. This not only ensures continuity of care during teleconsultations but also enhances the confidence of both clinicians and patients in the system. The psychological reassurance provided by dependable IT support fosters trust, encourages acceptance, and helps mitigate resistance to teleconsultation. By promptly addressing technical barriers and maintaining system reliability, IT support plays a critical role in ensuring the successful implementation and sustainability of teleconsultation services.

4.1.2 Recommendation 2: Execute process integration and establish clear guidelines

Section 3.4 reveals that teleconsultation, an innovation that is supposed to improve efficiency, also became a source of increased burden largely due to failure to accommodate the process disruption of technology use. Therefore, as the technology aspect is stabilized, process integration plans must be put in place to fully realize the teleconsultation benefits. Along with the process integration, guidelines and rules must also be implemented to ensure that there are roadmaps and signposts for physicians and other healthcare providers to follow while they navigate the change. The issues raised are consistent with the organizational and change management issues (Al Ahbabi & Ahmad, 2021) which currently remain an understudied area in the context of healthcare (Milella et al., 2021) and can be explored in future research.

4.1.3 Recommendation 3: Establish strategic partnerships and collaborations

While regulatory policies are critical for successful teleconsultation implementation, healthcare providers often have limited influence over such changes. Therefore, focusing on expanding external collaborations is recommended to increase the support and perceived value of teleconsultation (Steinhauser, 2019). As mentioned in Section 0 partnerships between urban and rural healthcare providers can enhance the perceived value of teleconsultation. These collaborations foster trust, facilitate resource sharing, and enable the exchange of technical expertise.

Such collaborations also create opportunities to address systemic barriers, including reimbursement challenges and payment support, by paving the way for partnerships with regulatory and financial stakeholders. By establishing strategic partnerships, healthcare providers can enhance the scalability, sustainability, and overall value of teleconsultation, even in the absence of immediate regulatory changes.

5 Conclusion and Future Research

This systematic review underscores that the successful implementation of teleconsultation hinges on prioritizing robust technological infrastructure while simultaneously addressing organizational and environmental dimensions. Reliable IT systems, real-time technical support, and targeted training programs for clinicians are foundational to fostering trust and confidence among healthcare providers. Equally critical are organizational changes, such as workflow integration, leadership engagement, and clear guidelines, which align teleconsultation practices with clinical objectives. Additionally, supportive regulatory frameworks and strategic external collaborations are pivotal in overcoming systemic barriers and ensuring the long-term sustainability of teleconsultation.

This review is subject to certain limitations. The inclusion criteria restricted the studies to those published between 2020 and 2024, which, while ensuring increased relevance to the post-COVID-19 landscape, may have excluded valuable insights from earlier foundational research. Future research should also explore the role of change management in teleconsultation implementation. More case studies and actual observational research would also be useful. Beyond the healthcare provider's perspective, future research should expand to include the views of healthcare decision-makers and implementation managers, whose roles are instrumental in scaling teleconsultation services. Decision-makers bring valuable insights into policy formulation, resource allocation, and long-term strategic planning, while implementation managers provide ground-level knowledge about operational challenges and practical solutions. Exploring their perspectives can enrich the understanding of implementation barriers and enablers, leading to more comprehensive frameworks for teleconsultation implementation.

References

Abdelghany, I. K., AlMatar, R., Al-Haqan, A., Abdullah, I., & Waheedi, S. (2024). Exploring healthcare providers' perspectives on virtual care delivery: insights into telemedicine services. *BMC Health Services Research*, *24*(1), 1. https://doi.org/10.1186/s12913-023-10244-w

Al Ahbabi, J. M., & Ahmad, S. Z. (2021). Using health-care technology to manage COVID-19 pandemic: a case of Al-Ain city hospital, UAE. *Emerald Emerging Markets Case Studies*, *11*(3), 1-18. https://doi.org/10.1108/EEMCS-10-2020-0356

Al-Alawy, K., & Moonesar, I. A. (2023). Perspective: Telehealth – beyond legislation and regulation. *SAGE Open Medicine*, *11*, 20503121221143223. https://doi.org/10.1177/20503121221143223

Aldawoud, F. L., Al-Zaher, F. A., Suhail, Z. H., Al Nasser, Z. H., Almubayidh, S. M., Alzahrani, F. Y., Alluwaif, H. R., Al Qanbar, N. H., & Aljezany, M. N. (2024). Obstacles and enablers for using telehealth services online in nursing practice: A systematic review. *World Journal of Advanced Research and Reviews*, *23*(3), 1742–1750. https://doi.org/10.30574/wjarr.2024.23.3.2846

Allen, J. D., Towne, S. D., Maxwell, A. E., DiMartino, L., Leyva, B., Bowen, D. J., Linnan, L., & Weiner, B. J. (2017). Measures of organizational characteristics associated with adoption and/or implementation of innovations: A systematic review. *BMC Health Services Research*, *17*(1), 591. https://doi.org/10.1186/s12913-017-2459-x

Allison, M. K., Curran, G. M., Walsh, W. A., Dworkin, E. R., & Zielinski, M. J. (2023). Factors Affecting Telemedicine Implementation in Emergency Departments and Nurses' Perceptions of Virtual Sexual Assault Nurse Examiner Consultation for Sexual Assault Survivors. *Journal of Forensic Nursing*, *19*(1), 41-49. https://doi.org/10.1097/JFN.0000000000000385

Anthony Jr, B. (2024). Examining the adoption of telehealth during public health emergencies based on technology organization environment framework. *Journal of Science and Technology Policy Management*, *15*(6), 1311-1338. https://doi.org/10.1108/JSTPM-05-2022-0079

Antonacci, G., Benevento, E., Bonavitacola, S., Cannavacciuolo, L., Foglia, E., Fusi, G., Garagiola, E., Ponsiglione, C., & Stefanini, A. (2023). Healthcare professional and manager perceptions on drivers, benefits, and challenges of telemedicine: results from a cross-sectional survey in the Italian NHS. *BMC Health Services Research*, *23*(1), 1115. https://doi.org/10.1186/s12913-023-10100-x

Baker, J. (2012). The Technology–Organization–Environment Framework. In Y. K. Dwivedi, M. R. Wade, & S. L. Schneberger (Eds.), *Information Systems Theory: Explaining and Predicting Our Digital Society, Vol. 1* (pp. 231-245). Springer New York. https://doi.org/10.1007/978-1-4419-6108-2_12

Baldwin, J. R., Pingault, J. B., Schoeler, T., Sallis, H. M., & Munafò, M. R. (2022). Protecting against researcher bias in secondary data analysis: challenges and potential solutions. *Eur J Epidemiol*, *37*(1), 1-10. https://doi.org/10.1007/s10654-021-00839-0

Barbosa da Silva, J., Espinal, M., Garcia-Saiso, S., Fitzgerald, J., Marti, M., Bascolo, E., Haddad, A. E., & D'Agostino, M. (2024). A digital transformation for primary health care. *Bulletin of the World Health Organization*, *102*(1), 2-2A. https://doi.org/10.2471/BLT.23.290726

Barney, A., Buckelew, S., Mesheriakova, V., & Raymond-Flesch, M. (2020). The COVID-19 Pandemic and Rapid Implementation of Adolescent and Young Adult Telemedicine: Challenges and Opportunities for Innovation. *Journal of Adolescent Health*, *67*(2), 164-171. https://doi.org/10.1016/j.jadohealth.2020.05.006

Brown, A. D., Kelso, W., Velakoulis, D., Farrand, S., & Stolwyk, R. J. (2023). Understanding Clinician's Experiences with Implementation of a Younger Onset Dementia Telehealth Service. *Journal of geriatric psychiatry and neurology*, *36*(4), 295-308. https://doi.org/10.1177/08919887221141653

Butt, Z., Kirsten, L., Beatty, L., Kelly, B., Dhillon, H., & Shaw, J. M. (2022). Barriers and enablers to implementing telehealth consultations in psycho-oncology. *Psycho-oncology*, *31*(8), 1365-1373. https://doi.org/10.1002/pon.5939

Dhunnoo, P., Kemp, B., McGuigan, K., Meskó, B., O'Rourke, V., & McCann, M. (2024). Evaluation of Telemedicine Consultations Using Health Outcomes and User Attitudes and Experiences: Scoping Review. *J Med Internet Res*, *26*, e53266. https://doi.org/10.2196/53266

Diel, S., Doctor, E., Reith, R., Buck, C., & Eymann, T. (2023). Examining supporting and constraining factors of physicians' acceptance of telemedical online consultations: a survey

study. *BMC Health Services Research*, *23*(1), 1128. https://doi.org/10.1186/s12913-023-10032-6

Etz, R. S., Solid, C. A., Gonzalez, M. M., Britton, E., Stange, K. C., & Reves, S. R. (2023). Telemedicine in Primary Care: Lessons Learned About Implementing Health Care Innovations During the COVID-19 Pandemic. *Annals of Family Medicine*, *21*(4), 297-304. https://doi.org/10.1370/afm.2979

Ewald, H., Klerings, I., Wagner, G., Heise, T. L., Stratil, J. M., Lhachimi, S. K., Hemkens, L. G., Gartlehner, G., Armijo-Olivo, S., & Nussbaumer-Streit, B. (2022). Searching two or more databases decreased the risk of missing relevant studies: a metaresearch study. *Journal of Clinical Epidemiology*, *149*, 154-164. https://doi.org/10.1016/j.jclinepi.2022.05.022

Federal Communications Commission. (2024). *Telehealth, Telemedicine, and Telecare: What's What?* https://www.fcc.gov/general/telehealth-telemedicine-and-telecare-whats-what

Fernández Coves, A., Yeung, K. H. T., van der Putten, I. M., & Nelson, E. A. S. (2022). Teleconsultation adoption since COVID-19: Comparison of barriers and facilitators in primary care settings in Hong Kong and the Netherlands. *Health policy (Amsterdam, Netherlands)*, *126*(10), 933-944. https://doi.org/10.1016/j.healthpol.2022.07.012

Gagnon, M.-P., Desmartis, M., Labrecque, M., Car, J., Pagliari, C., Pluye, P., Frémont, P., Gagnon, J., Tremblay, N., & Légaré, F. (2012). Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. *Journal of Medical Systems*, *36*(1), 241-277. https://doi.org/10.1007/s10916-010-9473-4

Garavand, A., Aslani, N., Nadri, H., Abedini, S., & Dehghan, S. (2022). Acceptance of telemedicine technology among physicians: A systematic review. *Informatics in Medicine Unlocked*, *30*, 100943. https://doi.org/10.1016/j.imu.2022.100943

Grata-Borkowska, U., Sobieski, M., Drobnik, J., Fabich, E., & Bujnowska-Fedak, M. M. (2022). Perception and Attitude toward Teleconsultations among Different Healthcare Professionals in the Era of the COVID-19 Pandemic. *International Journal of Environmental Research and Public Health*, *19*(18), 11532. https://doi.org/10.3390/ijerph191811532

Haines, K. J., Sawyer, A., McKinnon, C., Donovan, A., Michael, C., Cimoli, C., Gregory, M., Berney, S., & Berlowitz, D. J. (2023). Barriers and enablers to telehealth use by physiotherapists during the COVID-19 pandemic. *Physiotherapy*, *118*, 12-19. https://doi.org/10.1016/j.physio.2022.09.003

Harari, M. B., Parola, H. R., Hartwell, C. J., & Riegelman, A. (2020). Literature searches in systematic reviews and meta-analyses: A review, evaluation, and recommendations. *Journal of Vocational Behavior*, *118*, 103377. https://doi.org/10.1016/j.jvb.2020.103377

Hartling, L., Featherstone, R., Nuspl, M., Shave, K., Dryden, D. M., & Vandermeer, B. (2016). The contribution of databases to the results of systematic reviews: a cross-sectional study. *BMC Medical Research Methodology*, *16*(1), 127. https://doi.org/10.1186/s12874-016-0232-1

Hassan, A., Mari, Z., Gatto, E. M., Cardozo, A., Youn, J., Okubadejo, N., Bajwa, J. A., Shalash, A., Fujioka, S., Aldaajani, Z., Cubo, E., & the International Telemedicine Study, G. (2020). Global Survey on Telemedicine Utilization for Movement Disorders During the COVID-19 Pandemic. *Movement Disorders*, *35*(10), 1701-1711. https://doi.org/https://doi.org/10.1002/mds.28284

Hausman, A., & Stock, J. R. (2003). Adoption and implementation of technological innovations within long-term relationships. *Journal of Business Research*, *56*(8), 681-686. https://doi.org/https://doi.org/10.1016/S0148-2963(01)00313-7

Hosseini, S. M., Boushehri, S. A., & Alimohammadzadeh, K. (2024). Challenges and solutions for implementing telemedicine in Iran from health policymakers' perspective. *BMC Health Services Research*, *24*(1), 50. https://doi.org/10.1186/s12913-023-10488-6

Iliyasu, B. Z., Iliyasu, Z., Kwaku, A. A., Sani, A., Nass, N. S., Amole, T. G., Abdullahi, H. M., Abdullahi, A. U., Tsiga-Ahmed, F. I., Jibo, A. M., Mbbch, H. A. B., Salihu, H. M., & Aliyu, M. H. (2024). Acceptability of Teleconsultation Services for HIV Care in Nigeria: A Mixed Methods Study. *Telemedicine and e-Health*. https://doi.org/10.1089/tmj.2024.0196

Jallal, M., Berrada, K., Bouaddi, O., El Badisy, I., Charaka, H., Serhier, Z., Bennani Othmani, M., Nejjari, C., & Khalis, M. (2023). How Physicians in a Moroccan Tertiary Care Center Perceive Teleconsultation During COVID-19 Pandemic? *Telemedicine journal and e-health : the official journal of the American Telemedicine Association*, *29*(2), 284-292. https://doi.org/10.1089/tmj.2022.0094

Laub, N., Agarwal, A. K., Shi, C., Sjamsu, A., & Chaiyachati, K. (2022). Delivering Urgent Care Using Telemedicine: Insights from Experienced Clinicians at Academic Medical Centers. *Journal of general internal medicine*, *37*(4), 707-713. https://doi.org/10.1007/s11606-020-06395-9

Masood, S., Samad, Z., Nadeem, S., & Ghulam, U. (2024). Telemedicine During a Pandemic: Challenges, Opportunities & Way Forward at a Tertiary Care Hospital. *Hospital Topics*, *102*(1), 9-15. https://doi.org/10.1080/00185868.2022.2073304

Mathew, M. J. (2024). Literature search in systematic reviews: How much is good enough? *Clinical Epidemiology and Global Health*, *25*, 101485. https://doi.org/10.1016/j.cegh.2023.101485

Milella, F., Minelli, E. A., Strozzi, F., & Croce, D. (2021). Change and Innovation in Healthcare: Findings from Literature. *ClinicoEconomics and Outcomes Research*, *13*(null), 395-408. https://doi.org/10.2147/CEOR.S301169

Nagarathinam, S., Umadevi, R., Grace, A., & Saravanan, R. (2024). A cross-sectional study to assess the barriers to teleconsultations among doctors in Chennai. *Journal of Applied Pharmaceutical Research*, *12*, 29-34. https://doi.org/10.18231/j.joapr.2024.12.1.29.34

Nittas, V., & von Wyl, V. (2020). COVID-19 and telehealth: a window of opportunity and its challenges. *Swiss medical weekly*, *150*(1920), w20284-w20284.

Oliveira, T., & Martins, M. R. (2011). Literature Review of Information Technology Adoption Models at Firm Level. *1566-6379*, *14*(1), 110-121. https://academic-publishing.org/index.php/ejise/article/view/389

Osman, M. A., Schick-Makaroff, K., Thompson, S., Bialy, L., Featherstone, R., Kurzawa, J., Zaidi, D., Okpechi, I., Habib, S., Shojai, S., Jindal, K., Braam, B., Keely, E., Liddy, C., Manns, B., Tonelli, M., Hemmelgarn, B., Klarenbach, S., & Bello, A. K. (2019). Barriers and facilitators for implementation of electronic consultations (eConsult) to enhance access to specialist care: a scoping review. *BMJ Global Health*, *4*(5), e001629. https://doi.org/10.1136/bmjgh-2019-001629

Oudbier, S. J., Souget-Ruff, S. P., Chen, B. S. J., Ziesemer, K. A., Meij, H. J., & Smets, E. M. A. (2024). Implementation barriers and facilitators of remote monitoring, remote consultation and digital care platforms through the eyes of healthcare professionals: a review of reviews. *BMJ open*, *14*(6), e075833. https://doi.org/10.1136/bmjopen-2023-075833

Pussegoda, K., Turner, L., Garritty, C., Mayhew, A., Skidmore, B., Stevens, A., Boutron, I., Sarkis-Onofre, R., Bjerre, L. M., Hróbjartsson, A., Altman, D. G., & Moher, D. (2017).

Systematic review adherence to methodological or reporting quality. *Systematic Reviews*, *6*(1), 131. https://doi.org/10.1186/s13643-017-0527-2

Regragui, S., Abou Malham, S., Gaboury, I., Bois, C., Deville-Stoetzel, N., Maillet, L., Savoie, A., & Breton, M. (2023). Nursing practice and teleconsultations in a pandemic context: A mixed-methods study. *Journal of clinical nursing*, *32*(17-18), 6339-6353. https://doi.org/10.1111/jocn.16756

Ross, J., Stevenson, F., Lau, R., & Murray, E. (2016). Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). *Implementation Science*, *11*(1), 146. https://doi.org/10.1186/s13012-016-0510-7

Savira, F., Sanigorski, A., Namara, K. M., Hensher, M., Peeters, A., & Manias, E. (2024). Perspectives on telehealth implementation in Australia: An exploratory qualitative study with practice managers and general practitioners. *International Journal of Medical Informatics*, *188*, 105473. https://doi.org/10.1016/j.ijmedinf.2024.105473

Senyel, D., Senn, K., Boyd, J., & Nagels, K. (2024). A systematic review of telemedicine for neuromuscular diseases: components and determinants of practice. *BMC Digital Health*, *2*(1), 17. https://doi.org/10.1186/s44247-024-00078-9

Siddiqui, S., Farr, E., Dusto, N., Chen, L., Kocherginsky, M., Skelton, F., Verduzco-Gutierrez, M., & Lee, S. (2023). Telemedicine Use Among Physiatrists During the Early Phase of the COVID-19 Pandemic and Potential for Future Use. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association*, *29*(2), 242-252. https://doi.org/10.1089/tmj.2022.0030

Steinhauser, S. (2019). Network-Based Business Models, the Institutional Environment, and the Diffusion of Digital Innovations: Case Studies of Telemedicine Networks in Germany. *Schmalenbach Business Review*, *71*(3), 343-383. https://doi.org/10.1007/s41464-019-00076-9

Teixeira da Silva, J. A., & Daly, T. (2024). Against Over-reliance on PRISMA Guidelines for Meta-analytical Studies. *Rambam Maimonides Med J*, *15*(1). https://doi.org/10.5041/rmmj.10518

Theis, R. P., Dorbu, J. I., Mavrodieva, M. E., Guerrero, R. A., Wright, S. E., Donahoo, W. T., Modave, F., Carrasquillo, O., & Shenkman, E. A. (2024). Telehealth Implementation Response to COVID-19 in the OneFlorida+ Clinical Research Network: Perspectives of Clinicians and Health Systems Leaders. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association*, *30*(1), 268-277. https://doi.org/10.1089/tmj.2023.0029

Tornatzky, L. G., Fleischer, M., & Chakrabarti, A. K. (1990). *The Processes of Technological Innovation*. Lexington Books. https://books.google.co.nz/books?id=QWiwAAAAIAAJ

Turner, K., Bobonis Babilonia, M., Naso, C., Nguyen, O., Gonzalez, B. D., Oswald, L. B., Robinson, E., Elston Lafata, J., Ferguson, R. J., Alishahi Tabriz, A., Patel, K. B., Hallanger-Johnson, J., Aldawoodi, N., Hong, Y.-R., Jim, H. S. L., & Spiess, P. E. (2022). Health Care Providers' and Professionals' Experiences With Telehealth Oncology Implementation During the COVID-19 Pandemic: A Qualitative Study. *Journal of Medical Internet Research*, *24*(1), e29635. https://doi.org/10.2196/29635

van Dinter, R., Tekinerdogan, B., & Catal, C. (2021). Automation of systematic literature reviews: A systematic literature review. *Information and Software Technology*, *136*, 106589. https://doi.org/10.1016/j.infsof.2021.106589

Wright, K., Golder, S., & Lewis-Light, K. (2015). What value is the CINAHL database when searching for systematic reviews of qualitative studies? *Systematic Reviews*, *4*(1), 104. https://doi.org/10.1186/s13643-015-0069-4

14. SWIFT CRISIS RESPONSE: THE ROLES OF ORGANIZATIONAL MINDFULNESS, CITIZENSHIP BEHAVIOR AND DIGITAL INFRASTRUCTURE

Edward W.N. Bernroider
Vienna Univ. of Economics and Business
edward.bernroider@wu.ac.at

Everist Limaj
Vienna Univ. of Economics and Business
everist.limaj@wu.ac.at

Nikolaus Obwegeser
Bern University of Applied Sciences (BFH)
nikolaus.obwegeser@bfh.ch

Abstract
Recent external crises have revealed limitations of traditional business continuity planning in fostering organizational agility during unprecedented situations. Drawing on socio-technical perspectives, our study examines how two understudied factors, namely organizational mindfulness toward digital transformation (MDT) and organizational citizenship behavior (OCB), enable agile crisis responses supported by digital infrastructure capabilities. MDT reflects an organization's awareness of the transformative impact of technology, while OCB comprises voluntary, prosocial employee behaviors that enhance organizational learning and adaptability. Analyzing primary survey data, we demonstrate that MDT, loosely coupled digital systems, and OCB can be associated with enhanced organizational agility, which in turn supports competitive performance during a crisis.

Keywords
Organizational Agility; Crises; Mindfulness Towards Digital Transformation; Organizational Citizenship Behavior; Empirical Study

1 Introduction
Although prior work underscored the importance of organizational preparedness for crisis situations (e.g. El Idrissi et al., 2022; Fowler et al., 2007), organizations can only partly rely on mindful preparation. Business continuity plans, for example, are typically designed to address predictable disruptions or disasters, rather than sudden events like the Covid-19 pandemic, which introduced unprecedented environmental conditions. Therefore, plans and risk-averting infrastructures mindfully prepared pre-crisis might not fit the requirements during such an unprecedented crisis (Ye et al., 2022). New digital innovations typically emerge as a crisis progresses, which come with opportunities and risks for unique challenges to organizations and their employees (Neumannova et al., 2023).

The acceleration of digital transformation in a crisis context (Soto-Acosta, 2020) has two important aspects, which have been considered as general assumptions in this study. First, we anticipated that organizational mindfulness towards digital transformation (MDT), i.e., organizations being fully aware of the impact that technology has on an organization's processes and strategies, should

be in a better position to process these novel requirements that may not have been accounted for in traditional business continuity planning. Second, we assume that organizational citizenship behavior (OCB) (Davison et al., 2020) has become an crucial organizational learning capacity during the crisis when employees went beyond their duties to help colleagues in a discretionary manner. Whilst some studies have confirmed the importance of OCB with many organizations struggling with digital transformation in general, it should have received an even more prominent role during a crisis.

Therefore, the aim of our research is to explore new avenues for enhancing organizational agility and performance during unprecedented crises, and to investigate whether mindfulness towards digital transformation (MDT) helps organizations to make informed decisions about their digital infrastructure, and supports organizational agility to enable swift changes required. Additionally, we anticipate that voluntary supportive behavior plays an important role in this context. We thus set out to answer the following research questions in the crisis context, exemplified by the Covid-19 pandemic: (*RQ1) How is MDT related to the digital infrastructure and organizational agility? (RQ2) How does OCB influence organizational agility? (RQ3) What are the implications for competitive performance?* Combined, our integrated view on MDT and OCB in conjunction with technology-related changes recognizes that any organizational change is inherently socio-technical in nature and is shaped by issues around knowledge flows in interactions between people and technology (Fischbacher-Smith, 2017).

2 Theory and Hypotheses Development

Next, we will develop our conceptual background and the research model shown in Figure 1.

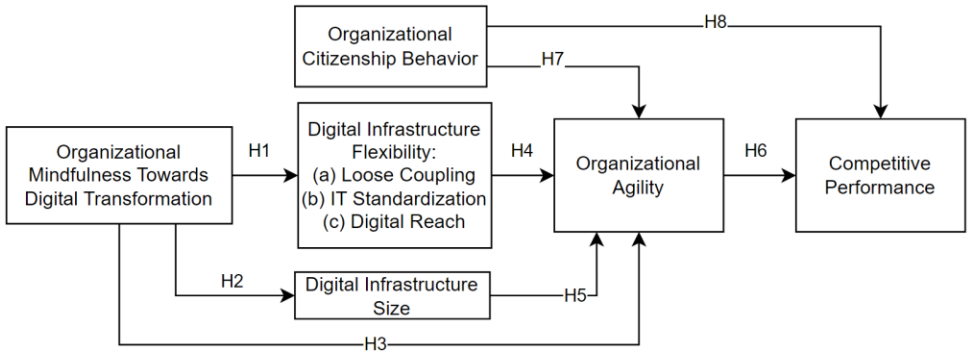

Figure 1. Developed Conceptual Model

2.1 Mindfulness Towards Digital Transformation

In the context of digital transformation, mindfulness means having a well-developed understanding of the positive and negative effects of IT investments on the entire organization and its culture, which can be achieved by regularly searching for opportunities of digital transformation, providing alternatives for decision-making, and anticipating and evaluating potential and realized effects (Li et al., 2021). Essentially, organizations that are mindful towards digital transformation have the capacity to stay alert to future technology changes and can quickly provide options aligned with strategic decisions in changing environments. Prior research has typically focused on the anticipation perspective of mindful resource preparation to safeguard

organizations, which in particular includes creating risk-aversion systems and accumulating sufficient internal resources prior to a crisis (Ye et al., 2022). In practice, pre-crisis risk-aversion behavior is addressed by business continuity and disaster recovery management (Herbane et al., 2004; Niemimaa et al., 2019). Contemporary research, however, has been shifting its attention to the importance of mindful behaviors, i.e., being consciously attentive and fully present in the current moment (Khan et al., 2024).

First, we turn to the role of MDT in conjunction with the focal organization's *digital infrastructure*, which reflects the arrangement through which various technology elements such as software, data, network and hardware are interlinked and used to provide IT services (Nyrhinen, 2008). It is commonly agreed that nowadays the digital infrastructure is crucial for maintaining operations and the achievement of sustainable competitive advantage as a core competency of an organization (Byrd & Turner, 2001; Mikalef et al., 2021). Contemporary studies on digital architectures emphasize the importance of flexibility (Mikalef et al., 2021; Tiwana & Konsynski, 2010; Xu et al., 2010), which can be defined as the as the degree to which the focal organization's digital resources are sharable and reusable (Byrd & Turner, 2000). Such flexibility is achieved through IT architecture modularity characterized by the interaction of loosely coupled systems, organization-wide IT standards and digital reach (Byrd & Turner, 2001; Mikalef et al., 2021; Tafti et al., 2013; Tiwana & Konsynski, 2010).

A key aspect of being mindful of digital transformation is the importance of continuous attentiveness and adaptability in incorporating digital technology that aligns with business requirements (Li et al., 2021). In order to facilitate these changes efficiently and effectively, it is imperative to establish a flexible digital infrastructure, which has long been recognized as strategically important (Mikalef et al., 2021). This flexibility enables organizations to remain open to emerging digital options, as being mindful of digital transformation requires. Similarly, mindful organization should be aware of current blind spots in the digital architecture and take actions for improvement (Butler & Gray, 2006), which extends to ensuring digital reach across the organization. Awareness of, and subsequent acting on quickly closing any significant gaps in the digital infrastructure will arguably require flexible IT infrastructures. As illustrated above, prior work has shown that flexible IT infrastructures can be fostered by a loosely coupled digital infrastructure, IT standardization and digital reach (Byrd & Turner, 2001; Mikalef et al., 2021; Tafti et al., 2013; Tiwana & Konsynski, 2010), we consequently assume:

H1: MDT is positively related with a digital infrastructure flexibility characterized by (a) loose coupling, (b) IT standardization, and (c) digital reach.

A mindful approach to digital transformation implies that organizations are geared towards regularly assessing their current digital infrastructure and identifying areas that can be improved to align with the overall strategy of the organization, and thereby ensuring that resources are allocated in a way that delivers the greatest value (Li et al., 2021). The implications of digital transformation can be cost savings by streamlining processes, reducing manual labor and optimizing resource utilization. However, it can be assumed that it can lead to increased IT costs as organizations invest in new digital technologies and their support. Prior research has shown that IT investments need an appropriate level of funding to ensure effective IT appraisals and implementation (Bernroider, 2013) as well as long-term IT operations despite potential overall

cost or efficiency savings for the organization as a whole (e.g. Endres et al., 2020; Guenzi & Habel, 2020). Hence, we hypothesize:

H2: MDT is positively related with size of an organization's digital infrastructure.

2.2 Organizational Agility

Organizational agility refers to an organization's capability to promptly and effectively respond to changes occurring within its internal or external environment, which involves the ability to swiftly adapt strategies, processes, and resources in order to tackle new challenges and seize opportunities (Sambamurthy et al., 2003). Prior research has presented various conceptualizations of organizational agility (Pulakos et al., 2019). We have adopted a dual perspective that differentiates between proactive and reactive components, focusing on capitalizing on new market opportunities and making operational adjustments, respectively (e.g. Li et al., 2020; Panda, 2022). The concepts of market capitalizing and operational adjustment align conceptually with the more conventional sensing and responding capabilities, which have also been reviewed in conjunction with organizational agility (Mao et al., 2021; Panda & Rath, 2018).

An agile organization is characterized by a flexible and adaptive culture, and streamlined decision-making processes allowing for a willingness to experiment and learn from mistakes (Limaj & Bernroider, 2022), especially in the context of a severe crisis (Arokodare & Falana, 2021). A high level of mindfulness toward digital transformation can help the organization to adapt to these changes more effectively by promoting a sense of openness, curiosity, and willingness to learn on the individual and collective level (Butler & Gray, 2006). Organizations that are mindful towards digital transformation are aware or can quickly assess investment opportunities regarding digital solutions and therefore have the ability to renew themselves strategically and stay flexible without sacrificing operations (Elali, 2021). On the individual level, mindfulness practices can increase innovative work behavior of employees as they recognize external changes, and the need for exploration and exploitation of new ideas quickly (Khan et al., 2024). This can be particularly valuable during a time of crisis, when organizations may need new solutions to meet sudden challenges. Thus, we propose:

H3: MDT is positively related with organizational agility.

Furthermore, organizational agility is likely to benefit from flexibility in digital infrastructures especially during a crisis, when swift changes must be accommodated by established structures. When an organization's digital infrastructure is designed on the basis of *loose coupling*, it should have also enhanced its organizational agility by allowing for greater flexibility and adaptability (Mooney & Ganley, 2006). Our second dimension on flexible IT infrastructures captures the role of *IT standardization*, which can also be directly linked with organizational agility. IT standardization reduces complexity and fosters harmonization, enabling enhanced planning capabilities and increased flexibility to accommodate new functions and improve the scalability of existing functions (Ahlemann et al., 2023; van Wessel, 2010). *Digital reach* should also positively affect organizational agility as it allows for the integration of data from various sources and enables organizations to bridge gaps and collaborate with internal and external stakeholders outside traditional boundaries. Digital reach improves IT service transparency and accessibility regardless of the actual location of the stakeholder in the network (Mikalef et al., 2021). Such digital external

and internal integration enables boundary spanning in various forms making organizations more agile by increasing organizational sensing and responding capabilities (Nazir & Pinsonneault, 2012). Especially in times of high environmental uncertainty, prioritizing loose coupling, IT standardization, and digital reach within an organization's digital infrastructure is considered crucial for enabling the organization to adapt more quickly to changes and respond to challenges (Mikalef et al., 2021). Therefore, we propose:

H4: A more flexible IT infrastructure, in terms of (a) loose coupling, (b) IT standardization, and (c) digital reach, is positively related with organizational agility.

Next, we discuss the role of digital infrastructure size for organizational agility. While IS research is likely to advise that c-level executives should invest in new IT solutions for various reasons (Tallon et al., 2022), certain investments into IT may also temper organizational agility (Queiroz et al., 2023), esp. due to legacy applications (Limaj et al., 2020; Wimelius et al., 2021). In general, however, expanding digital transformation should shift organizations towards more flexible structures and develop capacities that facilitate ongoing adaptation (Hanelt et al., 2021). For example, automation can be fostered across all business sections, from computer vision technology to improve agricultural automation (Tian et al., 2020), which should free up time to focus on higher value activities, such as strategic changes. Thus, we argue:

H5: The size of an organization's digital infrastructure is positively related with organizational agility.

Prior research has firmly established the importance of organizational agility reflected by innovation capabilities for organizational performance. For example, prior work has suggested that open innovation benefits organizations in terms of enhanced organizational performance (Singh et al., 2021) These results are further supported by prior research using capability-building hierarchies to establish organizational agility as a mediator for translating the positive effects of IT capabilities into enhanced organizational performance (Queiroz et al., 2018). This is particularly applies in terms of operational adjustment agility and market capitalizing agility (Liu et al., 2014; Panda, 2022). Hence, we assume
:
H6: Organizational agility positively affects competitive performance.

2.3 Organizational Citizenship Behavior

Organizational Citizenship Behavior (OCB), also termed extra-role behavior, is a term in the field of organizational psychology that describes the voluntary and discretionary actions of employees that exceed their official job duties (Bateman & Organ, 1983; Smith et al., 1983). It was reported that when employees beyond their formal duties to understand customer needs and rapidly collaborate to exchange know-how, they are better equipped to fulfill new demands and create customer value (Holbeche, 2019). In the IS context, innovative OCB behaviors include creating and offering workarounds to solve technical problems, e.g., to help colleagues solve Enterprise Resource Planning (ERP) system issues (Davison et al., 2020). Ultimately, OCB therefore becomes linked with social capital, which promotes collaborative knowledge creation and is a key driver of organizational agility, especially in response to crises (Al-Omoush et al., 2020). Based on these arguments, we posit that:

H7: OCB positively affects organizational agility.

Finally, we turn to the role of OCB in relation to competitive performance during a crisis. From a political perspective, OCB captures all organizationally relevant employee behaviors (Graham, 1991), which suggests that it can be inherently linked with individual performance. In terms of social exchange theory, the effects of such behaviors are theoretically expected to be mutually reinforcing throughout the organization, as individuals tend to reciprocate favors received, particularly if they experience a high level of job satisfaction (Bateman & Organ, 1983). Empirically, prior research corroborates this perspective. For instance, it was shown that OCB has a positive effect on employee performance especially through enhanced job satisfaction (Mahrani & Alwi, 2022). A literature review has generally confirmed that OCB affects competitive performance, whilst noting that further research is required to clarify the impacts of OCBs to account for different contexts (Podsakoff & MacKenzie, 2014). It should also not be overlooked that OCB has been considered as an implicit dimension of teamwork quality, which was related with increased performance (Xu et al., 2010). We thus assume:

H8: OCB positively affects competitive performance.

3 Data Collection and Analysis

3.1 Instrument

To develop the instrument, we first developed unambiguous conceptual definitions for each construct against the background of our crisis context, which involved reconceptualizations of definitions used in prior studies refined to fit our research context. All constructs have been measured reflectively with most based on items from previous studies (see Table 1).

Digital infrastructure size, a self-developed construct, assesses the extent of an organization's IT resources and investments using IT-focused FTEs, IT spending relative to revenue, and competitor comparisons. In line with current studies (e.g. Panda, 2022), we conceptualized Organizational Agility (OA) as a two-dimensional construct to capture its multifaceted nature in response to market changes and internal adjustments, thereby considering Market Capitalizing Capability (MCAP) and Operational Adjustment Capability (OADJ). After evaluating each dimension individually, we calculated a combined measure for a comprehensive representation of organizational agility, enabling a more effective analysis of its overall impact on performance during crises. Control variables were measured by single-items and included organizational age, organizational size and the level of cloud service utilization. To assess content validity and engage in scale refinements (MacKenzie et al., 2011), we conducted two rounds of pre-testing without including their responses in the analysis.

Construct (Code)	Definition	Adapted from
Mindfulness Towards Digital Transformation (MDT)	The extent to which an organization proactively identifies, strategically incorporates, and informs management about digital transformation opportunities.	(Li et al., 2021)

Organizational Citizenship Behavior (OCB)	The extent of employees' voluntary actions that go beyond job duties to assist colleagues and enhance organizational functioning.	(Piccolo & Colquitt, 2006)
Loose Coupling (LC)	The extent to which an organization's digital infrastructure components are designed to minimize the impact of internal changes on others elements.	(Tiwana & Konsynski, 2010)
IT Standardization (ITS)	The extent to which organization-wide IT standards, policies and compliance guidelines are in place.	(Weill & Ross, 2004)
Digital Reach (DR)	The extent to which IT services are visible, accessible, and deployable across different functions within the firm and outside its boundaries.	(Mikalef et al., 2021)
Digital Infrastructure Size (DIS)	The extent of an organization's IT assets including people and digital infrastructure.	Self-developed
Organizational Agility (OA)	The combined ability of an organization to proactively and reactively adapt internally to capitalize on emerging market opportunities and demand changes.	(Lu & Ramamurthy, 2011; Panda, 2022)
Comp. Perf. (COMP)	The degree to which an organization performs relative to its competitors.	(Rai & Tang, 2010)

Table 15. Definitions and Sources of Key Constructs

3.2 Data Collection and Analysis

The units of analysis for this international study were organizations headquartered in the DACH region (Germany, Austria, and Switzerland). Therefore our results should also be applicable to other advanced economies, especially within the European Union due to comparable institutional pressures, regulatory environments, and market dynamics. The target participants were alumni of executive academies or MBA programs who hold executive or management positions. An online survey tool was used to collect and process data, and potential respondents were contacted directly through emails and social media requests while being assured of confidentiality and anonymity in accordance with GDPR (EC, 2016). After removing incomplete, suspicious and outlier responses (Hair Jr. et al., 2022, pp. 61-84), 166 valid responses remained, which exceeded the minimum sample size requirements in terms of the ten times rule (Hair Jr. et al., 2022, p. 25). The data were collected during the height of the COVID-19 pandemic in late 2020. We executed the analysis by means of partial least squares structural equation modeling (PLS-SEM), which allowed us to concurrently evaluate the reliability and validity of the measures and estimate relations between constructs (Hair et al., 2019).

3.3 Bias Analysis

We used wave analysis (Van der Stede et al., 2006) to investigate the possibility of bias resulting from nonresponse in our survey based on demographic characteristics including gender, job role, mean age, industry, organizational age, number of full-time employees, and proportion of IT employees, but found no significant differences. To test for common method variance (CMV), we incorporated a single-item exogenous measure (age of respondent) as a theoretically unrelated

marker variable, connected to every endogenous latent variable in the structural model (Nitzl, 2016). We compared the estimated path model relationships with and without this marker variable and found no significant differences. All the theorized pathways maintained their levels of significance, which suggests no threat of CMV.

4 Results

4.1 Sample Demographics

The job roles of the respondents comprise C-suite executives (18.5%), vice Presidents (13.6%), directors (22.8%), and general managers (30.2%). Individual contributors and other roles account for the remaining 14.8%. Functional areas encompass a wide array of expertise, from manufacturing (30.8%) and IT (26.9%) to finance and insurance (10.9%), administrative and support activities (8.3%), and beyond. The organizations' age distribution reveals a focus on incumbents with 86.5% aged 11 years or older and a small percentage (1.8%) representing young start-ups (0-2 years). The sample mainly consists of large organizations (EU guidelines) with 60.4% employing over 250 FTEs.

4.2 Measurement Model

We evaluated the reliability and validity of our reflective measurement constructs following current guidelines (Hair Jr. et al., 2022; Manley et al., 2021). Standardized outer loadings, except ITS3, were significant (> 0.708) and latent variables explained at least 50% of an indicator's variance, indicating indicator reliability. All constructs met the AVE 0.5 threshold for convergent validity. However, we attribute this to the nature of the construct, which consists of three indicators with different scales and is self-developed. We assessed internal consistency reliability using Cronbach's alpha (Cronbach, 1951), composite reliability p_c (Chin, 2010), and Dijkstra-Henseler's rho pA (Dijkstra & Henseler, 2015). All constructs exceed the 0.7 threshold on all three measures (Manley et al., 2021), except for digital infrastructure size, which still shows a good internal reliability due to its high composite reliability pc and correlations with other constructs. Applying the Fornell-Larcker criterion and the heterotrait-monotrait (HTMT) approach showed highly satisfactory discriminant validity, with the square root of each construct's AVE greater than 0.5 and more than its highest correlation with any other construct, and all HTMT values below 0.85.

4.3 Structural Model

The structural model was evaluated by examining the size and statistical significance of the path coefficients. Effect size was assessed using the f^2 values of 0.02, 0.15, and 0.35 as small, medium, and large, respectively. As showed in Table 2, all hypotheses, except H4b, H4c (marginal), and H5 (marginal), were supported. MDT had a large positive effect on LC (ß=0.555, p<0.01) and medium positive effects on STA (ß=0.376, p<0.01) and DR (ß=0.438, p<0.01), supporting H1a-c. MDT had a small positive effect on ITS (ß=0.316, p<0.01) and OA (ß=0.348, p<0.01), supporting H2 and H3, respectively. Additionally, LC had a small positive effect on OA (ß=0.191, p<0.01), supporting H4a. OA had a small positive effect on COMP (ß=0.214, p<0.01), while OCB had a small positive effect on both OA (ß=0.208, p<0.01) and COMP (ß=0.232, p<0.01), supporting H7 and H8, respectively. The structural model accounted for 52% of OA's variance (i.e. $R^2_{adjusted}$=0.52), 30% of LC's variance, 20% of COMP's variance, 19% of DR's variance, 14% of STA's variance, and 9% of ITS's variance. Finally, collinearity was assessed (Hair Jr. et al., 2022, p. 191) and found not to be an issue, and the PLS model's predictive power was confirmed by outperforming the naïve LM benchmark in $_{Q2predict}$ and lower RMSE values for all indicators

(Sarstedt et al., 2021, pp. 36-37). Our examination also unveiled two significant impacts arising from our control variables. Specifically, organizational age manifested a subtle yet significant negative effect on competitive performance (ß=-0.163, p<0.01). In contrast, organizational size revealed a moderate and positive effect on competitive performance (ß=0.345, p<0.01). Finally, cloud utilization had no impact on competitive performance (ß=0.082, p<0.01).

5 Discussion

This study provides a novel integration of organizational capabilities that foster agile crisis responses by considering mindfulness towards digital transformation (MDT) in conjunction with digital infrastructure flexibility and size, and organizational citizenship behavior (OCB) in a major crisis context. To this end, research hypotheses (see Table 2) have been developed and examined based on data drawn from Germany, Austria, and Switzerland.

Besides validating the agility-performance link in the crisis context (H6), we contribute to agility research by demonstrating the importance of two organizational capabilities (MDT, OCB), which have received very limited consideration in prior literature. These can be developed and nurtured to rapidly and effectively respond to dynamic environments to address the next frontier of organizational agility (Tallon et al., 2022). In the crisis context, most previous studies have focused on the role of pre-crisis risk-aversion behavior, which can be implemented, e.g., by business continuity planning (Ameri & Musa, 2021; Herbane et al., 2004), systematic crises preparation (El Idrissi et al., 2022; Fowler et al., 2007) or, more generally, through developing IT resources and capabilities for continuously sensing and adapting to meet new requirements (e.g. Mikalef et al., 2021; Panda & Rath, 2018; Queiroz et al., 2023). The factors associated with organizational agility investigated in this study are different. They refer to organizational capacities that can be developed without having to rely on prepared contingency plans or communication protocols, which are likely to be outdated or limited, considering an unprecedented crisis situation with unforeseeable market changes and opportunities for digitalization. Instead, MDT entails being aware of digital options aligned with the organizational strategy in any given moment of time, which not only affects the organization's digital infrastructure (by making it more flexible and bigger, H1-2), but also fosters its capacity for swift changes, specially crisis-related (H3).

No.	Relationship	Path coeff. (ß)	Eff. size f^2	t-value	Verdict
H1 (a-c)	MDT is positively related with a digital infrastructure flexibility characterized by (a) loose coupling, (b) IT standardization, and (c) digital reach.	(a) 0.555 (b) 0.376 (c) 0.438	0.445 0.165 0.238	8.634*** 5.078*** 6.629***	Supported. (a-c)
H2	MDT is positively related with size of an organization's digital infrastructure.	0.316	0.111	4.900***	Supported.
H3	MDT is positively related with organizational agility.	0.348	0.137	4.388***	Supported.
H4 (a-c)	A more flexible IT infrastructure, in terms of (a) loose coupling, (b) IT standardization, and (c) digital reach, is positively related with organizational agility.	(a) 0.191 (b) 0.090 (c) 0.138	0.033 0.011 0.023	2.115* 1.324^{NS} 1.736^{T}	Supported for (a), not (b, c).

H5	The size of an organization's digital infrastructure is positively related with organizational agility.	-0.107	0.022	1.759^{T}	Not supported.
H6	Organizational agility positively affects competitive performance.	0.214	0.042	2.501*	Supported.
H7	OCB positively affects organizational agility.	0.208	0.064	2.684**	Supported.
H8	OCB positively affects competitive performance.	0.232	0.050	2.647**	Supported.

*$^{T}p<0.10$, *$p<0.05$, **$p<0.01$, ***$p<0.001$, ^{NS} Not significant*

Table 2. Summary of Hypotheses Tests

The significance of the socio-technical dimensions considered in this study also offer a contribution to IS strategy literature, which recognizes the importance of combining or complementing deliberate and emergent strategies (Parnell, 2000; Salmela & Spil, 2002). While organizations can prepare business continuity plans for anticipated incidents (Niemimaa et al., 2019), the necessity to adapt quickly and seize new market opportunities in unexpected ways may require more than solely relying on premeditated strategies (Ye et al., 2022). This calls for an element of emergent response. The importance of nurturing MDT in terms of positively affecting organizational agility on the basis of an intertwinement of the organizational and IS strategy supports this assertion.

Our findings highlight the importance of ensuring digital infrastructure flexibility and avoiding rigidity (Kim et al., 2018) which can act as a foundation for emergent strategy adaptations, or in other words prepare for organizational agility (Galliers, 2006). However, in terms what constitutes the most valuable characteristic of flexible digital infrastructures, our study emphasizes the role of loose coupling (H4a). Loose coupling allows for quickly updating or replacing single components without disrupting the digital infrastructure and consequently operations. The other two considered factors of flexible digital infrastructures, which prior work has considered as parts of a composite construct (Mikalef et al., 2021), do not clearly provide added value for organizational agility (H4b,c) in this study's crisis response context. In terms of digital reach, the positive effects are at least marginally significant ($p<.1$) and in the anticipated direction. However, there is no evidence supporting that IT standardization benefits organizational agility. Our conceptualization and operationalization has considered IT standardization based on organization-wide inbound IT standards (Paik et al., 2017) and compliance requirements. As indicated in section 2, such IT standardization could also lessen flexibility (Ahlemann et al., 2023). Seen as prohibitive global norms (Bernroider, 2021), they are likely to cause tensions between internal and external requirements for organizational changes. Consequently, it would be interesting to pursue further, whether balanced approaches, e.g., those aiming to resolve tensions between IT standardization and differentiation (Gregory et al., 2015), can benefit organizational agility.

While our findings on the effects of digital infrastructure size (H5) on organizational agility are only marginally significant ($p<.1$), they warrant attention, particularly because they indicate negative effects instead of the presumed positive effects. We initially argued for positive effects based on prior research that highlighted gains in automation (Tian et al., 2020) and enhanced business analytics capabilities (Lu & Ramamurthy, 2011). However, our findings lend support to opposing perspectives, suggesting that an increase in digital infrastructure size can lead to

complexity and thus cause greater integration challenges, particularly with legacy IT (Limaj et al., 2020; Wimelius et al., 2021). The resulting digital infrastructure complexity can make swift decisions for market-related changes and even internal operational adjustments harder.
Another main contribution of this study is the recognition of the significant impact that OCB has on enhancing organizational agility (H7) and competitive performance (H8) during a crisis. Whilst OCB has a long history in studies of organizations, IS research has generally neglected the potential of OCB for organizational effectiveness, especially in the context of change (Davison et al., 2020). Our research suggests that in times of crisis, organizations can benefit from employees who are willing to go beyond their formal duties and take the time to explain and solve problems, thereby mobilizing knowledge (Holbeche, 2019) and facilitating required rapid change.

Future work may expand on these findings by investigating the interrelated effects of social, human and structural capital on change performance (Neumannova et al., 2023). To conclude, we need to note key limitations, which include the use of self-reported, cross-sectional data from convenience sampling, translating into a need for replication with alternative designs.

References

Abdelghany, I. K., AlMatar, R., Al-Haqan, A., Abdullah, I., & Waheedi, S. (2024). Exploring healthcare providers' perspectives on virtual care delivery: insights into telemedicine services. *BMC Health Services Research*, *24*(1), 1. https://doi.org/10.1186/s12913-023-10244-w

Ahlemann, F., Dittes, S., Fillbrunn, T., Rehring, K., Reining, S., & Urbach, N. (2023). Managing in-company IT standardization: A design theory. *Information Systems Frontiers*, *25*(3), 1161-1178.

Al-Alawy, K., & Moonesar, I. A. (2023). Perspective: Telehealth – beyond legislation and regulation. *SAGE Open Medicine*, *11*, 20503121221143223. https://doi.org/10.1177/20503121221143223

Al-Omoush, K. S., Simón-Moya, V., & Sendra-García, J. (2020). The impact of social capital and collaborative knowledge creation on e-business proactiveness and organizational agility in responding to the COVID-19 crisis. *Journal of Innovation & Knowledge*, *5*(4), 279-288.

Al Ahbabi, J. M., & Ahmad, S. Z. (2021). Using health-care technology to manage COVID-19 pandemic: a case of Al-Ain city hospital, UAE. *Emerald Emerging Markets Case Studies*, *11*(3), 1-18. https://doi.org/10.1108/EEMCS-10-2020-0356

Aldawoud, F. L., Al-Zaher, F. A., Suhail, Z. H., Al Nasser, Z. H., Almubayidh, S. M., Alzahrani, F. Y., Alluwaif, H. R., Al Qanbar, N. H., & Aljezany, M. N. (2024). Obstacles and enablers for using telehealth services online in nursing practice: A systematic review. *World Journal of Advanced Research and Reviews*, *23*(3), 1742–1750. https://doi.org/10.30574/wjarr.2024.23.3.2846

Allen, J. D., Towne, S. D., Maxwell, A. E., DiMartino, L., Leyva, B., Bowen, D. J., Linnan, L., & Weiner, B. J. (2017). Measures of organizational characteristics associated with adoption and/or implementation of innovations: A systematic review. *BMC Health Services Research*, *17*(1), 591. https://doi.org/10.1186/s12913-017-2459-x

Allison, M. K., Curran, G. M., Walsh, W. A., Dworkin, E. R., & Zielinski, M. J. (2023). Factors Affecting Telemedicine Implementation in Emergency Departments and Nurses' Perceptions of Virtual Sexual Assault Nurse Examiner Consultation for Sexual Assault Survivors. *Journal of Forensic Nursing*, *19*(1), 41-49. https://doi.org/10.1097/JFN.0000000000000385

Ameri, M., & Musa, H. (2021). The impact of business continuity management on the performance of public organizations in UAE. *Journal of Legal, Ethical and Reg. Issues*, *24*, 1-13.

Anthony Jr, B. (2024). Examining the adoption of telehealth during public health emergencies based on technology organization environment framework. *Journal of Science and Technology Policy Management*, *15*(6), 1311-1338. https://doi.org/10.1108/JSTPM-05-2022-0079

Antonacci, G., Benevento, E., Bonavitacola, S., Cannavacciuolo, L., Foglia, E., Fusi, G., Garagiola, E., Ponsiglione, C., & Stefanini, A. (2023). Healthcare professional and manager perceptions on drivers, benefits, and challenges of telemedicine: results from a cross-sectional survey in the Italian NHS. *BMC Health Services Research*, *23*(1), 1115. https://doi.org/10.1186/s12913-023-10100-x

Arokodare, M. A., & Falana, B. R. (2021). Strategic agility and the global pandemic: The agile organizational structure, a theoretical review. *Information Management and Business Review*, *13*(1 (I)), 16-27.

Baker, J. (2012). The Technology–Organization–Environment Framework. In Y. K. Dwivedi, M. R. Wade, & S. L. Schneberger (Eds.), *Information Systems Theory: Explaining and Predicting Our Digital Society, Vol. 1* (pp. 231-245). Springer New York. https://doi.org/10.1007/978-1-4419-6108-2_12

Baldwin, J. R., Pingault, J. B., Schoeler, T., Sallis, H. M., & Munafò, M. R. (2022). Protecting against researcher bias in secondary data analysis: challenges and potential solutions. *Eur J Epidemiol*, *37*(1), 1-10. https://doi.org/10.1007/s10654-021-00839-0

Barbosa da Silva, J., Espinal, M., Garcia-Saiso, S., Fitzgerald, J., Marti, M., Bascolo, E., Haddad, A. E., & D'Agostino, M. (2024). A digital transformation for primary health care. *Bulletin of the World Health Organization*, *102*(1), 2-2A. https://doi.org/10.2471/BLT.23.290726

Barney, A., Buckelew, S., Mesheriakova, V., & Raymond-Flesch, M. (2020). The COVID-19 Pandemic and Rapid Implementation of Adolescent and Young Adult Telemedicine: Challenges and Opportunities for Innovation. *Journal of Adolescent Health*, *67*(2), 164-171. https://doi.org/10.1016/j.jadohealth.2020.05.006

Bateman, T. S., & Organ, D. W. (1983). Job satisfaction and the good soldier: The relationship between affect and employee “citizenship”. *Academy of management Journal*, *26*(4), 587-595.

Bernroider, E. W. N. (2013). Effective ERP adoption processes: the role of project activators and resource investments. *European Journal of Information Systems*, *22*(2), 232-250.

Bernroider, E. W. N. (2021). *Internal IT Control Systems and Global Regulation of Organizations: The Rise of Control Assemblages* International Conference on Information Resources Management (CONF-IRM), Linz, Austria.

Brown, A. D., Kelso, W., Velakoulis, D., Farrand, S., & Stolwyk, R. J. (2023). Understanding Clinician's Experiences with Implementation of a Younger Onset Dementia Telehealth Service. *Journal of geriatric psychiatry and neurology*, *36*(4), 295-308. https://doi.org/10.1177/08919887221141653

Butler, B. S., & Gray, P. H. (2006). Reliability, mindfulness, and information systems. *MIS quarterly*, *30*(2), 211-224.

Butt, Z., Kirsten, L., Beatty, L., Kelly, B., Dhillon, H., & Shaw, J. M. (2022). Barriers and enablers to implementing telehealth consultations in psycho-oncology. *Psycho-oncology*, *31*(8), 1365-1373. https://doi.org/10.1002/pon.5939

Byrd, T. A., & Turner, D. E. (2000). Measuring the Flexibility of Information Technology Infrastructure: Exploratory Analysis of a Construct. *Journal of Management Information Systems*, *17*(1), 167-208.

Byrd, T. A., & Turner, D. E. (2001). An exploratory examination of the relationship between flexible IT infrastructure and competitive advantage. *Information & Management*, *39*(1), 41-52.

Chin, W. W. (2010). How to write up and report PLS analyses. In *Handbook of Partial Least Squares* (pp. 655-690). Springer.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrica*, *16*(3), 297-334.

Davison, R. M., Ou, C. X., & Ng, E. (2020). Inadequate information systems and organizational citizenship behavior. *Information & Management*, *57*(6), 103240.

Dhunnoo, P., Kemp, B., McGuigan, K., Meskó, B., O'Rourke, V., & McCann, M. (2024). Evaluation of Telemedicine Consultations Using Health Outcomes and User Attitudes and Experiences: Scoping Review. *J Med Internet Res*, *26*, e53266. https://doi.org/10.2196/53266

Diel, S., Doctor, E., Reith, R., Buck, C., & Eymann, T. (2023). Examining supporting and constraining factors of physicians' acceptance of telemedical online consultations: a survey study. *BMC Health Services Research*, *23*(1), 1128. https://doi.org/10.1186/s12913-023-10032-6

Dijkstra, T. K., & Henseler, J. (2015). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics & Data Analysis*, *81*, 10-23.

EC. (2016). Directive 95/46/EC (General Data Protection Regulation). *Official Journal of the European Union*, *EU L 119*, 1-88.

El Idrissi, M., El Manzani, Y., Maatalah, W. A., & Lissaneddine, Z. (2022). Organizational crisis preparedness during the COVID-19 pandemic: an investigation of dynamic capabilities and organizational agility roles. *International Journal of Organizational Analysis*, *31*(1), 27-49.

Elali, W. (2021). The Importance of Strategic Agility to Business Survival During Corona Crisis and Beyond. *International Journal of Business Ethics and Governance*, *4*(2), 1-8.

Endres, H., Stoiber, K., & Wenzl, N. M. (2020). Managing digital transformation through hybrid business models. *Journal of Business Strategy*, *41*(6), 49-56.

Etz, R. S., Solid, C. A., Gonzalez, M. M., Britton, E., Stange, K. C., & Reves, S. R. (2023). Telemedicine in Primary Care: Lessons Learned About Implementing Health Care Innovations During the COVID-19 Pandemic. *Annals of Family Medicine*, *21*(4), 297-304. https://doi.org/10.1370/afm.2979

Ewald, H., Klerings, I., Wagner, G., Heise, T. L., Stratil, J. M., Lhachimi, S. K., Hemkens, L. G., Gartlehner, G., Armijo-Olivo, S., & Nussbaumer-Streit, B. (2022). Searching two or more databases decreased the risk of missing relevant studies: a metaresearch study. *Journal of Clinical Epidemiology*, *149*, 154-164. https://doi.org/10.1016/j.jclinepi.2022.05.022

Federal Communications Commission. (2024). *Telehealth, Telemedicine, and Telecare: What's What?* https://www.fcc.gov/general/telehealth-telemedicine-and-telecare-whats-what

Fernández Coves, A., Yeung, K. H. T., van der Putten, I. M., & Nelson, E. A. S. (2022). Teleconsultation adoption since COVID-19: Comparison of barriers and facilitators in primary care settings in Hong Kong and the Netherlands. *Health policy (Amsterdam, Netherlands)*, *126*(10), 933-944. https://doi.org/10.1016/j.healthpol.2022.07.012

Fischbacher-Smith, D. (2017). When organisational effectiveness fails: Business continuity management and the paradox of performance. *Journal of Organizational Effectiveness: People and Performance*, *4*(1), 89-107.

Fowler, K. L., Kling, N. D., & Larson, M. D. (2007). Organizational Preparedness for Coping With a Major Crisis or Disaster. *Business & Society*, *46*(1), 88-103.

Gagnon, M.-P., Desmartis, M., Labrecque, M., Car, J., Pagliari, C., Pluye, P., Frémont, P., Gagnon, J., Tremblay, N., & Légaré, F. (2012). Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. *Journal of Medical Systems*, *36*(1), 241-277. https://doi.org/10.1007/s10916-010-9473-4

Galliers, R. D. (2006). Strategizing for Agility: Confronting Information Systems Inflexibility in Dynamic Environments. In K. C. Desouza (Ed.), *Agile information systems* (pp. 1-15). Routledge.

Garavand, A., Aslani, N., Nadri, H., Abedini, S., & Dehghan, S. (2022). Acceptance of telemedicine technology among physicians: A systematic review. *Informatics in Medicine Unlocked*, *30*, 100943. https://doi.org/10.1016/j.imu.2022.100943

Graham, J. W. (1991). An essay on organizational citizenship behavior. *Employee responsibilities and rights journal*, *4*, 249-270.

Grata-Borkowska, U., Sobieski, M., Drobnik, J., Fabich, E., & Bujnowska-Fedak, M. M. (2022). Perception and Attitude toward Teleconsultations among Different Healthcare Professionals in the Era of the COVID-19 Pandemic. *International Journal of Environmental Research and Public Health*, *19*(18), 11532. https://doi.org/10.3390/ijerph191811532

Gregory, R. W., Keil, M., Muntermann, J., & Mähring, M. (2015). Paradoxes and the nature of ambidexterity in IT transformation programs. *Information Systems Research*, *26*(1), 57-80.

Guenzi, P., & Habel, J. (2020). Mastering the digital transformation of sales. *California Management Review*, *62*(4), 57-85.

Haines, K. J., Sawyer, A., McKinnon, C., Donovan, A., Michael, C., Cimoli, C., Gregory, M., Berney, S., & Berlowitz, D. J. (2023). Barriers and enablers to telehealth use by physiotherapists during the COVID-19 pandemic. *Physiotherapy*, *118*, 12-19. https://doi.org/10.1016/j.physio.2022.09.003

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2-24.

Hair Jr., J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. SAGE.

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.

Harari, M. B., Parola, H. R., Hartwell, C. J., & Riegelman, A. (2020). Literature searches in systematic reviews and meta-analyses: A review, evaluation, and recommendations. *Journal of Vocational Behavior*, *118*, 103377. https://doi.org/10.1016/j.jvb.2020.103377

Hartling, L., Featherstone, R., Nuspl, M., Shave, K., Dryden, D. M., & Vandermeer, B. (2016). The contribution of databases to the results of systematic reviews: a cross-sectional study. *BMC Medical Research Methodology*, *16*(1), 127. https://doi.org/10.1186/s12874-016-0232-1

Hassan, A., Mari, Z., Gatto, E. M., Cardozo, A., Youn, J., Okubadejo, N., Bajwa, J. A., Shalash, A., Fujioka, S., Aldaajani, Z., Cubo, E., & the International Telemedicine Study, G. (2020). Global Survey on Telemedicine Utilization for Movement Disorders During the COVID-19 Pandemic. *Movement Disorders*, *35*(10), 1701-1711. https://doi.org/https://doi.org/10.1002/mds.28284

Hausman, A., & Stock, J. R. (2003). Adoption and implementation of technological innovations within long-term relationships. *Journal of Business Research*, *56*(8), 681-686. https://doi.org/https://doi.org/10.1016/S0148-2963(01)00313-7

Herbane, B., Elliott, D., & Swartz, E. M. (2004). Business continuity management: time for a strategic role? *Long Range Planning*, *37*(5), 435-457.

Holbeche, L. (2019). Designing sustainably agile and resilient organizations. *Systems Research Behavioral Science*, *36*, 668-677.

Hosseini, S. M., Boushehri, S. A., & Alimohammadzadeh, K. (2024). Challenges and solutions for implementing telemedicine in Iran from health policymakers' perspective. *BMC Health Services Research*, *24*(1), 50. https://doi.org/10.1186/s12913-023-10488-6

Iliyasu, B. Z., Iliyasu, Z., Kwaku, A. A., Sani, A., Nass, N. S., Amole, T. G., Abdullahi, H. M., Abdullahi, A. U., Tsiga-Ahmed, F. I., Jibo, A. M., Mbbch, H. A. B., Salihu, H. M., & Aliyu, M.

H. (2024). Acceptability of Teleconsultation Services for HIV Care in Nigeria: A Mixed Methods Study. *Telemedicine and e-Health*. https://doi.org/10.1089/tmj.2024.0196

Jallal, M., Berrada, K., Bouaddi, O., El Badisy, I., Charaka, H., Serhier, Z., Bennani Othmani, M., Nejjari, C., & Khalis, M. (2023). How Physicians in a Moroccan Tertiary Care Center Perceive Teleconsultation During COVID-19 Pandemic? *Telemedicine journal and e-health : the official journal of the American Telemedicine Association*, *29*(2), 284-292. https://doi.org/10.1089/tmj.2022.0094

Khan, J., Jaafar, M., Mubarak, N., & Khan, A. K. (2024). Employee mindfulness, innovative work behaviour, and IT project success: the role of inclusive leadership. *Information Technology and Management*, *25*, 145-159.

Kim, M.-K., Park, J.-H., & Paik, J.-H. (2018). Factors influencing innovation capability of small and medium-sized enterprises in Korean manufacturing sector: Facilitators, barriers and moderators. *International Journal of Technology Management*, *76*(3-4), 214-235.

Laub, N., Agarwal, A. K., Shi, C., Sjamsu, A., & Chaiyachati, K. (2022). Delivering Urgent Care Using Telemedicine: Insights from Experienced Clinicians at Academic Medical Centers. *Journal of general internal medicine*, *37*(4), 707-713. https://doi.org/10.1007/s11606-020-06395-9

Li, H., Wu, Y., Cao, D., & Wang, Y. (2021). Organizational mindfulness towards digital transformation as a prerequisite of information processing capability to achieve market agility. *Journal of business research*, *122*, 700-712.

Li, L., Lin, J., Turel, O., Liu, P., & Luo, X. (2020). The impact of e-commerce capabilities on agricultural firms' performance gains: the mediating role of organizational agility. *Industrial Management & Data Systems*, *120*(7), 1265-1286.

Limaj, E., & Bernroider, E. W. N. (2022). A taxonomy of scaling agility. *The Journal of Strategic Information Systems*, *31*(3), 101721.

Limaj, E., Bernroider, E. W. N., & Ivanova, M. (2020). *Facing Legacy Information System Modernization in Scaling Agility in the Banking Industry: Preliminary Insights on Strategies and Non-technical Barriers* 41st International Conference on Information Systems (ICIS), India.

Liu, H., Song, D., & Cai, Z. (2014). *Knowledge management capability and firm performance: the mediating role of organizational agility* PACIS, Chengdu, China.

Lu, Y., & Ramamurthy, K. R. (2011). Understanding the link between information technology capability and organizational agility. *MIS quarterly*, *35*(4), 931-954.

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS quarterly*, *35*(2), 293-334.

Mahrani, S. W., & Alwi, S. (2022). The effect of Organizational Citizenship Behavior on Performance mediated by Job Satisfaction. *International Journal of Management and Education in Human Development*, *2*(04), 769-774.

Manley, S. C., Hair, J. F., Jr, R. I. W., & McDowell, W. C. (2021). Essential new PLS-SEM analysis methods for your entrepreneurship analytical toolbox. *International Entrepreneurship and Management Journal*, *17*, 1805–1825.

Mao, H., Liu, S., Zhang, J., Zhang, Y., & Gong, Y. (2021). Information technology competency and organizational agility: roles of absorptive capacity and information intensity. *Information Technology & People*, *34*(1), 421-451.

Masood, S., Samad, Z., Nadeem, S., & Ghulam, U. (2024). Telemedicine During a Pandemic: Challenges, Opportunities & Way Forward at a Tertiary Care Hospital. *Hospital Topics*, *102*(1), 9-15. https://doi.org/10.1080/00185868.2022.2073304

Mathew, M. J. (2024). Literature search in systematic reviews: How much is good enough? *Clinical Epidemiology and Global Health*, *25*, 101485. https://doi.org/10.1016/j.cegh.2023.101485

Mikalef, P., Pateli, A., & Wetering, R. v. d. (2021). IT architecture flexibility and IT governance decentralisation as drivers of IT-enabled dynamic capabilities and competitive performance: The moderating effect of the external environment. *European Journal of Information Systems*, *30*(5), 512-540.

Milella, F., Minelli, E. A., Strozzi, F., & Croce, D. (2021). Change and Innovation in Healthcare: Findings from Literature. *ClinicoEconomics and Outcomes Research*, *13*(null), 395-408. https://doi.org/10.2147/CEOR.S301169

Mooney, J. G., & Ganley, D. (2006). Enabling Strategic Agility Through Agile Information Systems: The Roles of Loose Coupling and Web Services Oriented Architecture1. In *Agile Information Systems* (pp. 97-109). Routledge.

Nagarathinam, S., Umadevi, R., Grace, A., & Saravanan, R. (2024). A cross-sectional study to assess the barriers to teleconsultations among doctors in Chennai. *Journal of Applied Pharmaceutical Research*, *12*, 29-34. https://doi.org/10.18231/j.joapr.2024.12.1.29.34

Nazir, S., & Pinsonneault, A. (2012). IT and firm agility: an electronic integration perspective. *Journal of the Association for Information Systems*, *13*(3), 2.

Neumannova, A., Bernroider, E. W., & Obwegeser, N. (2023). *Conceptualizing Digital Resilience: An Intellectual Capital Perspective* International Conference on Information Systems (ICIS), Hyderabad, India.

Niemimaa, M., Järveläinen, J., Heikkilä, M., & Heikkilä, J. (2019). Business continuity of business models: Evaluating the resilience of business models for contingencies. *International Journal of Information Management*, *49*, 208-216.

Nittas, V., & von Wyl, V. (2020). COVID-19 and telehealth: a window of opportunity and its challenges. *Swiss medical weekly*, *150*(1920), w20284-w20284.

Nitzl, C. (2016). The use of partial least squares structural equation modelling (PLS-SEM) in management accounting research: Directions for future theory development. *Journal of Accounting Literature*, *37*, 19-35.

Nyrhinen, M. (2008). IT infrastructure: structure, properties and processes. *Sprouts - Working Papers on Information Systems*, *6*(22), 1-24.

Oliveira, T., & Martins, M. R. (2011). Literature Review of Information Technology Adoption Models at Firm Level. *1566-6379*, *14*(1), 110-121. https://academic-publishing.org/index.php/ejise/article/view/389

Osman, M. A., Schick-Makaroff, K., Thompson, S., Bialy, L., Featherstone, R., Kurzawa, J., Zaidi, D., Okpechi, I., Habib, S., Shojai, S., Jindal, K., Braam, B., Keely, E., Liddy, C., Manns, B., Tonelli, M., Hemmelgarn, B., Klarenbach, S., & Bello, A. K. (2019). Barriers and facilitators for implementation of electronic consultations (eConsult) to enhance access to specialist care: a scoping review. *BMJ Global Health*, *4*(5), e001629. https://doi.org/10.1136/bmjgh-2019-001629

Oudbier, S. J., Souget-Ruff, S. P., Chen, B. S. J., Ziesemer, K. A., Meij, H. J., & Smets, E. M. A. (2024). Implementation barriers and facilitators of remote monitoring, remote consultation and digital care platforms through the eyes of healthcare professionals: a review of reviews. *BMJ open*, *14*(6), e075833. https://doi.org/10.1136/bmjopen-2023-075833

Paik, J.-H., Kim, M.-K., & Park, J.-H. (2017). The antecedents and consequences of technology standardizations in Korean IT small and medium-sized enterprises. *Information Technology and Management*, *18*(4), 293-304.

Panda, S. (2022). Strategic IT-Business Alignment Capability and Organizational Performance: Roles of Organizational Agility and Environmental Factors. *Journal of Asia Business Studies*, *16*(1), 25-52.

Panda, S., & Rath, S. K. (2018). Modelling the relationship between information technology infrastructure and organizational agility: A study in the context of India. *Global Business Review*, *19*(2), 424-438.

Parnell, J. A. (2000). Reframing the combination strategy debate: defining forms of combination. *Journal of Applied Management Studies*, *9*(1), 33-54.

Piccolo, R. F., & Colquitt, J. A. (2006). Transformational Leadership and Job Behaviors: The Mediating Role of Core Job Characteristics. *The Academy of Management Journal*, *49*(2), 327-340.

Podsakoff, P. M., & MacKenzie, S. B. (2014). Impact of organizational citizenship behavior on organizational performance: A review and suggestions for future research. *Oraganizational Citizenship Behavior and Contextual Performance*, 133-151.

Pulakos, E. D., Kantrowitz, T., & Schneider, B. (2019). What leads to organizational agility: It's not what you think. *Consulting Psychology Journal: Practice and Research*, *71*(4), 305.

Pussegoda, K., Turner, L., Garritty, C., Mayhew, A., Skidmore, B., Stevens, A., Boutron, I., Sarkis-Onofre, R., Bjerre, L. M., Hróbjartsson, A., Altman, D. G., & Moher, D. (2017). Systematic review adherence to methodological or reporting quality. *Systematic Reviews*, *6*(1), 131. https://doi.org/10.1186/s13643-017-0527-2

Queiroz, M., Tallon, P. P., Coltman, T., & Sharma, R. (2023). Conditional paths to business unit agility: corporate IT platforms and the moderating role of business unit IT autonomy. *European Journal of Information Systems*, *32*(4), 675-694.

Queiroz, M., Tallon, P. P., Sharma, R., & Coltman, T. (2018). The role of IT application orchestration capability in improving agility and performance. *The Journal of Strategic Information Systems*, *27*(1), 4-21.

Rai, A., & Tang, X. (2010). Leveraging IT capabilities and competitive process capabilities for the management of interorganizational relationship portfolios. *Information Systems Research*, *21*(3), 516-542.

Regragui, S., Abou Malham, S., Gaboury, I., Bois, C., Deville-Stoetzel, N., Maillet, L., Savoie, A., & Breton, M. (2023). Nursing practice and teleconsultations in a pandemic context: A mixed-methods study. *Journal of clinical nursing*, *32*(17-18), 6339-6353. https://doi.org/10.1111/jocn.16756

Ross, J., Stevenson, F., Lau, R., & Murray, E. (2016). Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). *Implementation Science*, *11*(1), 146. https://doi.org/10.1186/s13012-016-0510-7

Salmela, H., & Spil, T. A. (2002). Dynamic and emergent information systems strategy formulation and implementation. *International Journal of Information Management*, *22*(6), 441-460.

Sambamurthy, V., Bharadwaj, A., & Grover, V. (2003). Shaping agility through digital options: Reconceptualizing the role of information technology in contemporary firms. *MIS quarterly*, *27*(2), 237-263.

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2021). Partial Least Squares Structural Equation Modeling. In *Handbook of Market Research* (pp. 1-47). Springer.

Savira, F., Sanigorski, A., Namara, K. M., Hensher, M., Peeters, A., & Manias, E. (2024). Perspectives on telehealth implementation in Australia: An exploratory qualitative study with practice managers and general practitioners. *International Journal of Medical Informatics*, *188*, 105473. https://doi.org/10.1016/j.ijmedinf.2024.105473

Senyel, D., Senn, K., Boyd, J., & Nagels, K. (2024). A systematic review of telemedicine for neuromuscular diseases: components and determinants of practice. *BMC Digital Health*, *2*(1), 17. https://doi.org/10.1186/s44247-024-00078-9

Siddiqui, S., Farr, E., Dusto, N., Chen, L., Kocherginsky, M., Skelton, F., Verduzco-Gutierrez, M., & Lee, S. (2023). Telemedicine Use Among Physiatrists During the Early Phase of the COVID-19 Pandemic and Potential for Future Use. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association*, *29*(2), 242-252. https://doi.org/10.1089/tmj.2022.0030

Singh, S. K., Gupta, S., Busso, D., & Kamboj, S. (2021). Top management knowledge value, knowledge sharing practices, open innovation and organizational performance. *Journal of business research*, *128*, 788-798.

Smith, C., Organ, D. W., & Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of applied psychology*, *68*(4), 653.

Soto-Acosta, P. (2020). COVID-19 pandemic: Shifting digital transformation to a high-speed gear. *Information Systems Management*, *37*(4), 260-266.

Steinhauser, S. (2019). Network-Based Business Models, the Institutional Environment, and the Diffusion of Digital Innovations: Case Studies of Telemedicine Networks in Germany. *Schmalenbach Business Review*, *71*(3), 343-383. https://doi.org/10.1007/s41464-019-00076-9

Tafti, A., Mithas, S., & Krishnan, M. S. (2013). The effect of information technology–enabled flexibility on formation and market value of alliances. *Management Science*, *59*(1), 207-225.

Tallon, P. P., Queiroz, M., & Coltman, T. (2022). Digital-enabled strategic agility: The next frontier. *European Journal of Information Systems*, *31*(6), 641-652.

Teixeira da Silva, J. A., & Daly, T. (2024). Against Over-reliance on PRISMA Guidelines for Meta-analytical Studies. *Rambam Maimonides Med J*, *15*(1). https://doi.org/10.5041/rmmj.10518

Theis, R. P., Dorbu, J. I., Mavrodieva, M. E., Guerrero, R. A., Wright, S. E., Donahoo, W. T., Modave, F., Carrasquillo, O., & Shenkman, E. A. (2024). Telehealth Implementation Response to COVID-19 in the OneFlorida+ Clinical Research Network: Perspectives of Clinicians and Health Systems Leaders. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association*, *30*(1), 268-277. https://doi.org/10.1089/tmj.2023.0029

Tian, H., Wang, T., Liu, Y., Qiao, X., & Li, Y. (2020). Computer vision technology in agricultural automation—A review. *Information Processing in Agriculture*, *7*(1), 1-19.

Tiwana, A., & Konsynski, B. (2010). Complementarities between organizational IT architecture and governance structure. *Information Systems Research*, *21*(2), 288-304.

Tornatzky, L. G., Fleischer, M., & Chakrabarti, A. K. (1990). *The Processes of Technological Innovation*. Lexington Books. https://books.google.co.nz/books?id=QWiwAAAAIAAJ

Turner, K., Bobonis Babilonia, M., Naso, C., Nguyen, O., Gonzalez, B. D., Oswald, L. B., Robinson, E., Elston Lafata, J., Ferguson, R. J., Alishahi Tabriz, A., Patel, K. B., Hallanger-Johnson, J., Aldawoodi, N., Hong, Y.-R., Jim, H. S. L., & Spiess, P. E. (2022). Health Care Providers' and Professionals' Experiences With Telehealth Oncology Implementation During the COVID-19 Pandemic: A Qualitative Study. *Journal of Medical Internet Research*, *24*(1), e29635. https://doi.org/10.2196/29635

Van der Stede, W. A., Young, M. S., & Chen, X. C. (2006). Doing Management Accounting Survey Research. In C. S. Chapman, A. G. Hopwood, & M. D. Shields (Eds.), *Handbooks of Management Accounting Research* (Vol. 1, pp. 445-478). Elsevier.

van Dinter, R., Tekinerdogan, B., & Catal, C. (2021). Automation of systematic literature reviews: A systematic literature review. *Information and Software Technology*, *136*, 106589. https://doi.org/10.1016/j.infsof.2021.106589

van Wessel, R. (2010). Standardization & Standards. In *Toward Corporate IT Standardization Management: Frameworks and Solutions* (pp. 12-49). IGI Global.

Weill, P., & Ross, J. W. (2004). *IT governance: How top performers manage IT decision rights for superior results*. Harvard Business Press.

Wimelius, H., Mathiassen, L., Holmström, J., & Keil, M. (2021). A paradoxical perspective on technology renewal in digital transformation. *Information systems journal*, *31*(1), 198-225.

Wright, K., Golder, S., & Lewis-Light, K. (2015). What value is the CINAHL database when searching for systematic reviews of qualitative studies? *Systematic Reviews*, *4*(1), 104. https://doi.org/10.1186/s13643-015-0069-4

Xu, X., Zhang, W., & Barkhi, R. (2010). IT infrastructure capabilities and IT project success: a development team perspective. *Information Technology and Management*, *11*(3), 123-142.

Ye, D., Liu, M. J., Luo, J., & Yannopoulou, N. (2022). How to Achieve Swift Resilience: the Role of Digital Innovation Enabled Mindfulness. *Information Systems Frontiers*, *26*(2), 551-573.

15. THE PULSE OF THE PARIS OLYMPICS: SENTIMENT AND TOPIC MODELLING OF OPENING CEREMONY

Yunus Emre Bulut
Bogazici University
bulut.yemre@gmail.com

Nazim Taskin
Bogazici University
nazim.taskin@bogazici.edu.tr

Abstract

The study explores public sentiment and thematic discourse surrounding the Paris Olympics 2024 opening ceremony through the analysis of YouTube comments. Employing advanced natural language processing methods, the research leverages sentiment analysis using a RoBERTa-based model and topic modelling through Latent Dirichlet Allocation (LDA) and BERTopic. The analysis of over 20,000 comments highlights diverse audience reactions, uncovering themes of cultural

pride, artistic appreciation, and significant polarization on ideological and religious topics. Positive sentiments emphasize inclusivity, creative displays, and national pride, while negative sentiments critique perceived political messaging and religious insensitivity. Neutral observations reflect analytical engagement with the event.
The comparison of LDA and BERTopic reveals both overarching themes and nuanced insights, underscoring the value of complementary methodologies in capturing complex audience dynamics. The findings emphasize the challenges and opportunities of crafting global events that resonate across diverse cultural expectations, offering actionable recommendations for enhancing inclusivity and mitigating controversy in future international ceremonies.

Keywords
Large Language Models, DSRM, Sentiment Analysis, Topic Modelling, LDA, BERTopic, Paris Olympics Opening Ceremony

1. Introduction

1.1 Background, Gap, and Motivation

The rapid development of social media platforms like Meta (Instagram, Facebook), X (formerly Twitter), and YouTube has significantly transformed human interactions. These platforms enable users to share videos, exchange messages, and express opinions with a global audience. Among them, YouTube stands out as the leading video-sharing platform (Zhang, 2023), ranked as the second most visited website globally, with 122 million active daily users (Omnicore, 2023). Users upload approximately 500 hours of video every minute, and over five billion videos are watched each day (Omnicore, 2023). To facilitate content discovery, YouTube organizes videos into distinct categories and offers features, such as likes, dislikes, and comments for user engagement.

This explosion of user-generated content on platforms like YouTube offers a rich, dynamic repository of textual data, encompassing video descriptions, comments, and reviews. Analysing these interactions can reveal critical insights into public sentiments and thematic patterns, making advanced computational tools like topic modelling indispensable. Topic modelling provides a systematic approach to uncoverin hidden themes within large text collections, complementing related techniques such as sentiment analysis (SA). SA, a key area of natural language processing (NLP), focuses on extracting and analysing the sentiments present in user comments using NLP techniques (Hemmatian & Sohrabi, 2019). It involves various phases that are influenced by pre-processing steps, such as normalization, stemming, tokenization, and stop word removal. These steps ensure clean and structured data for subsequent classification of sentiments (Krouska, Troussas, & Virvou, 2016). Additionally, topic modelling techniques, such as Latent Dirichlet Allocation (LDA) and Correlated Topic Models (CTM), rely on assumptions about how topics are generated and their relationships with words. For instance, LDA assumes that documents consist of multiple topics, while CTM accounts for relationships between topics. In this study, BERTopic, an advanced neural topic modelling approach, will be applied alongside LDA to analyse online reviews and identify key themes, offering a comprehensive exploration of both thematic patterns and sentiment dynamics without imposing strict assumptions about topic structure (Mohanna & Basiouni, 2024).

The utility of these techniques becomes particularly compelling in the context of global events like the Paris Olympics 2024, a milestone that not only celebrates athletic excellence but also catalyses vibrant public discourse across digital platforms. Drawing global audiences, the Paris Olympics inspires discussions on inclusivity, representation, and the narratives surrounding athletes' achievements. Social media platforms, such as YouTube provide a space for real-time public engagement, capturing diverse sentiments and opinions as the event unfolds.

This study focuses on analysing public sentiment and thematic discourse within YouTube comments posted under the official video of the Paris Olympics 2024 opening ceremony. By utilizing Latent Dirichlet Allocation (LDA) and BERTopic for topic modelling, combined with sentiment analysis, it aims to uncover evolving public opinions and emerging themes related to inclusivity, representation, and the cultural significance of the event. These methodologies address key gaps in understanding how audience sentiments develop and how thematic narratives evolve throughout such a monumental global event.

Despite extensive research on the societal impact of the Olympics, there is limited understanding of how global audiences engage in real-time discourse during such events on platforms like YouTube (Mehra et al., 2024; Bandyopadhyay & Karmakar, 2024; Mello et al., 2023). To the best of our knowledge, this study is the first to focus specifically on analysing public sentiment and thematic discourse surrounding the Paris Olympics opening ceremony on YouTube. By combining advanced topic modelling techniques with sentiment analysis, this investigation not only offers new insights into public discourse but also contributes to a broader understanding of how global audiences experience and engage with the Olympics.

1.2. Research Objectives

This study seeks to explore public sentiment surrounding the Paris Olympics 2024, specifically as expressed in YouTube comments under the official video of the opening ceremony. The study employs Latent Dirichlet Allocation (LDA) and BERTopic for dual topic modelling to identify dominant and nuanced themes in YouTube discussions due to their high level of success for such tasks (Atagün, Hartoka, & Albayrak, 2021). This approach allows for a comprehensive analysis of topics such as inclusivity, representation, and the celebration of global athletic achievements.

Furthermore, the research examines public attitudes regarding inclusivity and representation, particularly how these themes are reflected in audience sentiments and commentary. By comparing LDA and BERTopic, the study also contributes methodological insights into the effectiveness of topic modelling. The Research questions are formulated as following:

- What does public sentiment reflect throughout key moments of the opening ceremony of Paris Olympics 2024?
- What are the primary themes in YouTube discussions about the Paris Olympics?

2. Literature Review

The extensive data generated on social media platforms has opened unprecedented opportunities for the academic exploration of human beliefs and behaviours. Social media serves as a naturally occurring repository of human interactions, providing valuable insights into patterns of communication and intent (Tyshchuk & Wallace, 2018). Researchers can now investigate human

behaviours from multiple dimensions, including attitudes, social norms, and perceived behavioural control, using advanced computational methods such as natural language processing and social network analysis (Tyshchuk & Wallace, 2018). Advanced data-driven techniques, such as sentiment analysis and topic modelling provide fresh insights into social media discussions. Sentiment analysis identifies subjective aspects within comments through various methods, including natural language processing (NLP) and machine learning (ML) algorithms. Using sentiment analysis techniques, decision-makers can track shifts in public opinion regarding businesses, events, products, services, and solutions on social media platforms. (Durga et al., 2023). These tools provide dashboards that display trends in positive, negative, and neutral sentiments for platforms like X, YouTube, etc., where many users express their views on various topics (Aloufi & Saddik, 2018). Effective sentiment analysis often relies on unsupervised learning techniques (Yoo, 2018). Studies indicate that using pre-trained sentiment libraries can yield reliable results even when analysing comments across diverse domains (Agrawal & Mittal, 2022; Lee, 2017). Numerous methods for text mining have been developed to uncover themes within extensive text datasets, with topic modelling emerging as the most widely used approach (Egger & Yu, 2022). This technique, rooted in statistical modelling and applied in ML and NLP, identifies latent subject themes within a text corpus (Guo, 2017). Popular algorithms for topic modeming include Latent Dirichlet Allocation (LDA), latent semantic analysis, Top2Vec, and BERTopic (Egger & Yu, 2022).

Although topic modeming methods are robust, most research tends to utilize a single model, with LDA being the most common choice (Gallagher, 2017). However, LDA has faced significant criticism regarding its effectiveness in analysing social media data (Egger & Yu, 2022). For instance, it often overlooks co-occurrence relationships when multiple subjects are present in a text (Jaradat & Matskin, 2019). Furthermore, noisy and sparse datasets, which lack essential statistical learning features, pose challenges for LDA (Chen, 2019). Recently, BERTopic has emerged as a more effective alternative in certain contexts, particularly for analysing short and concise textual data from social media platforms (Jeon et al., 2023). Researchers have demonstrated BERTopic's ability to derive meaningful themes from noisy datasets (Jeon, 2023). This method leverages sentence-transformers to extract document embeddings, supporting over 50 languages, and excels at identifying consistent linguistic patterns. Unlike LDA, which provides discrete topic models, BERTopic offers a continuous modeming approach, setting it apart (Egger & Yu, 2022). The model allows researchers to generate significant themes and incorporates hierarchical clustering, which aids in exploring specific topics and their interconnections, enabling a deeper analysis of a broad range of topics (Grootendorst, 2022).

3. Methodology, Design and Implementation

This study adopts the Design Science Research (DSR) methodology to develop the research problem and address the proposed research questions through structured mechanisms, following established guidelines in the field (Hevner & Chatterjee, 2010). DSR plays a crucial role in Information Systems (IS) by identifying emerging IT capabilities that extend beyond their traditional domains. The methodology involves both a process (activity) and a product (artifact), embodying both an action-oriented (verb) and object-based (noun) perspective (Walls, 1992). This dual nature highlights how artifacts are conceptualized and utilized to influence and interact with real-world systems.

In the context of IS research, DSR produces targeted IT artifacts designed to address significant organizational challenges (Hevner, 2004). Fundamentally, any designed object that inherently provides a solution to a well-defined research problem qualifies as a DSR artifact (Peffers, 2007). The application of DSR in this study ensures a systematic approach in designing, evaluating, and refining the analytical models—sentiment analysis, LDA, and BERTopic—to generate actionable insights from the YouTube comments dataset.

This study analysed a total of 20,204 YouTube comments, including replies, collected from official Paris Olympics Opening Ceremony videos hosted on mainstream media platforms. By examining these comments, the study aims to provide insights into public sentiment, highlight key thematic discussions, and explore the broader audience engagement with the event. The inclusion of replies enriches the dataset, allowing for a more comprehensive understanding of conversations and interactions surrounding the ceremony. The methodology consisted of three key components: Sentiment Analysis, Topic Modelling using Latent Dirichlet Allocation (LDA), and Topic Modelling using BERTopic.

3.1 Sentiment Analysis

Sentiment analysis was conducted to classify comments into three categories: Positive, Neutral, and Negative. The cardiffnlp/twitter-roberta-base-sentiment model, a RoBERTa-based transformer fine-tuned on social media text, was used due to its robustness in handling informal language, slang, and abbreviations.

Text preprocessing involved removing URLs, special characters, and excess whitespace. Additionally, a hybrid language detection method was implemented to focus the analysis on English-language comments. This included an initial pass using the langdetect library and a fallback to a zero-shot classification model for ambiguous cases. A negation-handling function was also integrated to manage phrases like “not happy,” ensuring proper interpretation of sentiment polarity by appending prefixes to negated words.

To optimize the process for a large dataset, comments were divided into batches and processed on a GPU to improve efficiency. Tokenization ensured that inputs adhered to the model’s maximum length of 512 tokens, retaining critical information while discarding excessive text. Sentiment outputs were mapped to user-friendly labels, and the final predictions were appended to the dataset for subsequent analysis.

3.2 Topic Modelling with LDA

Topic modelling was performed using Latent Dirichlet Allocation (LDA) to uncover key themes in the comments. This generative probabilistic method identifies latent topics based on word co-occurrence patterns, treating each comment as a mixture of topics. (Atagün, Hartoka, & Albayrak, 2021).

Text pre-processing included removing stopwords, URLs, and non-essential characters. The comments were converted into a document-term matrix using the CountVectorizer, limiting the vocabulary to the most frequent 1,000 words. This ensured a balance between computational efficiency and topic interpretability.

LDA was applied to both the full dataset and sentiment-specific subsets to explore thematic variations across Positive, Neutral, and Negative comments. The number of topics was predefined, typically ranging from five to ten, based on exploratory testing to maintain coherence without oversimplifying the themes. For each topic, the top keywords were extracted, and semi-automated labelling was conducted, supported by manual inspection, to assign descriptive names to the themes.

3.3 Topic Modelling with BERTopic

BERTopic, a transformer-based topic modelling method, was used to complement LDA by leveraging contextual embeddings and advanced clustering techniques. This approach provided deeper insights into the complexity of discussions in the dataset. (Wijanto, 2024).

The BAAI/bge-small-en SentenceTransformer model was used to generate embeddings for the comments, preserving semantic and contextual information. These embeddings were reduced using UMAP to facilitate clustering. HDBSCAN, a density-based clustering algorithm, identified distinct groups of comments, representing topics. Unlike LDA, BERTopic dynamically determined the number of clusters based on the dataset's structure, enabling more adaptable topic generation.

For topic representation, BERTopic employed multiple strategies. KeyBERT-inspired keyword extraction highlighted the most representative terms, while Maximal Marginal Relevance (MMR) ensured diversity in keyword selection. Labels for topics were generated using the LLaMA-2 model, which provided concise, human-readable summaries based on the extracted keywords.

The combination of sentiment analysis and topic modelling provided a comprehensive understanding of public reactions to the Paris Olympics Opening Ceremony. Sentiment analysis categorized comments effectively, while LDA and BERTopic uncovered thematic insights. LDA offered a reliable baseline for broad topic identification, whereas BERTopic captured nuanced and contextually rich themes, making it particularly suited for analysing complex discussions.

4. Evaluation and Results

The analysis of YouTube comments on the Paris Olympics opening ceremony highlights a blend of admiration, cultural pride, and significant polarization among viewers. Many celebrated the representation of national identities, particularly from regions like South Asia, and appreciated the creative and artistic displays. The ceremony served as a platform for cultural expression, with musical performances and visual storytelling standing out as key highlights. However, artistic choices also sparked controversy, with some viewers feeling alienated by themes they perceived as provocative or divisive. These differences reveal the complexities of creating a global event that caters to diverse cultural expectations.

In terms of sentiment distribution, the analysis revealed a predominance of negative sentiment, followed by neutral and then positive responses. Among a total of 20,204 comments, 12,646 were written in English, indicating that over half of the engagement came from English-speaking users.

Polarization was particularly pronounced around issues of ideology, religion, and representation. Discussions around LGBTQ+ themes and perceived political messaging divided audiences, with

some praising inclusivity and others criticizing it as excessive or polarizing. Similarly, religious symbolism sparked debate, as some appreciated its inclusion while others viewed it as mocking or inappropriate. These reactions underscore how the intersection of global values, cultural sensitivity, and creative expression can elicit both admiration and backlash, highlighting the challenges of uniting a global audience through such a high-profile event.

Sentiment	Count
Negative	5.347
Neutral	4.287
Positive	3.012

Table 1: Sentiment Distribution

The sentiment-based analysis of the Paris Olympics opening ceremony reveals a multifaceted narrative, reflecting the event's ability to captivate, inspire, and divide its global audience. By examining negative, positive, and neutral sentiments, distinct patterns emerge, shedding light on how viewers perceive the ceremony through various emotional and analytical lenses.

The positive sentiment highlights a strong sense of admiration for the ceremony's cultural, artistic, and nationalistic elements. Many viewers celebrated the creative displays, particularly the musical performances by iconic artists, as well as the remarkable execution of the event's staging. There was significant pride expressed by audiences from countries like India, Pakistan, and Bangladesh, emphasizing the inclusivity and global appeal of the Olympics. The tone was largely celebratory, emphasizing the event's success in delivering memorable experiences and instilling a sense of admiration in many.

Conversely, the negative sentiment reveals significant polarization, particularly around ideological, cultural, and religious themes. Criticism often centered on perceived political messaging, with some viewers reacting negatively to LGBTQ+ representation, labelling it excessive or inappropriate. Religious elements also sparked controversy, with some feeling that Christianity or spirituality was mocked or disrespected. Additionally, viewers expressed dissatisfaction with the overall execution, describing the event as disorganized, boring, or lacking coherence. These reactions illustrate the challenges of balancing artistic freedom and global cultural sensitivity in a high-profile event.

The neutral sentiment occupies a middle ground, reflecting analytical and observational commentary. Viewers explored specific details of the ceremony, such as comparisons to Eurovision, discussions of cultural symbols, and questions about performances. The tone was inquisitive and reflective, with less emotional investment, showing that many were engaging with the event on an intellectual rather than intuitive level. This sentiment also touched on themes of representation and symbolism, but with a more open-ended approach.

Overall, the synthesis of sentiments paints a picture of a globally impactful yet polarizing event. The ceremony successfully inspired many through its artistic and cultural displays while simultaneously provoking controversy and debate on ideological and religious grounds. Neutral discussions reflect the audience's deep engagement with the ceremony, showing that, whether through admiration, critique, or analysis, the event succeeded in sparking conversation and

connecting people worldwide. This complex web of reactions underscores the dual challenge and opportunity of crafting an event that resonates across diverse global audiences.
The topics from BERTopic reflect a wide range of viewer reactions to the Paris Olympics opening ceremony, capturing both the celebratory and critical aspects of audience engagement. Overall, the themes highlight the ceremony's ability to inspire national pride, spark artistic appreciation, and provoke controversy on sensitive topics.

On the positive side, many topics emphasize admiration for cultural representation and inclusivity. Viewers expressed joy and pride, particularly in discussions about the participation of nations like India and Bangladesh. This underscores the Olympics' role as a global unifier, where all nations receive a spotlight and resonate strongly with their audiences. Artistic performances were also a major highlight, with topics indicating praise for the event's creative elements and emotional moments. These discussions reflect how the ceremony was able to create a sense of unity and wonder for many viewers, aligning with the traditional goals of an Olympic opening.

On the critical side, the topics reveal significant controversy, particularly around religious and ideological themes. Discussions of "Last Supper Mockery" and references to Christianity highlight a strong negative reaction from viewers who felt certain elements were disrespectful or inappropriate. Similarly, the "Freak Show" and "Cringe Culture" topics indicate harsh criticism of the ceremony's tone and artistic choices, with some audiences perceiving it as over-the-top or out of sync with their expectations. These reactions show how the event's creative and symbolic elements, while bold, failed to connect with certain portions of the audience, resulting in sharp polarization.

The analysis from LDA and BERTopic reveals both similarities and differences in how audience reactions to the Paris Olympics opening ceremony are captured. Both methods identify key themes, such as cultural pride, artistic appreciation, and polarization around ideological and religious topics. For example, both approaches highlight the positive engagement from audiences in countries like India and Bangladesh, as well as criticism of perceived religious insensitivity and controversial artistic choices. This alignment underscores the consistency in capturing the overarching sentiments.

However, the approaches differ in granularity and focus. LDA provides broader, more general themes (e.g., LGBTQ+ representation, religious sensitivity, artistic brilliance), while BERTopic breaks down the discussion into more specific and labelled clusters, such as "Last Supper Mockery," "Freak Show," or "Love for Bangladesh." BERTopic's specificity allows for deeper exploration of distinct moments or controversies, whereas LDA excels in grouping broader emotional and thematic categories. Together, these methods provide a comprehensive understanding, balancing detailed event-specific insights with overarching audience sentiments.

5. Discussion

This study conducted sentiment analysis and topic modeling on public responses from YouTube comments regarding the Paris Olympics 2024 opening ceremony. Employing RoBERTa-based sentiment analysis along with LDA and BERTopic for topic modeling, our findings show nuanced sentiment classifications and thematic insights. Here, we contextualize our results by comparing them with those from related studies using similar methodological frameworks

Fig 1: All Comments Wordcloud

.Our sentiment analysis approach, which utilized a RoBERTa-based model specifically fine-tuned for social media data, aligns well with the findings from Arnob et al. (2023). Arnob et al. employed RoBERTa alongside distilBERT and XLNet models on Twitter data related to FIFA World Cup 2022, achieving robust classification accuracy. Similar to their findings, our results demonstrated the capability of RoBERTa to handle informal, colloquial, and short-text data effectively. However, an important difference was that Arnob et al. (2023) incorporated explainability with LIME to elucidate model decisions clearly, something our current study did not integrate. This suggests potential future enhancement by including explainable AI methods to deepen interpretability and improve transparency of sentiment predictions.

When compared to the study by Karki et al. (2024), which explored sentiment dynamics during the FIFA World Cup 2018 using CatBoost for classification and BERT-based topic modeling, several key differences emerge. Karki et al. found significant fluctuations in negative sentiments correlated closely with team losses, highlighting how sports outcomes directly influence fan emotions. In contrast, our analysis revealed ideological and cultural controversies (such as debates surrounding LGBTQ+ representation and religious symbolism) as primary drivers of negative sentiment, rather than event outcomes themselves. The thematic nature of our event—cultural and artistic versus sports competition—could largely explain this discrepancy. Additionally, Karki et al.'s (2024) integration of geographical and temporal segmentation in their analysis allowed for a more granular understanding of sentiment fluctuations, suggesting that incorporating similar segmentation in future Olympic studies might enhance the depth of insights.

Our findings also align closely with those from Bandyopadhyay and Karmakar (2024), who analyzed Olympic-related Twitter data using BERT-based models, achieving high accuracy (99.23%). Their use of advanced deep learning architectures such as CNN, BiLSTM, and BERT demonstrates that transformer-based models are highly effective for capturing nuanced public sentiment regarding Olympic events. Our similar success with RoBERTa reinforces the conclusion that transformer-based models robustly perform sentiment analysis in social media contexts involving global events.

6. Implications

The results of this study provide critical insights for both the organizers of global events like the Paris Olympics and for researchers interested in understanding public discourse in a digital context. The implications highlight the importance of managing public perception, addressing the diverse expectations of a global audience, and leveraging social media data for continuous improvement.

The polarization observed in the sentiment analysis, particularly around ideological and religious themes, underscores the importance of cultural sensitivity. Olympic organizers should consider inclusive storytelling that carefully balances artistic freedom with respect for cultural and religious diversity. Implementing audience feedback loops during the creative process can help pre-emptively identify potentially divisive elements.

The mixed reactions related to controversial themes indicate the need for deeper community engagement initiatives. Organizers could create dialogue opportunities post-event, such as open forums or social media discussions, to address viewer concerns and explain creative choices. These efforts could help bridge gaps in understanding and reduce the negative sentiment that arises from perceived cultural or ideological insensitivity.

7. Future Recommendations

Future research should consider incorporating additional platforms beyond YouTube, such as Instagram, TikTok, or Facebook, to capture a broader spectrum of public sentiment. Different platforms host different demographic segments, and this could provide a more nuanced understanding of audience reactions.

To enhance global representativeness, future studies should also expand the dataset to include multilingual comments. Incorporating non-English data would allow researchers to explore cultural variations in sentiment and uncover insights from diverse linguistic communities.
Another important dimension is the temporal and geographic breakdown of sentiment. Segmenting sentiment data by time (e.g., before, during, and after the event) and location can reveal dynamic patterns and regional sentiment trends. This information could be valuable for researchers seeking to study the evolution of public opinion and for event organizers aiming to tailor content to specific audiences.

In terms of topic modelling, greater transparency in topic labelling is essential. Clarifying how topics are named—whether through manual inspection, automated summarization, or a hybrid approach—will make the process more reproducible and interpretable, thereby strengthening the credibility of findings. Future research should also consider integrating Explainable AI (XAI) techniques. The use of methods such as LIME or SHAP would improve transparency by revealing how models make classification decisions. This would increase trust in model outputs and make the insights more accessible and actionable for both academic and applied audiences, including event organizers, media analysts, and cultural commentators.

Lastly, to complement social media analytics, qualitative methods such as interviews and surveys could also be employed to gather deeper, context-rich insights directly from participants. These

methods would allow researchers to explore motivations, emotional responses, and interpretations that may not be fully captured through text analysis alone.

References

Agrawal, S. R., & Mittal, D. (2022). Optimizing customer engagement content strategy in retail and E-tail: Available on online product review videos. Journal of Retailing and Consumer Services, 67, 102966.

Aloufi, S., & Saddik, A. E. (2018). Sentiment identification in football-specific tweets. IEEE Access, 6, 78609–78621. https://doi.org/10.1109/ACCESS.2018.2885117

Arnob, S. S., Shikder, M. A. A., Ovey, T. A., Rhythm, E. R., & Rasel, A. A. (2023). Analyzing public sentiment on social media during FIFA World Cup 2022 using deep learning and explainable AI. *2023 26th International Conference on Computer and Information Technology (ICCIT)*.

Atagün, E., Hartoka, B., & Albayrak, A. (2021). Topic modeling using LDA and BERT techniques: Teknofest example. Proceedings of the 6th International Conference on Computer Science and Engineering (UBMK), 660–664. https://doi.org/10.1109/ubmk52708.2021.9558988

Bandyopadhyay, I., & Karmakar, R. (n.d.). Deep learning-based sentiment analysis of Olympics tweets. *Department of Computer Science, The University of Burdwan*

Chen, Y. L., Chang, C. L., & Yeh, C. S. (2017). Emotion classification of YouTube videos. Decision Support Systems, 101, 40–50. https://doi.org/10.1016/j.dss.2017.05.014

Chen, Y., Zhang, H., Liu, R., Ye, Z., & Lin, J. (2019). Experimental explorations on short text topic mining between LDA and NMF-based schemes. Knowledge-Based Systems, 163, 1–13.

Durga, S. L. V., Anandhi, R. J., Singh, N., Rana, A., Kalra, R., & Abdullah, S. A. A. (2023). Mining social media data for sentiment analysis and trend prediction. Proceedings of the 2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), 1557-1562. https://doi.org/10.1109/UPCON59197.2023.10434704

Egger, R., & Yu, J. (2022). A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. Frontiers in Sociology, 7, Article 886498.

Gallagher, R. J., Reing, K., Kale, D., & Ver Steeg, G. (2017). Anchored correlation explanation: Topic modeling with minimal domain knowledge. Transactions of the Association for Computational Linguistics, 5, 529–542.

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794.

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent Dirichlet allocation. Tourism Management, 59, 467–483.

Hemmatian, F., & Sohrabi, M. K. (2019). A survey on classification techniques for opinion mining and sentiment analysis. Artificial Intelligence Review, 52(3), 1495–1545. https://doi.org/10.1007/s10462-017-9599-6

Hevner, A., & Chatterjee, S. (2010). Design science research in information systems. In Design research in information systems (pp. 9-22). Springer. Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS quarterly, 75-105.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS quarterly, 75-105.

Jaradat, S., & Matskin, M. (2019). On dynamic topic models for mining social media. In Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining (pp. 209–230). Springer.

Jeon, E., Yoon, N., & Sohn, S. Y. (2023). Exploring new digital therapeutics technologies for psychiatric disorders using BERTopic and PatentSBERTa. Technological Forecasting and Social Change, 186, 122130.

Karki, S., Parasher, S., Divedi, P., Aggarwal, A., & Mittal, A. (2024). Comprehensive analysis of sports discourse: Sentiment dynamics, derogatory speech patterns, and topic modeling during FIFA 2018 World Cup. *Preprint*.

Krouska, A., Troussas, C., & Virvou, M. (2016). The effect of preprocessing techniques on Twitter sentiment analysis. 7th International Conference on Information, Intelligence, Systems and Applications (IISA). https://doi.org/10.1109/IISA.2016.7785373

Lee, C. S., Osop, H., Goh, D. H. L., & Kelni, G. (2017). Making sense of comments on YouTube educational videos: A self-directed learning perspective. Online Information Review, 41(5), 611–625.

Mehra, V., Singh, P., Bharany, S., & Sawhney, R. S. (2024). *Sports, crisis, and social media: A Twitter-based exploration of the Tokyo Olympics in the COVID-19 era*. *Social Network Analysis and Mining, 14*(56). https://doi.org/10.1007/s13278-024-01218-9

Mello, C., Cheema, G. S., & Thakkar, G. (2023). *Combining sentiment analysis classifiers to explore multilingual news articles covering London 2012 and Rio 2016 Olympics*. *International Journal of Digital Humanities, 5*, 131–157. https://doi.org/10.1007/s42803-022-00052-9

Mohanna, S., & Basiouni, A. (2024). Consumer's cognitive and affective perceptions of artificial intelligence (AI) in social media: Topic modelling approach. Journal of Electrical Systems, 20(3), 1317–1326. https://doi.org/10.52783/jes.3539

Omnicore. (n.d.). YouTube by the numbers: Stats, demographics & fun facts. Retrieved from https://www.omnicoreagency.com/youtube-statistics/

Peffers, 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.

Tyshchuk, Y., & Wallace, W. A. (2018). Modeling human behavior on social media in response to significant events. IEEE Transactions on Computational Social Systems, 5(2), 444–455. https://doi.org/10.1109/TCSS.2018.2815786

Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (1992). Building an information system design theory for vigilant EIS. Information systems research, 3(1), 36-59.

Wijanto, M. C., Widiastuti, I., & Yong, H.-S. (2024). Topic modeling for scientific articles: Exploring optimal hyperparameter tuning in BERT. International Journal on Advanced Science, Engineering and Information Technology, 14(3), 912-919.

Yoo, S., Song, J., & Jeong, O. (2018). Social media contents-based sentiment analysis and prediction system. Expert Systems with Applications, 105, 102–111.

.

16. THE ROLE OF AI-POWERED PERSONALISATION THROUGHOUT THE PURCHASE DECISION-MAKING PROCESS ON ONLINE MARKETPLACES

Nodumo Tembani
University of Cape Town
nodumo.tembani@alumni.uct.ac.za

Maureen Tanner
University of Cape Town
mc.tanner@uct.ac.za

Papama Mtambeka
University of Cape Town
papama.mtambeka@uct.ac.za

Abstract

This study explores the role of Artificial Intelligence (AI)-powered personalisation tools in facilitating consumers' purchase decision-making process within the online shopping

environment. Recently, there has been a rapid growth in online shopping, and retailers have adopted AI-powered personalisation tools such as recommendation systems, chatbots and dynamic pricing to assist and support their consumers with the purchase decision. The techniques and tools required to achieve personalisation have been studied, however, their influence and supportive role during the consumer decision-making journey remains unexplored. The research addresses this gap by assessing how AI-powered personalisation tools facilitate the purchase decision from the problem recognition phase to the post-purchase phase. The findings from the research suggest that AI-powered personalisation can provide valuable support throughout four of the five phases of the consumers' decision-making process through tools AI-powered personalised recommendations, AI-driven contextual recommendations and comparison tools, AI-driven comparison analysis tools, AI-driven personalised reminders and in-cart recommendations. However, tools like AI-driven chatbots and Dynamic pricing are deemed as less relevant by consumers. Post-purchase, AI support is not particularly effective for consumers, who find AI-driven chatbots and review requests to be unhelpful in handling complex queries. The findings emphasise the importance of AI-powered personalisation but also raise a need for improvements to align with customer expectations

Keywords
Artificial Intelligence, AI-Personalisation, Decision-Making Process, Online Marketplace.

1. Introduction
Personalisation refers to the provision of goods and services to consumers that fit their specific requirements (Huang & Zhou, 2019). This technique is particularly prevalent in the online marketplace, where consumers can be offered a convenient shopping experience as well as a variety of choices and product-related information. Recently, online retailers have provided consumers with even more convenience and efficiency by introducing Artificial Intelligence-based (AI) tools called product recommendation agents. These relate to AI-enabled personalisation tools that give individualised suggestions to consumers based on their requirements and specifications (Xiao & Benbasat, 2018). These tools can anticipate consumers' needs, provide tailored product recommendations, and optimise pricing, yet their influence across the entire consumer decision-making process remains unexplored.

The integration of Artificial Intelligence (AI) with personalisation techniques has transformed customers' online shopping experience, allowing for effective marketing initiatives to be actioned at the most strategic time (Gao & Liu, 2023). Using tools such as recommendation agents, AI-powered personalisation can tailor product recommendations to an individual's browsing and purchase history. This is achieved by collecting consumer data and manipulating and translating this data to find opportunities that meet consumers' wants and needs (Mashood et al., 2023). Therefore, AI-driven personalisation can play a significant role in consumers' decision-making (Mohsin, 2024). It is therefore not surprising that e-commerce businesses are allocating a significant number of resources towards the use of AI-powered personalisation.

Given the advent of AI-personalisation and the significant rise in consumer purchases online, it is important to understand consumer decision-making in this context (Mohsin, 2024). For instance, Kotyrlo et al. (2024) report that there is a low level of trust in AI technologies among online

consumers, caused mainly by concerns around security and data privacy. They also report that consumers are often not aware of the practical applications of AI in that context. Although there is ample literature on the technical capabilities of AI within the e-commerce sector, there is a gap in available research on how AI-powered personalisation technologies facilitate decision-making. Few studies address how consumers perceive and interact with these technologies throughout the decision-making process. It is therefore important to understand the role that AI personalisation plays in facilitating purchase decisions across the different stages of the consumer purchase decision-making journey.

To address this gap, the study investigates how AI technologies support consumers in their decision-making journey from the need awareness to post-purchase. The following research question has been formulated to address the research problem: How is AI-powered personalisation used to support the consumer purchase decision-making process?

Using a qualitative approach comprising of semi-structured interviews of online shoppers, the study investigated the role of AI-powered personalisation throughout the five stages of consumer decision-making namely: Problem Recognition, Information Search, Evaluation of Alternatives, Purchase Decision and Post Purchase Evaluation.

The paper contributes to practice by providing valuable insight to businesses looking to enhance their e-commerce marketplaces through AI-driven personalisation strategies. The paper does this through its in-depth analysis of AI technologies, highlighting benefits and key limitations, particularly regarding the effectiveness of chatbots in post-purchase support. It also contributes to existing literature by bridging the gap where the influence of AI-personalisation on the consumer decision-making process is underexplored. This gap is addressed by assessing the influence across all five stages of the Engel-Kollat-Blackwell (EKB) model.

The paper is structured as follows. Section 2 is the literature review, providing an overview of AI technologies, how AI enables personalisation, personalisation techniques, and their role in consumer decision-making. Section 3 details the research design, while Section 4 discusses the findings derived from qualitative interviews. The paper is concluded in section 5, summarizing key contributions, limitations and suggestions for future research.

2. Literature Review

With the many technological advances that have been introduced in today's digital era, consumers' decision-making in the online marketplace has undergone major changes. Online retailers are now delivering individualised and personalised recommendations, content and promotions tailored to preferences using technology like AI (Yoon & Lee, 2021). In this section, we will delve into the intricacies of AI-enabled personalisation, exploring its mechanisms and implications. We also examine existing research on the impact of AI-personalisation on consumer decision-making, providing insights into the evolving landscape of online marketplaces.

2.1 Artificial Intelligence

AI is frequently defined as a system's capacity to accurately understand external data, assimilate knowledge from it, and use these insights to accomplish predefined objectives and tasks (Haenlein & Kaplan, 2019). Recently, there has been an increased interest in AI technologies, together with

data analytics, due to the need for companies to be increasingly competitive (Ameen, Tarhini, Reppel, & Anand, 2021). The technologies relevant to AI-powered personalisation relate to Machine Learning (ML) for recommendation systems, Deep Learning (DL) for enhancing recommendation systems and Natural Language Processing (NLP) for real-time assistance chatbots.

AI encompasses a range of technologies that enable machines to process information, learn from data, and make autonomous decisions. These models learn from the data and assist organisations with predictive analytics (Pallathadka et al., 2023). ML algorithms evolve through repeated exposure to data, enhancing their accuracy in predicting outcomes (Castillo & Taherdoost, 2023). The algorithms find patterns and trends in large datasets, training the machine to make recommendations (Helm et al., 2020). In e-commerce, the algorithms identify patterns in the behaviours of consumers and generate personalised recommendations based on browsing history and purchase behaviour, thereby improving user experience and engagement (Nwanna et al., 2025).

A more advanced subset of ML, which is DL, uses multilayered neural networks to refine recommendations, enhance chatbot interactions, and improve predictive analysis (Rahman et al., 2024). ML consists of algorithms that continuously improve performance by identifying patterns in large datasets and training the machine to make autonomous recommendations (Helm et al., 2020). Multilayered neural networks are artificial neural networks resembling the human brain that consist of neurons that can process and transmit information Unlike traditional ML, which does a multistep feature extraction and classification, DL integrates these processes into a single computational model increasing the precision and efficiency of the applications (Kumar et al., 2021). This has led to more sophisticated personalisation strategies with the refined chatbot interactions and optimized recommendation engines enabled by deep learning.

Another critical AI technology applied in e-commerce is NLP, which is the ability of computers to recognise voice or text in human (natural) language instead of mathematical formulas or computers (Bhardwaj et al., 2024). NLP-driven AI systems are essential for chatbots, virtual assistants, and voice-search functionalities, as they can process and respond to customer queries (Al-Ebrahim et al., 2024; Mehta & Devarakonda, 2018). Some examples of NLP applications are language translation and in-text next-word suggestions made available in messaging apps, automated customer support and product recommendations (Mashood et al., 2023). The effectiveness of NLP in consumer interactions depends on how sophisticated the underlying models are and how well they can adapt to diverse linguistic patterns and user preferences (Kufel et al., 2023).

2.2 AI Applications in e-Commerce marketplaces

AI applications in e-commerce have leveraged the advanced techniques enabled by the above-mentioned technologies to personalise user experiences, streamline transactions and enhance decision-making. E-commerce marketplaces have increasingly integrated chatbots powered by NLP and DL to improve the customer experience. Chatbots have helped to facilitate real-time interactions, assisting customers with queries, product recommendations and tracking orders (Antonio et al., 2022; Kediya et al., 2024) In the context of decision-making for an online purchase,

consumers require chatbots to provide recommendations and assistance at an optimal level of simplicity providing a seamless flow in interactions (Silva et al., 2023).

One of the most widely used applications of AI in e-commerce relates to recommendation systems, which makes use of collaborative filtering, content-based filtering, and hybrid approaches for personalising suggestions (Castillo & Taherdoost, 2023; Raji et al., 2024). These systems provide targeted recommendations by analysing consumer behaviours, browsing history and past purchases. In the context of decision-making for an online purchase, recommendation systems are intended to introduce consumers to new products and personalise recommendations to simplify the decision-making process by narrowing the choices (Bhardwaj et al., 2024; Rahman et al., 2024).

Dynamic Pricing is another AI powered application in e-Commerce marketplaces. Dynamic Pricing is powered by ML algorithms and allows for the adjustment of product prices based on consumers' profiles and location, product configuration and time of purchase (Kopalle et al., 2023). The prices are adjusted in real time based on demand, consumer behaviour, and competitor pricing. With dynamic pricing, ML algorithms evaluate market conditions to optimise pricing strategies and recognise consumer patterns and behaviours making personalised offers (Rizwan et al., 2024). An example is how companies like Uber in response to real-time supply and demand dynamics will adjust prices.

2.3 AI-Enabled Personalisation

Personalisation is the tailoring of products and or services to the preferences of individual customers (Huang, & Zhou, 2019). AI-powered personalisation enhances the online shopping experience by offering consumers individualised recommendations, targeted marketing and predictive analytics (Bawack et al., 2022; J. Huang & Zhou, 2019). Online retailers have assisted consumers with even more convenience and efficiency by introducing AI tools called product recommendation agents; these are a type of AI-enabled personalisation tool that gives individualised suggestions to consumers based on their requirements and specifications (Xiao & Benbasat, 2018).

AI-enable personalisation has introduced many benefits to both online retailers and online shoppers. Studies by Huang and Rust (2021) and Nwanna et al. (2025), found that AI-powered personalisation has redefined the consumer experience due to the relevant recommendations provided, increasing consumer engagement, trust, and loyalty. Contextual recommendations which involve providing recommendations during searches provide relevant recommendations enhancing customer satisfaction (Zimmermann et al., 2005).

3. Engel-Kollat-Blackwell (EKB) Model of Consumer Decision-Making

The Engel-Kollat-Blackwell (EKB) model of consumer decision-making in this paper serves as a foundational framework for understanding the stages that consumers go through when making purchase decisions. The model equates shopping by a consumer as a problem-solving process (Engel et al., 1978; Tešić & Bogetić, 2022). It is based on and expands John Dewey's five-stage problem-solving process, applying it specifically to consumer behaviour by incorporating information processing, environmental influences and decision outcomes (Lappeman et al., 2020). The consumer decision-making process in this model is broken into five phases: problem

recognition, information search, evaluation of alternatives, purchase decision, and post-purchase evaluation (Darley et al., 2010; Tešić & Bogetić, 2022).

In the first phase, the consumer receives informational inputs from marketing and other stimuli which get processed and turned into a problem that is recognised (Vishesh, 2020). This *problem recognition* stage is influenced by both internal and external factors. AI-driven recommendation engines, behavioural targeting and personalised notifications play a crucial role in triggering problem recognition (Lappeman et al., 2020; Vishesh, 2020). For example, recommendation systems will make product suggestions that the consumer may not have actively sought, through analysing past behaviour.

After the problem has been recognised, the consumer will evaluate the solution by performing an *information search* through things such as reading reviews, browsing company websites or talking to friends (Virdi et al., 2020). This phase is significantly enhanced and personalised by AI through contextual recommendations and AI assistants. With this information, the consumer compares different products based on different factors such as price, quality, and customer reviews, moving the process to the evaluation of alternatives stage (Lappeman et al., 2020). This phase is often influenced by external persuasion tactics such as AI driven recommendations.

The *evaluation of alternatives* phase is where the consumer questions if this is the correct product or service they require, or if there is something different that could better satisfy their need with the considerations of price, quality, and features (Lappeman et al., 2020). Customers value the process of comparing numerous products by using a variety of information sources to identify the best choice. However, with the abundance of products available online and the expanding number of online merchants, choosing the right item can be a challenging task (Yoon & Lee, 2021). Comparison Analysis and Dynamic Pricing are the relevant AI-enabled personalisation are particularly relevant tools to customers at this stage. Comparison Analysis is when the recommendation system shows the consumer details of similar or complementary products simplifying the evaluation process (Moura, Reis, & Rodrigues, 2021). Dynamic Pricing is the ability for AI applications to charge different rates for the same product to different customers, depending on individualised personal attributes (Enache, 2021).

Once the alternatives have been evaluated, the consumer moves towards the *purchase decision* phase, where they decide whether to buy the product (Virdi et al., 2020; Vishesh, 2020). AI personalisation strategies facilitate seamless transactions and last-minute purchasing nudges to further refine this stage. Some examples are in-cart recommendations where complementary products are suggested (Gao & Liu, 2023). Once the purchase has been done, the process moves to the final stage which is the *post-purchase evaluation* phase. This stage assesses whether the product meets the consumers' needs and plays a critical role in consumer satisfaction, repeat purchases and brand loyalty (Lappeman et al., 2020). This is where AI-powered chatbots handle things like refunds, returns and tracking orders (Silva et al., 2023).

The study by Ashman, Solomon, and Wolny (2015) affirms the adaptability of the model to participatory, technology-driven cultures, making it a relevant and robust framework for analysing consumer decision-making in digital environments. The paper highlights how digital tools shift consumers from passive recipients to active participants in the decision-making journey, and AI-

powered tools are a part of this shift. Therefore, supporting the use of the model as a valid framework in modern, tech-driven contexts such as AI-personalisation in e-commerce.

4. Research Design

A qualitative research approach was employed, using semi-structured interviews to gather insights from online shoppers. Due to the interpretivist stance of the study, a qualitative method is most suitable for exploring subjective consumer experiences. The study took a deductive research approach to align with the EKB consumer decision-making model. A cross-sectional research strategy was implemented, collecting the data at a single point in time. Given the focus on understanding individual perceptions and experiences with AI-driven personalisation, we chose semi-structured interviews as the primary data collection method. A semi-structured approach provided flexibility and a deeper probe into the experiences of the participants while ensuring that the themes were consistent (Adams, 2015).

Sixteen participants aged 18 to 45, from different regions and who are active online shoppers, were selected for the study employing purposive sampling (Hennink & Kaiser, 2022). The interviews were conducted virtually, using Microsoft Teams and Google Meet, with recordings done on the platforms. The interview recordings were then transcribed and analysed to understand how AI tools influenced their shopping behaviour. Each participant was given an acronym ranging from PT01 to PT16.

Data was analysed systematically, focusing on recurring themes in AI-driven consumer decision-making. The thematic analysis, using NVivo software, was used to code and identify patterns across all the responses, categorising the influence of AI at each stage of the consumer decision-making process (Braun & Clarke, 2022). The major themes, such as targeted recommendations and real-time assistance, were grouped according to the phase of the decision-making process they fall into.

Ethical approval was obtained, and informed consent was provided by the participants before the interviews. This ensured that the study adhered to the research integrity principles set by the Commerce Ethics and Research Committee.

5. Findings and Discussion

Findings from this study reveal that AI-powered personalisation plays a significant role during the consumer decision-making process. During the ***problem recognition phase***, the study identified insights surrounding the supportive role of AI in helping consumers recognise a problem or establish a need. For instance, participants reported that AI-powered personalised recommendations like push notifications for discounts or product suggestions, played a crucial role in prompting needs they had not initially considered: *"I liked a post from [online store] with cute sandals and then literally sandals started popping up on my Instagram from other stores, so I ended up wanting a new pair of sandals"* [P02]. This finding aligns with Kotyrlo et al. (2024)'s study which found that AI-powered advertising significantly impacts consumer awareness and purchase intentions. Their study highlights how the ability of AI to anticipate consumer needs leads to increased engagement and a higher likelihood of purchase.

However, AI-powered personalisation can also lead to misidentification, which can generate frustration and annoyance amongst consumers who do not like receiving communications and recommendations that they consider irrelevant. An example is [P06], showing dissatisfaction with the personalised system saying *"My plus size friend sent me a link for something that she wanted once and then I never stopped receiving plus size content. I always get those plus size adverts on my Instagram and on my Facebook – it feels like I'm always being this identified"*. This is in line with Yaiprasert and Hidayanto (2024) who stated that over-personalisation may lead to consumer disengagement.

AI-powered Chatbots were not found to be as effective during the problem recognition phase. Instead, consumers primarily interact with this feature once a need has already been established. Some participants that mentioned they would consult chatbots to confirm specific details about products, but only after becoming interested in an item.

In the ***information search phase***, consumers value AI-driven contextual recommendations that offer comparable products and relevant alternatives when they were searching for specific items. These AI recommendations enhance the search process by providing more options, allowing consumers to evaluate a broader selection of alternatives and make an informed decision: *"They're very good at giving other suggestions that are similar in nature"* [P04]. Contextual recommendations and comparison tools facilitate the information search phase, while dynamic pricing and product comparisons aid in the evaluation of alternatives. For instance, [P09] stated that during the information search phase *"the ability to do a comparison of features is widely appreciated, especially for high involvement purchases like electronics and appliances"*. This reinforces findings from Gao and Liu (2023), who state that AI personalisation refines search relevance making online shopping more efficient.

AI-powered Chatbots offering real-time assistance proved to be both beneficial and frustrating for the consumers interacting with them during the information search phase. While AI-powered Chatbots can help with straightforward inquiries, several participants noted that these tools lacked sufficient knowledge for detailed product comparisons, which diminished their usefulness during in-depth information searches.

In the ***evaluation of alternatives phase***, AI-personalisation techniques like AI-driven comparison analysis tools can provide consumers with the ability to analyse and compare product features, prices and reviews. However, some degree of transparency is required to reassure consumers that their best interests are indeed considered. For instance, some participants did not fully trust these tools, preferring to conduct independent research due to concerns about potential biases in the automated comparisons: *"I do not use those tools, I don't trust them to be unbiased, so I prefer to do my own comparison"* [P011]. This lack of trust showed a need for transparency in the AI-based comparisons, aligning with Necula and Păvăloaia (2023), who found that consumers perceive AI-generated rankings as often favouring sponsored content over organic recommendations.

This line of thinking is also reflected in how consumers perceive the use of dynamic pricing. Many participants reported that they often cleared cookies or adjusted search times to avoid price hikes, highlighting both the effectiveness and perceived manipulative aspect of dynamic pricing. P6 mentioned that when they notice a spike in flight *prices "I know they're trying to trick me because*

maybe I've searched for those flights before and didn't think to clear my cookies". This aligns with the findings by Rizwan et al. (2024), who found that although dynamic pricing algorithms influence consumer agency, they also raise concerns about manipulation. However, Enache (2021), offers a counterargument stating that dynamic pricing is an essential aspect of e-commerce as it benefits consumers by enabling price optimisation based on demand and competition. The findings from this study indicate that while this is true from a business point of view, consumers are increasingly wary of AI-driven price fluctuations.

During the ***purchase decision phase***, AI-driven personalised reminders and AI-driven in-cart recommendations can effectively nudge consumers toward completing transactions. This was particularly true for online platforms supplying essential or urgent products like groceries. However, for non-urgent products, participants tended to ignore reminders, and some even found them annoying, thus proving that relevance of the need is an important contributing factor to the purchase decision.

Concerning in-cart recommendations, participants reported that these suggestions often "*prompted [them] to add extra items to their cart"* [P10], indicating the relevance of personalised upselling and cross-selling in the final purchasing decision. This finding is confirmed by Silva et al. (2023), who found that these strategies promote an increase in the conversion rate by personalising the final purchase experience. However, while this study found AI personalisation to be effective for most participants, some view these AI nudges as aggressive marketing tactics to *"get them to buy"* [P01]. As such, they choose to ignore these recommendations: *"I am strong willed. If it's in my cart, then I'm buying but it will depend on what I can afford. If it's not meant to be its not meant to be. If I lose it and they restock I will go for it. But their recommendation generally wont sway my decision"* [P01].

The findings reveal that the use of AI-driven personalisation techniques during the post-purchase decision stage needs to be reviewed and improved to truly generate consumer satisfaction and loyalty. For instance, participants have indeed received AI-driven review requests, but these have had very limited effectiveness in generating feedback. Most participants indicated that they only reviewed products or services that significantly exceeded or fell short of their expectations: *"Post purchase if I'm going to end up interacting with an AI tool, it's most probably because something went wrong with the service I signed up for or the product I bought isn't working as intended"* [P04]. Some mentioned they will only engage in reviews if it will be to their benefit via a discount or monetary compensation. Based on these findings it is apparent that review requests could be more effective if they were not so *"tedious to complete"* [P07] and possibly included some incentive or reward. Post-purchase support via AI-powered chatbots received mixed reviews. While helpful for simple inquiries, chatbots were often inadequate for handling complex customer service issues, leading to consumer frustration. These findings highlight both the strengths and limitations of that specific AI-powered personalisation technique. While AI effectively supports decision-making in earlier stages, post-purchase engagement requires further refinement to meet consumer expectations.

6. Limitations and Future Research

The study posed two potential limitations; Firstly, the sample size was small and only limited to a specific demographic, i.e. Ages 20-44. This has the potential of limiting the generalisability of the

findings to larger populations, consistent with interpretivist aims. Secondly, there was a large reliance on the opinions of the participants, which could introduce an element of bias in the findings. A deeper look into AI tools and their influence on consumer behaviours and consumer satisfaction using a larger, more diverse sample is recommended.

Lastly, a study on personalisation in post-purchase support interactions, assessing how AI personalisation can be successfully integrated into post-purchase support is needed. While AI-driven tools support problem recognition, information search, evaluation of alternatives and purchase decisions, improvements are needed in post-purchase support to ensure continued consumer satisfaction. Future research should explore ways to refine chatbot interactions and address trust concerns related to AI-driven personalisation.

7. Conclusion

This study was conducted to provide an understanding of the supportive role that AI-powered personalisation techniques could play across the consumer decision-making process. The study offered an analysis of AI techniques applied to each stage of decision-making process, giving insights into how these techniques can affect the shopping experience of online shoppers.

The findings indicate that AI-powered personalisation techniques can play a significant role in four of the five stages of the Engel-Kollat-Blackwell (EKB) consumer decision model. AI-powered personalised recommendations in the form of push notifications and behavioural targeting play a supportive role in the problem recognition phase, generating needs that encourage consumers to interact with retailers. However, these techniques may also give rise to product misidentification, which can cause frustration and dissatisfaction amongst consumers. During the information search phase, AI-driven contextual recommendation and comparison tools can assist consumers identify comparable and alternative products. While AI-driven chatbots can be effective for simple queries during the information search phase, this tool is not useful for more complex ones and does not assist consumers during the decision-making phase. During the evaluation of alternatives phase, AI-driven comparison analysis can help consumers compare feature, prices and reviews. However, the use of dynamic pricing is not as useful due to concerns about marketing manipulation. For the purchase decision phase, AI-driven personalised reminders and AI-driven in-cart recommendations are particularly relevant to finalise the transaction. While AI effectively supports decision-making in earlier stages, post-purchase engagement requires further refinement to meet consumers' expectations. AI-driven review requests and AI-driven chatbots were not found to be useful to consumers. Review requests are mostly ignored and post-purchase AI-driven chatbots are seemingly not equipped to effectively resolve consumers' issues.

References

Abrokwah-Larbi, K., & Awuku-Larbi, Y. (2024). The impact of artificial intelligence in marketing on the performance of business organizations: evidence from SMEs in an emerging economy. *Journal of Entrepreneurship in Emerging Economies*, *16*(4), 1090–1117. https://doi.org/10.1108/JEEE-07-2022-0207

Adams, W. C. (2015). Conducting Semi-Structured Interviews. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), *Handbook of Practical Program Evaluation: Fourth Edition* (pp. 492–505). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119171386.CH19

Al-Ebrahim, M. A., Bunian, S., & Nour, A. A. (2024). Engineered Science Recent Machine-Learning-Driven Developments in E-Commerce: Current Challenges and Future Perspectives. *Eng. Sci*, *28*(1), 28. https://doi.org/10.30919/es1044

Antonio, R., Tyandra, N., Nusantara, L. T., Anderies, & Agung Santoso Gunawan, A. (2022). Study Literature Review: Discovering the Effect of Chatbot Implementation in E-commerce Customer Service System Towards Customer Satisfaction. *2022 International Seminar on Application for Technology of Information and Communication: Technology 4.0 for Smart Ecosystem: A New Way of Doing Digital Business, ISemantic 2022*, 296–301. https://doi.org/10.1109/ISEMANTIC55962.2022.9920434

Ashman, R., Solomon, M. R., & Wolny, J. (2015). An old model for a new age: Consumer decision making in a participatory digital culture. Journal of Customer Behaviour, 14(2), 127-146. doi:10.1362/147539215x14373846805743

Bawack, R. E., Wamba, S. F., Carillo, K. D. A., & Akter, S. (2022). Artificial intelligence in E-Commerce: a bibliometric study and literature review. *Electronic Markets*, *32*(1), 297–338. https://doi.org/10.1007/s12525-022-00537-z

Bhardwaj, S., Kangri Vishwavidyalaya, G., Sharma, N., Sharma, K., & Verma, V. (2024). *Enhancing Customer Targeting in E-Commerce and Digital Marketing through AI-Driven Personalization Strategies 4 Enhancing Customer Targeting in E-Commerce and Digital Marketing through AI-Driven Personalization Strategies*. https://doi.org/10.1201/9781003450443-4

Braun, V., & Clarke, V. (2022). Conceptual and Design Thinking for Thematic Analysis. *Qualitative Psychology*, *9*(1), 3–26. https://doi.org/10.1037/QUP0000196

Castillo, M. J., & Taherdoost, H. (2023). The Impact of AI Technologies on E-Business. *Encyclopedia*, *3*(1), 107–121. https://doi.org/10.3390/ENCYCLOPEDIA3010009

Darley, W. K., Blankson, C., & Luethge, D. J. (2010). Toward an integrated framework for online consumer behavior and decision making process: A review. *Psychology and Marketing*, *27*(2), 94–116. https://doi.org/10.1002/MAR.20322

Enache, M. C. (2021). Machine Learning for Dynamic Pricing in e-Commerce. *Annals of Dunarea de Jos University of Galati. Fascicle I. Economics and Applied Informatics*, *27*(3), 114–119. https://doi.org/10.35219/EAI15840409230

Engel, J.F., Blackwell, R.D., & Kollat, D.T. (1978). *Consumer Behavior* (3rd ed.). Dryden Press

Gao, Y., & Liu, H. (2023). Artificial intelligence-enabled personalization in interactive marketing: a customer journey perspective. *Journal of Research in Interactive Marketing*, *17*(5), 663–680. https://doi.org/10.1108/JRIM-01-2022-0023

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. https://doi.org/10.1177/0008125619864925

Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., Spitzer, A. I., & Ramkumar, P. N. (2020). Machine Learning and Artificial Intelligence: Definitions, Applications, and Future Directions. *Current Reviews in Musculoskeletal Medicine*, *13*(1), 69–76. https://doi.org/10.1007/S12178-020-09600-8

Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine*, *292*, 114523. https://doi.org/10.1016/J.SOCSCIMED.2021.114523

Huang, J., & Zhou, L. (2019). The dual roles of web personalization on consumer decision quality in online shopping: The perspective of information load. *Internet Research*, *29*(6), 1280–1300. https://doi.org/10.1108/INTR-11-2017-0421/FULL/PDF

Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, *49*(1), 30–50. https://doi.org/10.1007/S11747-020-00749-9/TABLES/4

K. Mashood, Kayani, H. U. R., A. A. Malik, & A. Tahir. (2023). ARTIFICIAL INTELLIGENCE RECENT TRENDS AND APPLICATIONS IN INDUSTRIES. *Pakistan Journal of Science*, *75*(02). https://doi.org/10.57041/PJS.V75I02.855

Kediya, S., Mohanty, V., Gurjar, A., Chouhan, N., Sharma, R., & Golar, P. (2024). Chatbots in Customer Service: A Comparative Analysis of Performance and Customer Satisfaction. *2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry, IDICAIEI 2024*. https://doi.org/10.1109/IDICAIEI61867.2024.10842848

Kopalle, P. K., Pauwels, K., Akella, L. Y., & Gangwar, M. (2023). Dynamic pricing: Definition, implications for managers, and future research directions. *Journal of Retailing*, *99*(4), 580–593. https://doi.org/10.1016/J.JRETAI.2023.11.003

Kotyrlo, O., Naboka, R., Nestor, V., Tyshko, D., & Panasenko, O. (2024). Ways to use artificial intelligence to improve the personalisation of marketing strategies and improve the effectiveness of communication with consumers. *Multidisciplinary Reviews*, *8*(Special Issue). https://doi.org/10.31893/MULTIREV.2024SPE074

Kufel, J., Bargieł-Łączek, K., Kocot, S., Koźlik, M., Bartnikowska, W., Janik, M., Czogalik, Ł., Dudek, P., Magiera, M., Lis, A., Paszkiewicz, I., Nawrat, Z., Cebula, M., & Gruszczyńska, K. (2023). What Is Machine Learning, Artificial Neural Networks and Deep Learning?-Examples of Practical Applications in Medicine. *Diagnostics (Basel, Switzerland)*, *13*(15). https://doi.org/10.3390/DIAGNOSTICS13152582

Kumar, M., Sharma, R. K., & Sethi, I. (2021). Machine Learning in Image Analysis and Pattern Recognition. *Machine Learning in Image Analysis and Pattern Recognition*, 112. https://doi.org/10.3390/BOOKS978-3-0365-1713-1

Lappeman, J. R., Egan, P., Rightward, G., & Ramogase, T. (2020). Marketing to the South African Consumers. In *Marketing to the South African Consumer*. University of Cape Town Libraries. https://doi.org/10.15641/0-7992-2548-8

Mehta, N., & Devarakonda, M. V. (2018). Machine learning, natural language programming, and electronic health records: The next step in the artificial intelligence journey? *Journal of Allergy and Clinical Immunology*, *141*(6), 2019-2021.e1. https://doi.org/10.1016/J.JACI.2018.02.025

Mohsin, S. K. (2024). The Influence of AI-Driven Personalization on Consumer Decision-Making in E-Commerce Platforms. *Al-Rafidain Journal of Engineering Sciences*, *2*(2), 249–261. https://doi.org/10.61268/AJS54S12

Necula, S.-C., & Păvăloaia, V.-D. (2023). AI-Driven Recommendations: A Systematic Review of the State of the Art in E-Commerce. *Applied Sciences 2023, Vol. 13, Page 5531*, *13*(9), 5531. https://doi.org/10.3390/APP13095531

Nwanna, M., Offiong, E., Ogidan, T., Fagbohun, O., Ifaturoti, A., & Fasogbon, O. (2025). AI-Driven Personalisation: Transforming User Experience Across Mobile Applications. *Artif Intell Mach Learn & Data Sci*, *3*. https://doi.org/10.51219/JAIMLD/maxwell-nwanna/425

Pallathadka, H., Ramirez-Asis, E. H., Loli-Poma, T. P., Kaliyaperumal, K., Ventayen, R. J. M., & Naved, M. (2023). Applications of artificial intelligence in business management, e-commerce and finance. *Materials Today: Proceedings*, *80*, 2610–2613. https://doi.org/10.1016/J.MATPR.2021.06.419

Rahman, M. A., Al-Imran, M., Nabi, N., Pabel, A. H., Md, ✉, Sufian Mozumder, A., Sweet, R., Islam, Z., Nazmul, M., Miah, I., Naznin, R., & Kawsur, M. (2024). Unleashing Deep Learning: Transforming E-commerce Profit Prediction with CNNs Journal of Business and Management Studies Unleashing Deep Learning: Transforming E-commerce Profit Prediction with CNNs. *Article in Journal of Business and Management Studies*. https://doi.org/10.32996/jbms.2024.6.2.12

Raji, M. A., Bukola Olodo, H., Oke, T. T., Addy, W. A., Ofodile, O. C., & Oyewole, A. T. (2024). E-commerce and consumer behavior: A review of AI-powered personalization and market trends. *GSC Advanced Research and Reviews*. https://doi.org/10.30574/gscarr.2024.18.3.0090

Rizwan, A., Shraah, A. Al, Alonazi, W. B., Sanchez-Chero, M., Khokhar, K. S., & Rosas-Prado, C. (2024). Consumer-Driven Decision-Making Service Model Selection Using Contextualization and Personalization for Consumer Electronics in E-Commerce. *IEEE Transactions on Consumer Electronics*, *70*(1), 1664–1671. https://doi.org/10.1109/TCE.2024.3367296

Silva, S. C., De Cicco, R., Vlačić, B., & Elmashhara, M. G. (2023). Using chatbots in e-retailing – how to mitigate perceived risk and enhance the flow experience. *International Journal of Retail and Distribution Management*, *51*(3), 285–305. https://doi.org/10.1108/IJRDM-05-2022-0163/FULL/PDF

Tešić, D., & Bogetić, Z. (2022). Models of Consumer Behaviour: A Literature Review. *Journal of Faculty of Economics Brcko*, *16*(1), 21–30. https://doisrpska.nub.rs/index.php/ZREFB/article/view/10190/9848

Virdi, P., Kalro, A. D., & Sharma, D. (2020). Online decision aids: the role of decision-making styles and decision-making stages. *International Journal of Retail and Distribution Management*, *48*(6), 555–574. https://doi.org/10.1108/IJRDM-02-2019-0068

Vishesh. (2020). Customer decision-making process and the effect of marketing on the final purchase decision. *Journal of Management Research and Analysis*, *5*(3), 304–311. https://doi.org/10.18231/2394-2770.2018.0048

Xiao, B., & Benbasat, I. (2018). An empirical examination of the influence of biased personalized product recommendations on consumers' decision making outcomes. *Decision Support Systems*, *110*, 46–57. https://doi.org/10.1016/J.DSS.2018.03.005

Yaiprasert, C., & Hidayanto, A. N. (2024). AI-powered in the digital age: Ensemble innovation personalizes the food recommendations. *Journal of Open Innovation: Technology, Market, and Complexity*, *10*(2). https://doi.org/10.1016/J.JOITMC.2024.100261

Yoon, N., & Lee, H. K. (2021). AI Recommendation Service Acceptance: Assessing the Effects of Perceived Empathy and Need for Cognition. *Journal of Theoretical and Applied Electronic Commerce Research 2021, Vol. 16, Pages 1912-1928*, *16*(5), 1912–1928. https://doi.org/10.3390/JTAER16050107

17. TRANSFORMING CUSTOMER SERVICE IN E-COMMERCE: THE CASE OF HEPSIBURADA'S BESPOKE PLATFORM

Yavuz Selim Hindistan
Ozyegin University
yavuzselim.hindistan@ozyegin.edu.tr

Tayfun İskender
Hepsiburada
ayfun.iskender@hepsiburada.com

Handan Derya Ercan
Bogazici University
handan.ercan@std.bogazici.edu.tr

Hüseyin Özdemir
Hepsiburada
huseyin.ozdemir@hepsiburada.com

Rıdvan Yıldırım

Hepsiburada
ridvan.yildirim@hepsiburada.com

Abstract
One of the most integral aspects of the rapidly growing e-commerce sector is the effective and timely utilization of communication channels for managing all pre-sales and post-sales services provided to customers. This study examines the transformation of customer service processes within the e-commerce sector, focusing specifically on the case of Hepsiburada, one of the leading e-commerce retailers of Turkiye. The study consists of the analysis of the firm's customer service platform transformation through localization. The authors used a phenomenal single-case study approach. The study findings demonstrate that with proper planning and cross-functional coordination, a tailored solution offers strategic advantages over more generic alternatives in customer service environments that demand speed, context awareness, and seamless omnichannel integration. As a result of this project, operational efficiency, customer service agents, and customer satisfaction results reached the global benchmark levels. Findings offer insights to academics and guidance to practitioners aiming to localize customer service platforms in fast-growing e-commerce markets like Türkiye. Case study outcomes benefit businesses aiming to implement similar customer service architecture to improve operational efficiency, customer loyalty, and long-term profitability in the highly competitive e-commerce industry.

Keywords
e-commerce, customer service, microservice, digital transformation, case study

1. Introduction
Digital transformation is the strategic development of digital business models. It is aimed at generating and capturing value (Kane et al., 2015). The rapid rate of technological advancement compels businesses to adopt continuous transformation cycles to remain competitive and operationally resilient. This transformation extends beyond mere technology adoption, especially in the e-commerce industry. E-commerce has evolved from online activity into a dominant global retail since the 1990s. Global e-commerce sales exceeded USD 6.3 trillion in 2023. It is expected to surpass USD 8 trillion by 2027 (Statista, 2024). There is a fundamental shift in organizational strategies and customer engagement paradigms (Verhoef et al., 2021). These changes are evident in customer service and call center operations, where responsiveness, efficiency, and personalization are critical for success.

Speed, scalability, and tailored customer experiences are important in e-commerce. Therefore, digital transformation has become essential rather than optional. Customer service platforms have transformed from traditional telephony systems to omnichannel ecosystems while aligning with improving customer expectations for seamless, cross-channel interactions (Zendesk, 2024). Global e-commerce leaders such as Amazon, Alibaba, and Walmart have redesigned customer service through the use of AI-driven call routing, predictive analytics, cloud-native architectures, and multilingual chatbots. These innovations lead to reduced response times, higher customer satisfaction, and improved operational continuity (Sprinklr, 2024; Call Center Studio, 2025).

Türkiye has seen rapid growth in its e-commerce market in parallel with global trends. It represents a 115% year-on-year increase (ETİD, 2024). The market further expanded to ₺1.86 trillion with a 132% increase and is expected to grow from USD 10.7 billion in 2022 to USD 15.7 billion by 2027, at a compound annual growth rate (CAGR) of 7.55% (Statista, 2024; GlobeNewswire, 2023). Türkiye is positioned as one of the fastest-growing e-commerce markets in the EMEA region with a tech-savvy population and widespread smartphone usage. The top three market share ranking companies are Hepsiburada, Trendyol, and Sahibinden. (Mordor Intelligence, 2024).

Traditional call centers often face problems due to limitations in scalability and integration across digital platforms. Yet many firms in emerging markets continue to rely on legacy systems. These systems are often monolithic, rigid, and poorly suited to fast-changing customer needs. On the contrary to legacy system, off-the-shelf solutions offer basic functionality but may not support local requirements or integration with national systems (Christiawan, 2016). As a result, companies face bottlenecks in support operations and limited flexibility in service design. Existing studies compare packaged and bespoke systems, usually focusing on cost, usability, and implementation time. However, there is little empirical evidence on how a firm transitions from a globally standardized legacy system to a localized microservices architecture—especially in the context of e-commerce customer services in Türkiye. The literature lacks concrete insights into how architectural choices affect customer-facing processes, how localization efforts shape operational efficiency, and how firms integrate omnichannel design in real-time service delivery. This study addresses these gaps through a single case analysis of Hepsiburada, one of Türkiye's largest e-commerce platforms.

D-MARKET Electronic Services & Trading (Hepsiburada) is one of Türkiye's first and the most prominent e-commerce retailers. This study aims to explore the strategic and operational dynamics behind the transformation of Hepsiburada's customer service infrastructure. In particular, the following research questions are addressed:

RQ1: Why did Hepsiburada decide to replace its globally standardized legacy customer service platform with a locally developed, tailor-made solution?
RQ2: How has this transformation impacted operational efficiency and customer satisfaction within Hepsiburada's customer service operations?

The answers to these questions aim to offer practical insights into how large-scale digital transformation initiatives can be implemented locally to improve customer experience, operational performance, and competitive advantage in the e-commerce industry.

2. Literature Review

Large-scale e-commerce companies face strategic decisions regarding their information systems architecture and customer service operations in today's digitally competitive environment. One of the critical decisions is choosing between globally standardized, off-the-shelf applications and bespoke, tailor-made systems developed in-house to meet specific business needs.

2.1. Off-the-Shelf versus Bespoke Systems in E-commerce

Off-the-shelf platforms are commercially available software packages designed for a broad market. These systems offer advantages such as lower initial investment, rapid deployment, and a

mature feature set (Xu, 2024; Clarity Ventures, n.d.). They benefit from frequent updates, broad support communities, and proven performance across various business contexts. However, enterprises that aim to innovate or differentiate based on customer service capabilities encounter limitations in customization and scalability (Fu et al., 2020).

In contrast, bespoke solutions are developed internally or in collaboration with software vendors to align closely with a firm's unique processes and long-term goals. Teng et al. (2024) highlight how combining end-user models with cloud-based AI services facilitates context-specific service delivery, particularly in environments requiring real-time, intelligent responses. Fu et al. (2020) similarly found that bespoke customer service platforms are better equipped to manage complex inquiry routing and personalized customer interactions. Those features are critical for high-volume, multilingual e-commerce operations.

2.2. Strategic Rationale for Bespoke Development

The transition to tailor-made platforms is often driven by a need for agility, data sovereignty, and limitations associated with legacy systems. Xu (2024), in an analysis of Alibaba's transformation, emphasized that standardized systems frequently have problems delivering the degree of personalization required by today's consumers. In addition, such platforms may not fully comply with national regulations or support localized business practices, especially in emerging markets like Türkiye, where localization is a strategic necessity (Xu, 2024; Christiawan, 2016).

In line with this, Hui et al. (2024) show that mass customization capabilities—such as modular architecture, customer-driven design, and internal flexibility—positively impact organizational sustainability and performance, particularly under competitive pressure and the demands of cross-border e-commerce.

Christiawan (2016) also conducted a comparative assessment that shows while packaged software offers usability and lower upfront costs, bespoke software provides stronger alignment with organizational goals, especially in scenarios requiring integration, scalability, and proprietary process flows. The study evaluates the alternatives across factors such as ownership, delivery time, cost, and long-term sustainability. Similarly, Teng et al. (2024) argue that blending cloud and edge computing within a custom architecture enhances scalability while retaining operational flexibility, critical for dynamic e-commerce environments.

2.3. Performance Impact and Challenges

Bespoke systems often involve higher development costs and resource requirements, despite their strategic advantages. The ICS-Assist implementation at Alibaba (Fu et al., 2020) demonstrated that custom systems can lead to significant improvements in agent efficiency and inquiry resolution times. However, such gains depend heavily on robust data infrastructure and organizational readiness.

Moreover, Teng et al. (2024) highlight that performance improvements from custom platforms are conditional on internal commitment and alignment with long-term digital transformation goals. Bespoke systems must not only be technically sound but also supported by a culture and structure conducive to sustained innovation.

Table 1 summarizes the main pros and cons of bespoke versus global off-the-shelf applications, based on the literature.

Aspect	Bespoke Systems	Global Off-the-Shelf Applications
Customization	Fully tailored to business processes and user needs (Christiawan, 2016; Xu, 2024)	Limited; may require costly customizations (Christiawan, 2016)
Scalability & Flexibility	Easily adaptable to internal growth and local regulations (Hui et al., 2024)	May face integration issues; less agile in adapting to unique needs (Fu et al., 2020)
Implementation Time	Long development lifecycle; requires dedicated teams (Christiawan, 2016)	Rapid deployment; usable out-of-the-box (Teng et al., 2024)
Cost	High upfront investment, but may reduce long-term costs if well aligned (Christiawan, 2016)	Lower initial cost, but high maintenance/license fees and hidden costs are possible
Ownership & IP	Full control and ownership over source code (Christiawan, 2016)	Limited; usually subject to End User License Agreements
Integration	Seamless integration with in-house systems (Teng et al., 2024)	May need additional middleware or adaptors (Fu et al., 2020)
Vendor Lock-in	Avoids dependency on third-party providers (Christiawan, 2016)	Risk of lock-in and service disruption if vendor ends support (Goodcore, 2012)
Support & Maintenance	The internal team maintains the system; and quicker fixes.	The internal team maintains the system; quicker fixes
Innovation Potential	High – enables strategic experimentation and innovation (Xu, 2024)	Limited – confined to the vendor's update roadmap
Risk	Higher risk during development; depends on internal capacity (Christiawan, 2016)	Lower risk if the product is stable and widely adopted

Table 1 - Comparison of Bespoke and Off-the-Shelf Applications (Authors' own work)

3. Method

3.1. Selection of Case Study Method

For this study, an explanatory single case study design was chosen as an empirical inquiry to explore and extend knowledge about the adoption of a local call center platform solution. Yin (2009) defines a case study as a qualitative investigation of a contemporary phenomenon within its real-life context. Siggelkow (2007) identifies three key uses for case studies: motivation, inspiration, and illustration. Given the high volume of online customer interactions in the e-commerce sector, examining customer service transformation in this industry is both representative and insightful for broader digital transformation efforts. Accordingly, we adopted a case study approach following Yin (2009) and incorporated insights from seminal works on single case studies, such as Eisenhardt & Graebner (2007), Siggelkow (2007), and Ozcan et al. (2017). The study was conducted in six main phases: planning, designing, preparing, collecting, analyzing, and sharing results as outlined in Yin (2009).

3.2. Case Selection and Research Design

In the planning phase, we formulated research questions focusing on "how" and "why" inquiries, as these are particularly suited for case studies where the researcher has little control over ongoing

events (Yin, 2009; Eisenhardt, 2007). Yin (2009) describes different case study designs: holistic (single unit of analysis), embedded (multiple units of analysis) single case designs, holistic multiple-case designs, and embedded multiple-case designs. This study employed an embedded single case design focusing on the customer service process of Hepsiburada, which operates a hybrid model that integrates both direct sales and a third-party marketplace. Hepsiburada manages a complex ecosystem involving millions of customers and an extensive product portfolio. They sustain market leadership and foster long-term customer loyalty with efficient handling of customer inquiries, issue resolution, and seamless support (Hepsiburada, 2024). In the literature, single case study selection is justified based on the case's criticality, extremeness, typicality, revelatory power, or longitudinal potential (Yin, 2009). The choice of Hepsiburada is grounded in its typicality and revelatory power with its generalizability for other industries. System design consists of the following three levels.
1-System level: transformation, integration, customization. 2-Organizational level: alignment and user adoption. 3-Customer level: operational performance and customer experience.

3.3. Data Collection

Our study adopts an interpretivist ontological worldview, which assumes the existence of multiple realities, along with a subjectivist epistemology (Rashid et al., 2019). To ensure a structured approach, we developed a case study protocol, listing the required data sources before engaging in data analysis. Data were collected through the project documents, internal reports, and meetings with project owners and relevant executives. We interviewed with four project participants with the following titles: R&D Senior Project Manager (SPM), Head of Software Development (HSD), Software Development Unit Head (SDUH), Developer (DEV) respectively. Interviews were performed both face-to-face and online and took between 40 – 65 minutes. Additional data were collected through e-mail interviews as suggested by Ozcan et al. (2017). We meticulously reviewed secondary sectoral reports and brand-new internal data such as ongoing customer service project documents which are integral to digesting the whole customer service process. As recommended by Yin (2009), we began with the end in mind, outlining the structure of the final report before initiating data collection.

3.4. Data Analysis and Interpretation

Our study follows a phenomenon-driven research approach and aims to provide an insightful and representative transformation journey. Eisenhardt & Graebner (2007) argue that researchers employing this approach should frame their study around the significance of the phenomenon. In this study, the research question was intentionally broad to allow for flexibility, justified by the importance of the phenomenon. The analysis phase followed Yin's (2009) explanation building approach, focusing on holistic themes related to system transformation, organizational alignment, and customer experience. Specifically, we identified how system-level changes (transformation, integration, customization), organizational-level actions (alignment, user adoption), and customer-level outcomes (performance and satisfaction) interacted in the transformation project. Our analysis consists of the processes related to the transformation project, subsequent phases of further customer service projects are considered out of the scope of this study. We iteratively sought additional data sources while refining and sharing findings for feedback (Eisenhardt and Graebner, 2007).

4. Results

4.1. Decision-Making Rationale for Bespoke vs. Off-the-Shelf Solutions

The decision to build a bespoke system, rather than customize an off-the-shelf solution, was informed by internal assessments, competitive analysis, and long-term strategic goals. While market solutions offered speed and feature depth, they often fell short of meeting local regulatory needs, workflow integration requirements, and cultural service expectations.

SDUH stated that the bespoke solution provides the following competitive advantages. 1- Greater control over system customization. "In the old system customization was limited. Now, we've overcome those limits and can integrate as we wish.", 2-Elimination of recurring license fees, 3-Easier compliance with Türkiye-specific data regulations, and 4-Better alignment with Hepsiburada's service philosophy. Moreover, bespoke architecture was found essential by all participants, because it allowed the company to design systems around their unique needs, ensuring better fit, faster adaptation, and long-term scalability.

4.2. System Architecture Transformation

The main theme of the findings centered on changes in the digital architecture within customer service, driven by the need for higher performance outcomes. The SDUH stated that the transformation of Hepsiburada's customer service infrastructure was underpinned by a strategic shift from a monolithic, sequential system to a modular microservice architecture. As illustrated in Figure 1a, the legacy system followed a rigid workflow: a customer inquiry first triggered a ticket creation (#1), followed by a separate data request to the system (#2), and finally the display of relevant transaction data (#3). This linear structure caused latency—particularly under high load—leading to performance bottlenecks and delayed responses.

During the interviews, HSD depicted the project architecture. When a support ticket was assigned to an agent, the agent would begin by asking the customer clarifying questions to understand the nature of the inquiry. At this stage, retrieving the necessary data—such as order history, payment status, or delivery status—depended on a manual request initiated by the agent (Step #2). However, the system's actual performance bottleneck occurred during Step #3, when the data screens were loaded. Since the interface could not load until the requested information was processed, any delay in system response became highly visible and disruptive. This issue was amplified in cases where the customer was unsure about which order they were referring to, provided an incorrect order number, or failed to answer confirmation questions accurately. These uncertainties extended the interaction duration, as agents were forced to repeat queries or wait for alternate data to load. In such scenarios, the dependency between each step and the lack of real-time data access led to significantly longer handling times. The cumulative impact of these delays resulted in longer customer wait queues, increased abandonment rates, and reduced satisfaction for both customers and agents. For support staff, repeated interruptions, lagging screens, and time pressure negatively affected both performance metrics and morale, highlighting the limitations of the legacy system in managing dynamic customer interactions.

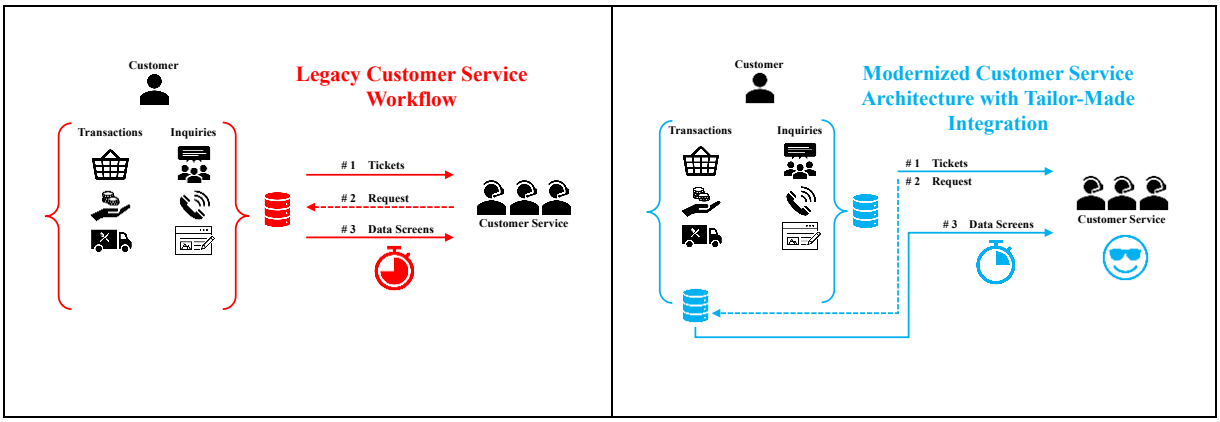

Figure 1: (a)Workflow of Previous System & (b) New System (Authors' own work)

Figure 1b depicts the newly developed system. It is designed around microservice principles, allowing for parallel processing of tickets and data retrieval. By decoupling components and leveraging API-first design, customer information and tickets are prepared simultaneously. This reduces screen load time to under 4 seconds, aligning Hepsiburada's performance with global benchmarks (Zendesk, 2024).

In real terms, this architectural upgrade significantly enhanced both customer and agent experiences. When a customer initiates contact—regardless of whether they provide an accurate order number or clear explanation—the agent is immediately equipped with contextual data, including recent transactions, delivery status, and payment history. Rather than waiting for the system to respond to separate queries, agents now view preloaded, real-time data as soon as the ticket is opened. For instance, even if a customer confuses their orders or is unable to recall specific confirmation details, the agent can seamlessly guide the conversation using the visual history already available on-screen. This eliminates unnecessary back-and-forth and allows the agent to respond more confidently and efficiently. As a result, interaction times are reduced, queue lengths are minimized, and agents can focus on resolution rather than navigating technical constraints, leading to a smoother and more satisfying experience for all parties involved.

Additionally, a one-time data migration process laid the technical foundation for this shift. Over 1TB of data was transferred from the legacy system to a new PostgreSQL environment using API-based extraction. A delta-feeding mechanism was implemented to maintain real-time consistency across both systems during the transition, ensuring business continuity and reporting integrity. This exemplifies enterprise-grade migration practices (Unity Communications, 2025).

Further value was delivered through persistent omnichannel continuity. Inquiries originating from voice, chat, or messaging channels are now processed within a unified pipeline, enabling customers to switch between platforms without repeating their issues. This aligns with global standards in seamless, context-aware support (VoiceSpin, 2024).

Table 2 summarizes the key sentences and areas for improvement from the most impactful technical advancements during this transformation.

Feature	Hepsiburada Previous System	Hepsiburada New System	Global Benchmarks
System Architecture	Private company cloud CRM	Microservice	Microservice (Industry Standards)
Screen Load Time	>17 seconds	Reduced <4 seconds	<5 seconds (Zendesk, 2024)
Information Source Integration	API-based integration with routing	API-based integration with routing	API-centric service meshes
Omnichannel Continuity	Only ticket, screen system, no omnichannel	Seamless transition across voice & text	AI-driven context-aware transitions (VoiceSpin, 2024)
Reporting During Migration	Cloud system	Dual system feed maintained for reporting	Continuous access via data pipelines

Table 2: Comparison of Systems (Authors' own work)

The migration to a microservice-based, parallel-processing system yielded clear and measurable performance benefits. The most immediate improvement was observed in screen load times, which had previously constituted a major bottleneck in the legacy workflow. By enabling real-time access to customer data, the new architecture allowed agents to respond faster and manage tickets more effectively, resulting in smoother and more efficient support operations. Post-deployment analysis revealed substantial gains across multiple key performance indicators:

- A reduction of approximately 20 seconds in minimum customer response time,
- Improvements in both Average Handle Time (AHT) and Monthly Monitoring Analysis (MMA) scores,
- Increased agent efficiency and engagement, enabled by faster interfaces and more streamlined workflows.

These results not only exceeded internal targets but also brought Hepsiburada's performance in line with global customer service benchmarks. According to industry standards reported by Zendesk (2024) and Sprinklr (2024), high-performing contact centers typically achieve screen load times below 4 seconds, AHT between 4 to 6 minutes, and agent efficiency scores exceeding 90%. Hepsiburada's post-transformation metrics indicate alignment with, or even exceed global benchmarks, validating the effectiveness of the bespoke system and its underlying architectural improvements.

4.3. Omnichannel Integration and Channel Customization

As part of the broader transformation, Hepsiburada introduced a role-based announcement system to streamline internal communication. By delivering notifications tailored to user roles (e.g., agents, supervisors, technical teams), the system enhances situational awareness and reduces information overload, especially during peak periods or outages. One of the most innovative developments was the WhatsApp integration as an inbound support channel. Hepsiburada implemented custom bot logic that preserved conversational context across extended interactions.

These bots, optimized for WhatsApp behavior, ensured fluid, human-like exchanges and efficient handoffs to human agents when necessary.

In addition, the company addressed inefficiencies caused by fragmented support systems. The SDUH explained: “We had an issue with duplicate tickets. For example, a customer ordered something on Hepsiburada. Then they called both Hepsiburada and HepsiJet. This created two separate tickets on different systems for the same problem. Two agents ended up working on the same issue separately.”

4.4. Organizational Alignment and User Adaptation

Technical modernization alone could not have guaranteed the project's success without effective organizational alignment. Instead of imposing a top-down switch, the project team adopted a phased rollout strategy. Selected early adopters were onboarded, and their positive experiences with the new platform—such as enhanced UI responsiveness and clearer workflows—created organic momentum for broader adoption.

Status updates were received from agents via feedback screens. We collect both application-specific feedback and requests submitted through the shared communication channel, and in parallel, we also gather and evaluate the emotional states of the representatives. Hepsiburada added a simple daily survey to gauge the emotional state of customer service agents during their shifts. Agents are asked the question "How are you feeling today?" and can choose from six emoji options: (😀 Happy, 😝 Energetic, 😎 Confident, 😐 Indifferent, 😟 Upset, 😡 Angry). They can also select "I don’t want to answer today" if they prefer not to respond.

4.5. Early AI Integration and Customer Experience

In addition to backend efficiency, the new system now supports AI-enriched features that contribute to scalable and personalized customer service. Hepsiburada began collecting conversational data to develop natural language processing (NLP) models in Turkish. Advancements in Turkish language tools have enabled effective disambiguation and parsing with high accuracy (Sak et al., 2008), paving the way for robust AI applications in customer service. Early tests of the platform’s AI layer show promising results, with the intent detection module achieving around 90% accuracy on free-form customer queries. This high level of performance allows the system to interpret requests without relying on rigid, predefined dialogue flows—supporting natural and responsive interactions (Mohammad et al., 2023). Moreover, these AI capabilities feed into agent training, service content optimization, and feedback loops, supporting a self-learning, adaptive service model (Zendesk, 2024). Together, these features move Hepsiburada closer to an intelligent, scalable, and proactive customer service environment.

The AI integration supports customer service quality as follows:

1- The system recognizes customer needs more quickly and with high accuracy.
2- If a phone call is interrupted, the conversation continues seamlessly on WhatsApp, ensuring uninterrupted customer support.
3- The system continuously collects customer feedback and behavioral data, which informs future service improvements and helps increase long-term customer satisfaction and loyalty.
4- The system enables agents to make proactive calls to customers, ensuring faster service through an optimized interface. Specifically, if agents anticipate that the service delivery may

exceed the estimated standard time, they proactively reach out to the customer without waiting for the customer to initiate contact. This approach not only enhances the overall customer experience but also improves service efficiency.

Furthermore, AI and automation are increasingly integral to modern customer service transformation. By 2025, more than 80% of service teams are expected to adopt generative AI (IBM, 2024). These technologies enable intelligent routing, virtual assistants, sentiment analysis, and 24/7 availability—streamlining operations and enhancing personalization. As platforms evolve, the ability to embed and adapt such technologies may further strengthen the case for bespoke solutions in high-growth digital businesses (VoiceSpin, 2024; Unity Communications, 2025).

As a result, Hepsiburada's decision to develop a bespoke system was based on its potential to serve as a typical example for other industries, while revealing challenges and solutions relevant to them. At the system level, the company moved from a slow, sequential legacy system to a modular design that improved integration, allowed for system customization, and reduced delays. At the organizational level, user adoption was supported through phased rollouts, feedback mechanisms, and agent input. At the customer level, performance improved with faster data retrieval, reduced wait times, and seamless support across channels, delivering a more efficient and satisfying customer experience.

5. Discussion

As we consider our results and prior studies (e.g., Christiawan, 2016), we argue that the main reason is that globally standardized solutions offer limited alignment with internal workflows and local regulations. On the other hand, a bespoke system, being developed entirely in-house, offers complete customization and enables the creation of a user interface specifically designed to meet the unique needs and expectations of both agents and customers. Similarly, scalability and flexibility are key concerns. A larger enterprise's growth in customer volume and support complexity requires an appropriate system that can evolve with internal demands and legal frameworks. The legacy system often requires external middleware for integrations. This requires costly support from specific consultants as well as time constraints. Integration needs can be handled easily when the new platform is developed internally. We observed this firsthand in Hepsiburada's seamless omnichannel interactions, which incorporated AI and bot integrations. In terms of implementation time, the tailor-made system requires a longer development cycle and dedicated resources. However, this investment can be justified by long-term gains such as avoiding dependency on third-party providers and gaining full ownership of its platform. Although the initial cost of developing the bespoke system is higher, the company eliminates recurring license fees and reduces hidden costs tied to adapting third-party systems.

Our findings suggest that bespoke systems become a practical choice when companies encounter customization challenges. Although the off-the-shelf solutions enable fast deployment and meet baseline needs, large enterprises such as Hepsiburada can not get effective support for localized regulations or customer service practices.

As a result of these changes, we observed that Hepsiburada's operational efficiency improved across multiple dimensions. Agent productivity increased due to faster interfaces and clearer

workflows. Customer satisfaction was enhanced as customers experienced quicker support and seamless transitions across communication channels, such as WhatsApp and voice calls. In addition, the system's ability to collect user feedback added a new dimension to performance monitoring.

Off-the-shelf customer service applications provide quick setup and broader accessibility, while bespoke systems offer deeper integration, strategic flexibility, and better alignment with enterprise goals, though often at the cost of increased complexity and upfront resource requirements. For e-commerce enterprises in competitive markets like Türkiye, bespoke systems can deliver greater long-term strategic value by enabling differentiated, localized, and high-touch customer experiences. Hepsiburada's experience illustrates that while off-the-shelf solutions may offer short-term convenience, tailor-made systems provide a stronger alignment with business strategy, local context, and service philosophy. The transformation not only yielded measurable performance improvements but also enhanced the organization's internal capabilities and technological autonomy, underscoring the value of investing in bespoke platforms for e-commerce companies in fast-changing, high-volume, and locally nuanced environments.

6. Conclusion

This study aims to showcase the reason behind the transformation of Hepsiburada's customer service infrastructure. We focus on two major concerns; first, why the company made a decision to replace its globally standardized legacy platform with a locally developed, tailor-made solution, and second, how the transformation project impacted the company's operational efficiency and customer satisfaction.

The new system at Hepsiburada opens opportunities for innovation, particularly in areas like natural language processing (NLP), AI-driven recommendations, and sentiment analysis, which are expected to enhance both customer service and employee training. Future plans include tools to prevent fraud, automate the closure of expired requests, and improve operational efficiency. Areas for further research may focus on advancing NLP capabilities, leveraging customer service platforms for product promotion, and balancing automation with interpersonal communication. Ongoing projects in speech recognition, agent performance management, and intelligent communication platforms are poised to provide additional benefits. As chatbot solutions improve, e-commerce users will experience more satisfying interactions, driving client acquisition and market share growth.

This study fills the gap in the literature by understanding the digital transformation journey of a leading e-commerce company, focusing on its shift from multi-purpose, general solutions to custom-developed systems. It contributes to strategies and approaches applicable not only within the e-commerce sector but also across other industries, offering valuable insights for both theoretical literature and managerial practice. However, the study also points out the shortcomings. As a single case study, the results may not be totally generalizable. Quantitative research designs like longitudinal studies should be utilized to collect data and describe the patterns of change to see the causal relationships.

References

Aberdeen, Trudie. "Yin, RK (2009). Case study research: Design and methods. Thousand Oaks, CA: Sage. *The Canadian Journal of Action Research*, 14(1), 69–71.
Call Center Studio. (2025). AI in Turkish Contact Centers. Retrieved from https://www.callcenterstudio.com
Christiawan, E. (2016). Comparative assessment of custom developed and packaged software in accomplishing business's objective [Bachelor's thesis, Murdoch University, Singapore]. ResearchGate. https://www.researchgate.net/publication/330846264
Clarity Ventures. (n.d.). Custom ECommerce Platform vs. Off-the-Shelf. https://www.clarity-ventures.com/articles/custom-ecommerce-platform-vs-off-the-shelf-clarity
Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, 50(1), 25–32.
ETİD (Elektronik Ticaret İşletmecileri Derneği). (2024). Türkiye E-Ticaret Raporu 2023. Retrieved from https://www.etid.org.tr
Fu, M., Guan, J., Zheng, X., Zhou, J., Lu, J., Zhang, T., Zhuo, S., Zhan, L., & Yang, J. (2020). ICS-Assist: Intelligent Customer Inquiry Resolution Recommendation in Online Customer Service for Large E-Commerce Businesses. arXiv preprint arXiv:2008.13534. https://arxiv.org/abs/2008.13534
GlobeNewswire. (2023). Turkey Ecommerce Market Databook 2023. https://www.globenewswire.com
Goodcore Software. (2012). Six Benefits of Custom Software for Your Business. https://www.goodcore.co.uk
Hepsiburada. (2024). Retrieved from https://investor.hepsiburada.com/en/about-hepsiburada/women-entrepreneurs
Hodusoft. (2024). Omnichannel Contact Center Software. Retrieved from https://www.hodusoft.com
Hui, G., Al Mamun, A., Masukujjaman, M., Makhbul, Z. K. M., & Ali, M. H. (2024). The relationship between mass customization and sustainable performance: The role of firm size and global E-commerce. *Heliyon,* 10, e27726. https://doi.org/10.1016/j.heliyon.2024.e27726
IBM. (2024). The State of AI in Customer Service. Retrieved from https://www.ibm.com/think/insights/customer-service-future
Kane, G. C. (2015). Digital Transformation: A Framework for the Future of Work. *MIT Sloan Management Review*, 56(4), 1–12.
Matt, C., Hess, T., & Benlian, A. (2015). Digital transformation strategies. *Business & Information Systems Engineering*, 57(5), 339–343.
Mohammad, R., Favell, O., Shah, S., Cooper, E., & Vakaj, E. (2023). Utilization of open intent recognition models for customer support intent detection. arXiv preprint arXiv:2307.16544.
Mordor Intelligence. (2024). Turkey E-commerce Market - Growth, Trends, and Forecasts. Retrieved from https://www.mordorintelligence.com/industry-reports/turkey-ecommerce-market
Ozcan, P., Han, S., & Graebner, M. E. (2017). Single cases: The what, why, and how. In The Routledge companion to qualitative research in organization studies (pp. 92-112). Routledge.

Rashid, Y., Rashid, A., Warraich, M. A., Sabir, S. S., & Waseem, A. (2019). Case study method: A step-by-step guide for business researchers. *International Journal of Qualitative Methods*, 18, 1609406919862424.
Sak, H., Güngör, T., & Saraçlar, M. (2008). Turkish language resources: Morphological parser, morphological disambiguator, and web corpus. In International Conference on Natural Language Processing (GoTAL 2008) (pp. 417–427). Springer.
Siggelkow, N. (2007). Persuasion with case studies. *Academy of Management Journal*, 50(1), 20–24.
Sprinklr. (2024). The Future of Contact Centers. Retrieved from https://www.sprinklr.com
Statista. (2024). Retail e-commerce sales worldwide from 2014 to 2027. Retrieved from https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales/
StockTitan. (2025). Hepsiburada Announces the Third Bond Issuance of Hepsifinans. Retrieved from https://www.stocktitan.net/news/HEPS
Teng, L., Liu, Y., Liu, J., & Song, L. (2024). End-Cloud Collaboration Framework for Advanced AI Customer Service in E-commerce. arXiv preprint arXiv:2410.07122. https://arxiv.org/abs/2410.07122
Unity Communications. (2025). AI Transformation in BPO Services. Retrieved from https://www.unity-connect.com
Verhoef, P. C. (2021). Digital Transformation and Marketing. Routledge.
VoiceSpin. (2024). Voice and AI for Customer Success. Retrieved from https://www.voicespin.com/blog/call-center-automation-trends/
Xu, H. (2024). Analysis of Digital Transformation Problems of Alibaba E-commerce Platform and Countermeasures to Enhance Personalized Services. https://shsconferences.org/articles/shsconf/pdf/2024/28/shsconf_dsm2024_01003.pdf
Yin, R. K. (2009). Case study research: Design and methods (Vol. 5). Sage.
Zendesk. (2024). CX Trends 2024 Report. Retrieved from https://www.zendesk.com/blog/ai-customer-service-statistics

18R RETHINKING DOCTORAL RESEARCH AND SUPERVISION IN THE GENERATIVE AI ERA: A RESEARCH AGENDA FOR BUSINESS SCHOOLS

Melika Soleimani
Auckland University of Technology
melika.soleimani@aut.ac.nz

Kasuni Weerasinghe
Auckland University of Technology
kasuni.weerasinghe@aut.ac.nz

Maryam Mirzaei
Auckland University of Technology
maryam.mirzaei@aut.ac.nz

David J. Pauleen
National Chung Cheng University
d.pauleen@protonmail.com

Abstract

The integration of Generative AI (GenAI) in higher education has been explored at undergraduate and master's levels, but its impact on doctoral education remains underexamined. This research-in-progress outlines a research agenda to investigate how PhD students and supervisors engage with GenAI, particularly in the context of supervision practices and institutional expectations, to enhance research productivity while upholding ethical and pedagogical standards. It proposes a symbiotic intelligence framework where human reasoning and GenAI capabilities complement each other. This paper reviews existing literature, presents a theoretical foundation, and sets the stage for empirical investigation. It aims to reimagine doctoral education, redefine success metrics, and promote reflective habits and responsible AI engagement.

Key words

Doctoral education, Generative AI (Gen AI), Symbiotic intelligence, Responsible AI integration, Ethics

1. Introduction

The rapid advancement of Generative Artificial Intelligence (GenAI) is transforming research across disciplines and offering new tools and approaches for scholars to enhance their investigative capabilities. (Foroughi et al., 2024). GenAI has the potential to facilitate interactive engagement and generate novel insights that can influence the research process itself (Mishra, Oster, & Henriksen, 2024). Compared to the many studies on undergraduate and master's level education as well as more advanced academic research (Bhullar, Joshi, & Chugh, 2024; Deng, Jiang, Yu, Lu, & Liu, 2025), GenAI's impact on doctoral studies remains underexplored. This gap is particularly significant given that doctoral research demands originality, critical analysis, and ethical rigor—factors that GenAI can both support and potentially undermine. As GenAI technologies evolve, understanding their integration, benefits, limitations, and ethical implications becomes increasingly critical for the academic community.

Overreliance on AI-generated content can undermine scholarly rigor, while outright restrictions may stifle innovation (Sun, Boudouaia, Zhu, & Li, 2024). Universities must strike a balance by fostering AI literacy among doctoral candidates, equipping them with the skills to critically evaluate AI-generated outputs while upholding ethical research practices (Ng, Leung, Chu, & Qiao, 2021). However, existing doctoral training programs often emphasize technical proficiency without sufficiently addressing the intellectual and ethical dimensions of AI-driven research. Furthermore, doctoral supervision often overlooks the development of practical wisdom, or *phronesis*, which fosters critical reflection on the broader social and ethical implications of

research (Intezari, Pauleen, & Rooney, 2016). We suggest an integrative approach can better equip candidates and supervisors to navigate the complex ethical terrain of AI in academia.

To bridge this gap, we present a review of the research-based GenAI literature, including the limited work on doctoral research. We propose the development of a symbiotic intelligence framework (Desolda, Esposito, Lanzilotti, Piccinno, & Costabile, 2024) guided by wisdom principles advocated for PhD study (Intezari, Pauleen, & Rooney, 2016), in which human reasoning and judgement and AI capabilities complement each other. Symbiotic intelligence refers to a mutually beneficial relationship between humans and AI, wherein both parties collaborate closely, leveraging their respective strengths to achieve outcomes neither could accomplish alone (Desolda, Esposito, Lanzilotti, Piccinno, & Costabile, 2024). This collaborative enhancement is what characterizes the framework as "intelligent," as it enables a level of performance and decision-making that surpasses what each could achieve independently (Samala et al., 2024). Wisdom principles in doctoral research refers to the application of ethical, reflective, and practical wisdom in conducting research that not only advances knowledge but also contributes positively to societal well-being (Intezari, Pauleen, & Rooney, 2016). By embedding reflective practices and ethical considerations into AI-assisted research, this approach seeks to ensure scholarly integrity. This research examines how PhD students and supervisors navigate GenAI, identifying the key competencies and institutional policies needed to support responsible AI use in research.

The study will investigate how to wisely integrate GenAI into PhD research by streamlining tasks such as literature reviews, data analysis, and drafting (Lo, Hew, & Jong, 2024; Lin & Chen, 2024), while tackling issues like plagiarism, content authenticity, and misuse of GenAI (Wise, Emerson, Van Luyn, Dyson, Bjork, Thomas, 2024; Tlili et al., 2023). The overarching goal is to understand how socially constructed perceptions, within supervisory relationships and institutional contexts, influence how GenAI is ethically adopted, interpreted, and meaningfully applied in doctoral education. To achieve this, the study explores the following research question: *What ethical, cultural, and academic guidelines should inform the responsible integration of GenAI into PhD research?* The findings will support a framework that strengthens ethical reasoning and reflective practices in doctoral training, preparing researchers for an AI-influenced academic landscape.

2 Background

This section explores the interplay between emerging GenAI technologies, institutional responses from universities and journals, and evolving dynamics in student–supervisor interactions. Together, these elements shape how GenAI is understood, adopted, and governed in doctoral education.

2.1 Technology

GenAI is the latest disruptive force redefining how research is conducted (Henriksen, Mishra, Woo, & Oster 2025). Unlike conventional research tools that primarily retrieve and summarize existing knowledge, GenAI possesses generative capabilities, enabling it to create new content and simulate human-like conversations that facilitate knowledge exchange (Henriksen, Mishra, Woo, & Oster 2025). This interactivity allows researchers to refine their understanding of complex topics through iterative engagement, making GenAI more than just an information retrieval system (Mishra, Oster, & Henriksen, 2024). Its multimodal capabilities (e.g., processing text, speech, and

visual data) expand its application across diverse disciplines (Kasneci et al., 2023). As an epistemic technology, GenAI is becoming indispensable in research workflows, supporting literature reviews, writing support, automation of citations, and assisting with qualitative and quantitative analysis (Henriksen, Mishra, Woo, & Oster 2025; Wang, Wang, & Su 2024). Table 1 presents some widely known researcher support AI tools and their features.

Feature	Elicit	SciSpace	Jenni AI	Perplexity AI	Notebook LLM
AI-powered writing	-		✓ ★		✓ ★
Natural language question-answering	✓ ★	✓ ★	-	✓ ★	✓ ★
Source attribution	✓ ★	✓ ★	✓ ★	✓ ★	✓ ★
Evidence synthesis	✓ ★	-	-	✓ ★	✓ ★
Collaboration tools	✓ ★	★	-	-	★
Interactive notebooks	★	-	-	-	✓ ★
Systematic literature review	★	★	★	-	-
Research question formulation	✓ ★				
Data analysis & visualization	★		-	-	-
Survey creation	★				
Plagiarism checking	-	-	★	-	-
LaTeX & formatting	-	★	-	-	-
✓ Included Free | ★ Paid version only |					

Table 1: Examples of existing research support platforms (as of March 2025)

Table 1 shows most of the free versions of the AI tools support literature reviews with more advanced features available in paid versions. While the features are attractive (e.g., deep academic literature search for systematic literature reviews) the rigor and validity of the results are uncertain. The lack of transparency and explainability raises concerns about the tools' reliability and may hinder both researcher development and study validity.

2.2 Institutional Responses

In response to Gen AI and its emerging impact on research and publishing, Business and Management journals have converged around a core set of ethical guidelines for the use of GenAI in research writing: GenAI must not replace original scholarly work, cannot be credited as an author, and any use must be transparently disclosed. Table 2 summarizes current recommendations for authors on appropriately acknowledging AI usage.

Beyond author responsibilities, major publishers also emphasize guidelines for editors and reviewers essentially instructing them to rely on their expertise rather than AI tools. Maintaining strict confidentiality is also essential, including avoiding uploading manuscript content (fully or partially) to external platforms or services. Universities are also developing policies to manage the rise of GenAI. In Australia and New Zealand, universities are introducing policies that promote transparency and ethical use of GenAI, requiring students and supervisors to declare AI involvement in research and cautioning against inputting sensitive or culturally significant information (Monash University Graduate Research, 2023; The University of Auckland, n.d.).

Criteria / Publishing Platform	Elsevier	Taylor & Francis	Academy of Management (AOM)	SAGE Publishing	Wiley Publishing
Language polishing, grammar, minor editing	✓	✓	✓	✓	✓
Idea generation	✓	✓	✓	✓	✓
Summarizing or paraphrasing	✗	✓	✗	✓	✗
Initial drafting of manuscript	✗	✗	✗	✓	✗
Formatting assistance	✗	✗	✓	✗	✗
Generating core intellectual or scientific content	✗	✗	✗	✗	✗
Crediting AI as author or co-author	✗	✗	✗	✗	✗
Using AI to create/alter images *(unless clearly documented in approved methods)*	✗	✗	✗	✗	✗
Concealing AI usage	✗	✗	✗	✗	✗
Key: ✓ = Permitted | ✗ = Not Permitted					

Table 2: Summary of AI recommendations from Publishers

2.3 Student-Supervisor Responses

These policies by academic journals and universities raise questions about what constitutes authentic scholarly work and cautions PhD students about the limits of ethically leveraging technology without compromising critical thinking and research quality. Unfortunately, the growing use of these tools may lead novice researchers astray. Research highlights that students' appreciation for GenAI grows through critical and reflective use. However, supervisors' emphasis on potential risks can contribute to students' reluctance to openly engage with GenAI. This negatively influences student openness and appreciation for the productive integration of AI into doctoral research (Harding & Boyd, 2024). Because banning GenAI could hinder innovation, a more balanced approach is necessary: one that harnesses its efficiency and creative potential while ensuring alignment with academic rigor and ethical responsibilities (Henriksen, Mishra, Woo, & Oster, 2025). This growing reliance on AI-driven tools in research further underscores the need for practical guidance in their responsible use (Henriksen, Mishra, Woo, & Oster, 2025). Practical wisdom emphasizes intellectual and ethical virtues to enable sound judgment, reflective thinking, and ethical decision-making (Intezari, Pauleen, & Rooney, 2016). This can help ensure that AI integration enhances rather than undermines scholarly integrity.

In summary, universities must equip doctoral candidates with AI literacy skills to enable responsible engagement with these tools, helping them distinguish between appropriate applications and potential misuse (Ng, Leung, Chu, & Qiao, 2021). Beyond technical skills, programs should also foster broader intellectual, ethical, and professional competencies, such as practical reasoning and argumentation (Shulman, 2007). This aligns with the broader imperative to integrate intellectual, ethical, and collaborative principles into research training. A symbiotic intelligence framework, in which human reasoning and AI capabilities reinforce each other, can lead to more robust research outcomes (Samala et al., 2024).

3 Research Methodology

3.1 Theoretical Foundations

While the broader context of GenAI use in doctoral education is acknowledged, this study narrows its focus to supervision practices and institutional engagement, where conceptual guidance is particularly valuable. This focus is grounded in two key theoretical foundations. First, it draws on wisdom principles to reflect on doctoral research (Intezari, Pauleen, & Rooney, 2016) and its application of Generative AI (GenAI), using these insights to inform the development of a symbiotic intelligence framework (Desolda, Esposito, Lanzilotti, Piccinno, & Costabile, 2024). This framework promotes a collaborative relationship between human and artificial intelligence—emphasizing ethical alignment, reflective practice, and the augmentation of human judgment rather than its replacement. Second, the research employs the Theory of Sociotechnical Representation (TSR) to explore the perspectives and experiences of doctoral students and supervisors regarding GenAI. TSR integrates information systems (IS) views of technology, grounded in sociotechnical perspectives, with social psychology insights from social representations theory to understand how individuals perceive technology and how these perceptions influence its adoption and use (Weerasinghe, Pauleen, Scahill, & Taskin, 2022). TSR emphasizes that technological phenomena are co-created through social interactions; the representations formed (and co-formed) by individuals such as students and supervisors play a significant role in shaping how technologies like GenAI are adopted and integrated. This perspective offers a structured lens for examining how people construct, share, and act upon their understandings of emerging technologies (Weerasinghe, Pauleen, Taskin, & Scahill, 2023).

3.2 Research Design

This study adopts a qualitative design to investigate the integration of GenAI in PhD research and supervision in New Zealand Business Schools. It focuses on PhD students at various stages and their primary supervisors. The study consists of three phases. Phase 1, presented in this paper, involves a systematic literature review on GenAI in research, doctoral education, and relevant social science literature, as well as an exploration of GenAI tools, journal policies, and university guidelines. Phase 2 includes semi-structured interviews with supervisors (n=20), doctoral students (n=20), and research Deans (n=5) to examine their views on the benefits, challenges, and ethical concerns of using GenAI in doctoral work. Participants are not paired to avoid bias. Interviews (45–60 minutes) will cover usage patterns, institutional guidance, and perceived risks and opportunities.

Phase 3 is conducted in the form of action learning (Pauleen, Corbitt, & Yoong, 2007) with doctoral students. In this phase we will introduce three GenAI tools to the participants, who are doctoral students in their second or third year of study (n=10). They will be requested to use one of the tools and journal their experience for a period of 6 weeks (minimum of 10 interactions). We will then conduct follow-up interviews where they will be asked to reflect on the experience journaled. In Phase 3 the students who participate may not be those who are interviewed in Phase 2 to keep the data collection phases independent of each other. This study's qualitative approach supports in-depth insights to inform a symbiotic intelligence framework for responsible GenAI use in PhD research.

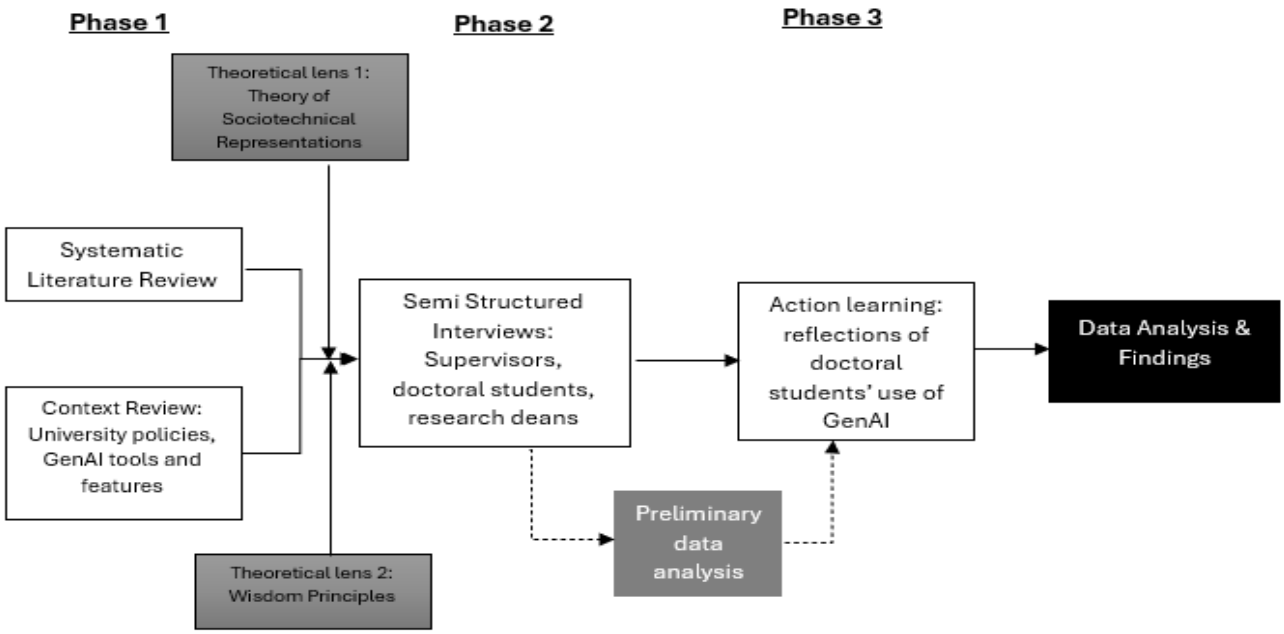

Fig 1: Research Design

4 Conclusion and Future Directions

This study explores the integration of GenAI in doctoral education, emphasizing the need for a balanced approach that enhances research productivity and creativity while maintaining ethical and pedagogical integrity. As a research-in-progress, this study proposes a symbiotic intelligence framework based on the TSR perspective, aiming to redefine PhD studies and supervision in the GenAI era by positioning GenAI as an enabler of human intellect. It seeks to uncover patterns of adoption and resistance, explore co-created norms, examine tensions in perspectives, and inform institutional policies and training for responsible GenAI integration. By applying TSR, we can move beyond just technical discussions about GenAI and explore how social interactions, shared meanings, and academic culture shape its adoption and use among doctoral students and supervisors. This perspective ensures that discussions about GenAI in academia are holistic, inclusive, and grounded in real experiences. Understanding these perspectives will offer insights into fostering reflective habits, ethical reasoning, and responsible AI engagement, contributing to a more thoughtful integration of GenAI in PhD research. The significance of this study lies in promoting a human-centred PhD experience that combines academic rigor with GenAI's learning potential. It reimagines the supervisory role as a collaborative process that balances AI-driven tasks with ethical reasoning and reflective human judgment. It also redefines success metrics by emphasizing balanced AI contributions, ethical considerations, reflective practices such as questioning AI limitations, comparing AI outputs with human judgment, and reflecting on AI's influence.

References

Bhullar, P. S., Joshi, M., & Chugh, R. (2024). ChatGPT in higher education - a synthesis of the literature and a future research agenda. Education and Information Technologies, 29,

21501–21522.
Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers and Education*, *227*(December 2024), 105224.
Desolda, G., Esposito, A., Lanzilotti, R., Piccinno, A., & Costabile, M. F. (2024). From human - centered to symbiotic artificial intelligence : a focus on medical applications. In *Multimedia Tools and Applications* (Issue 0123456789). Springer US. https://doi.org/10.1007/s11042-024-20414-5
Foroughi, B., Iranmanesh, M., Ghobakhloo, M., Senali, M. G., Annamalai, N., Naghmeh-Abbaspour, B., & Rejeb, A. (2024). Determinants of ChatGPT adoption among students in higher education: the moderating effect of trust. *Electronic Library*, *43*(1), 1–21.
Harding, D., & Boyd, P. (2024). Generative AI and Phd Supervision: a Covert Third Wheel or a Seat At the Table? *Association of Researchers in Construction Management, ARCOM 2024 - Proceedings of the 40th Annual Conference*, *September*, 91–100.
Henriksen, D., Mishra, P., Woo, L., & Oster, N. (2025). *The Education Doctorate in the Context of Generative Artificial Intelligence : Epistemic Shifts and Challenges to Practical Wisdom*. *10*(1), 73–79.
Intezari, A., Pauleen, D., & Rooney, D. (2016). Rediscovering Philosophia: The PhD as a Path to Enhancing Knowledge, Wisdom and Creating a Better World. *Journal of Business Ethics Education*, *13*, 147–168.
Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*(February).
Lin, Z., & Chen, H. (2024). Investigating the capability of ChatGPT for generating multiple-choice reading comprehension items. *System*, *123*(May), 103344.
Lo, C. K., Hew, K. F., & Jong, M. S. yung. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda. *Computers and Education*, *219*(June), 105100.
Mishra, P., Oster, N., & Henriksen, D. (2024). Generative AI, Teacher Knowledge and Educational Research: Bridging Short- and Long-Term Perspectives. *TechTrends*, *68*(2), 205–210.
Monash University Graduate Research. (2023). *Guidance Note on the use of Generative AI in research training & Examination*. https://www.monash.edu/graduate-research/support-and-resources/resources/guidance-on-generative-ai
Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*, 100041.
Pauleen, D. J., Corbitt, B., & Yoong, P. (2007). Discovering and articulating what is not yet known: Using action learning and grounded theory as a knowledge management strategy. *Learning Organization*, *14*(3), 222–240.
Samala, A. D., Rawas, S., Wang, T., Reed, J. M., Kim, J., Howard, N. J., & Ertz, M. (2024). Unveiling the landscape of generative artificial intelligence in education: a comprehensive taxonomy of applications, challenges, and future prospects. In *Education and Information Technologies* (Issue 0123456789). Springer US.
Shulman, L. (2007). Response to Comments: Practical Wisdom in the Service of Professional

Practice. *Educational Researcher*, *36*(9), 560–563.
Sun, D., Boudouaia, A., Zhu, C., & Li, Y. (2024). Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study. *International Journal of Educational Technology in Higher Education*, *21*(1).
The University of Auckland. (n.d.). *Advice for students on using Generative Artificial Intelligence in coursework*. Retrieved March 22, 2025, from https://www.auckland.ac.nz/en/students/forms-policies-and-guidelines/student-policies-and-guidelines/academic-integrity-copyright/advice-for-student-on-using-generative-ai.html
Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1).
Wang, N., Wang, X., & Su, Y. S. (2024). Critical analysis of the technological affordances, challenges and future directions of Generative AI in education: a systematic review. *Asia Pacific Journal of Education*, *44*(1), 139–155.
Weerasinghe, K., Pauleen, D., Scahill, S., & Taskin, N. (2022). Theory of Sociotechnical Representations: A Novel Approach to Understanding Technology Perspectives. *PACIS 2022 Proceedings*. https://aisel.aisnet.org/pacis2022/314
Weerasinghe, K., Pauleen, D., Taskin, N., & Scahill, S. (2023). Alignment of Big Data Perceptions Across Levels in Healthcare: The case of New Zealand. *Australasian Journal of Information Systems*, *27*.
Wise, B., Emerson, L., Van Luyn, A., Dyson, B., Bjork, C., & Thomas, S. E. (2024). A scholarly dialogue: writing scholarship, authorship, academic integrity and the challenges of AI. *Higher Education Research and Development*, *43*(3), 578–590.

19R. EXPLORING THE HUMAN-CENTRIC PERSPECTIVES OF SMART WAREHOUSES TECHNOLOGY ADOPTION – THE CASE OF NEW ZEALAND

Maryam Mirzaei
Auckland University of Technology
maryam.mirzaei@aut.ac.nz

Benjamin Dehe
Auckland University of Technology
Benjamin.Dehe@aut.ac.nz

Abstract

This ongoing study explores the adoption and implementation of smart warehouses in New Zealand's logistics sector through the lens of Socio-Technical Systems Theory (STST). By investigating the interrelated components of people, technologies, infrastructure, organisational structures, processes, cultural goals, and broader socio-human factors, this research leverages both secondary data— comprehensive database sourced from New Zealand government reports, industry publications, corporate case studies, and reputable news outlets—and primary data from industry interviews with warehouse managers. The research so far identifies two dominant automation models: the 'Split Task Model', where technology handles repetitive tasks while humans oversee complex functions, and the 'Collaborative Model', where technology enhances human performance making tasks more efficient, accurate, or safe. Findings indicate that automation in New Zealand is reshaping rather than replacing jobs, prompting role evolution and a growing focus on up-skilling.

Keywords

Automation, Smart warehouse, Socio-Technical Systems Theory, Human Centricity

1. Introduction

Warehousing is a cornerstone of modern supply chains, serving as a critical hub for the storage, management, and distribution of goods. Warehouse storage and retrieval have traditionally relied on manual processes and human labour. However, as supply chains become more complex and demand for efficiency grows, the concept of the "smart warehouse" has emerged as a transformative development in the industry. A smart warehouse leverages Industry 4.0 technologies, to enhance operational efficiency, accuracy, and flexibility. These technologies enable real-time inventory tracking, predictive maintenance, autonomous material handling, and seamless integration with broader supply chain systems. By reducing reliance on manual labour and optimising workflows, smart warehouses offer significant cost savings, faster order fulfilment, and improved overall performance (Geest et al., 2021). While smart warehousing technology is well-documented, the human aspect is less explored (Grosse et al., 2023). The integration of advanced technologies alters traditional job roles, requiring adaptation and raising concerns around job displacement, resistance to change, and the need for workforce upskilling. Moreover, managerial perspectives, organisational readiness, and broader socio-technical considerations shape how smart warehouses are implemented and accepted. Accordingly, this research investigates the following question: How can socio-technical systems theory (STST) be applied to guide the adoption and implementation of human-centric smart warehouses within New Zealand's logistics sector?

New Zealand has historically been cautious in adopting smart warehouses, often lagging behind other developed economies in embracing automation (Clark, 2024). However, driven by acute labour shortages, rising e-commerce demands, and the quest for enhanced productivity, New Zealand's warehousing and logistics sector is experiencing a gradual but notable shift towards "smart" operations. The warehousing sector, now a NZ$1.8 billion industry, is at a turning point

(IBISWorld, 2024). This study positions New Zealand as a compelling case to examine the human side of smart warehouse adoption. As industries worldwide grapple with automation's impact on the workforce, New Zealand's transition offers valuable insights into the challenges, perceptions, and opportunities shaping the future of warehousing. While the study is bounded by its local context, the technologies, practices, and organisational dynamics examined are common across warehouse settings globally, making the findings broadly transferable. By exploring employee experiences, managerial decision-making, and organisational readiness, this research identifies key challenges and enablers of smart warehouse implementation. The findings will guide practitioners and policymakers in ensuring a seamless transition to automation while prioritising workforce development and organisational change.

2.2. State-of-the-field

As smart warehouses gain popularity, there is no single agreed-upon definition. Zhen and Li (2022) describe them as "mostly unmanned, paperless, and automated" highlighting their contrast with traditional, labour-intensive warehouses with fragmented systems. There are basic key features highlighted in the literature, which include i) connectivity: offering continuous communication between networked devices such as software, sensors, and equipment; ii) automation: reducing the use of people with autonomous material handling devices; and iii) real-time analytics: where process data are retrieved in real-time and use to simplify decision-making. iv) Adaptability and intelligence are also mentioned in literature as key features, enabling dynamic adjustments with minimal delay and human input (Min, 2023; Geest et al., 2022). Therefore, despite various definitions, scholars and practitioners agree in essence that a smart warehouse largely utilises digital technology in improving the management and monitoring of the core activities and optimising their functioning.

Several enabling technologies support the transformation of traditional warehouses into smart warehouses. Central to this shift is the Internet of Things (IoT) and cyber-physical systems (CPS), which integrate physical assets with digital systems for real-time tracking and adaptive management (Lee et al., 2017; Liu et al., 2018; Geest et al., 2022). Automated identification technologies such as RFID, voice recognition, and AS/RS improve product identification accuracy and streamline inventory management (Edouard et al., 2022; Geest et al., 2022). Robotics, including AMRs, cobots, and AGVs, automate manual tasks, enhancing efficiency, safety, and flexibility (Barker, 2022). When combined with augmented reality (AR), these systems further reduce errors and improve spatial interaction (Stoltz et al., 2017). Wearable technologies like smart glasses support real-time guidance and reduce cognitive load, though their adoption must address user health and acceptance concerns (Epe et al., 2024). While AI is embedded in the above enablers such as on-device inference in IoT sensors, speech recognition in voice-directed picking, or vision guidance in cobots and smart-glasses, it can also operate as an independent analytics layer that drives demand forecasting, dynamic slotting, and real-time exception handling in warehouse-management systems; together with data analytics and digital-twin simulations, these AI capabilities deliver predictive insights and decision support, enabling smarter and more responsive warehouse operations (Lee et al., 2017). These technologies together are the prime facilitators of the evolution of conventional warehouses into smart warehouses defined earlier.

Human-centricity refers to the idea that prioritises human needs, values, and experiences in the design, implementation, and operation of systems, technologies, and organisational processes.

While Industry 4.0 has driven automation, the concept of Industry 5.0 highlights the importance of a human-centred approach to digitalisation, ensuring technology serves societal and workforce needs rather than displacing them (Ivanov, 2022). This includes designing systems that support workers' well-being, enable skill development, and promote seamless human-technology collaboration (Grosse et al., 2023). Ethical AI and Human Factors Engineering, which emerge from concerns over AI and automation, focus on designing systems that enhance human decision-making, safety, and work satisfaction (Grote, 2023). Human-centric smart warehouses deploy automation that enhances workers' capabilities, embedding ergonomic design that prioritises well-being and empowerment (Hermawati et al., 2024), and succeed only when an organisation's skills, culture, and infrastructure are prepared for intelligent logistics (Mahroof, 2019). In smart warehousing, human-centricity is critical to addressing concerns about job displacement, upskilling, and the social implications of automation, making it an essential consideration for sustainable and ethical technological adoption.

3. Overview of the Methodology (an ongoing study)

3.1. Theoretical Framework – Socio-Technical Systems Theory (STST)

While existing research often focuses on technological advancements, our study adopts a human-centric approach. We examine workforce transformation and organizational adaptation through the lens of Socio-Technical Systems Theory (STST) which emphasizes the critical alignment of technical and social components to ensure successful integration (Cherns, 1976). STST offers a holistic lens for exploring the complexities of smart warehouse adoption under the Industry 4.0 and Industry 5.0 umbrella by treating social and technical factors as interconnected rather than independent. STST has been applied in various studies to explore this interplay. Marcon et al. (2024) and Vlachos et al. (2024) emphasize that aligning human and technological elements—through collaboration and feedback loops—enhances outcomes in smart warehouse environments. Sony and Naik (2020) argue that Industry 4.0 must extend beyond technical solutions to address human factors such as skill development, collaboration, and adaptability. They stress the need for adaptable technologies that foster meaningful human-technology interaction. Sarker et al. (2019) review multiple socio-technical models—ranging from user-focused to organisational and societal levels—and recommend selecting frameworks that best align with the research context. Hence, drawing on Sony and Naik's (2020) STST model, this study adopts six interrelated constructs—people, infrastructure, processes, technologies, socio-human aspects, and organisational culture and goals—to guide analysis.

3.2. Research Design

To explore smart warehouse adoption and implementation in New Zealand and address the posit research question we adopt a qualitative design. Consistent with established qualitative methods (Creswell, 2014), this study employs a rigorous qualitative approach, integrating secondary research with ongoing primary data collection to ensure robust findings. Secondary data based on comprehensive database sourced from New Zealand government reports, industry publications, corporate case studies, and reputable news outlets, forms the foundation of this paper. These sources were identified and filtered through an online search using strict inclusion criteria: relevance to smart warehousing in New Zealand, publication date post-2020, and credibility of the issuing institution or publisher. The data was analysed using thematic coding, informed by six STST constructs (Sony and Naik, 2020), to identify patterns across social and technical domains. Table 1 provides some examples of this analysis process.

Primary data are being collected via semi-structured interviews with warehouse managers. The interim results reported here draw mainly on a large corpus of secondary sources. As additional interviews are completed, we will iteratively code them, compare the themes with the secondary evidence, and reconcile interpretations across the research team until saturation is reached. This phased, triangulated design (Denzin, 1978) maintains methodological rigour while letting the study adapt as new insights emerge.

STST Construct	NZ Examples, Quotes, and Evidence
Goals Strategic objectives, performance metrics, alignment across departments and industry partners, focusing on safety, productivity, efficiency, competitiveness, and workforce development.	Established in 2018, the Future of Work Tripartite Forum—a partnership between government, Business New Zealand, and the New Zealand Council of Trade Unions—addresses labor market changes and outlines strategic priorities to strengthen economic resilience and workforce preparedness (Ministry of Business, Innovation and Employment, n.d.) NZ Post partnered with DataRobot to implement AI-driven predictive analytics and image classification, substantially improving warehouse safety (DataRobot, 2024). Similarly, The Warehouse Group collaborated with MicroStrategy to integrate advanced analytics and AI chatbot insights, optimizing its warehouse (MicroStrategy, 2024).
Technology Tools and systems for automation, analytics, control, and integration.	Fonterra employs a Robotic Process Automation (RPA) bot named Enid, designed to tackle repetitive administrative tasks like resolving quantity mismatches in the SAP system (Fonterra, n.d.a). New Zealand companies are leveraging wearable technology, such as exoskeletons and wearable robotic devices (Public, 2023).
People Human skills, competencies, evolving roles, workforce adaptation, and retraining initiatives triggered by automation.	Automation is unlikely to trigger significant job losses in New Zealand but will transform nature of work, necessitating enhanced education, training and support. (NZ Productivity Commission, 2020). Mainfreight's use of autonomous mobile robots (AMRs) highlights the evolving role of workers toward technical oversight. (Mainfreight, n.d.). AS Colour has introduced a Multishuttle GTP system, boosting worker productivity through implied reskilling initiatives (Dematic, 2024)
Infrastructure Physical and digital systems enabling automation, including connectivity, IoT sensors, robotics, and integrated IT networks.	In 2022, Fonterra upgraded to a new generation of AGVs supplied by Dematic, featuring advanced navigation lasers and hazard detection systems for safe 24/7 operations (Dematic, 2022). Woolworths has introduced micro-fulfilment centers equipped with automated storage units capable of holding up to 11,000 grocery items, concentrate on keeping the robotic systems stocked with inventory (Woolworths Group, 2024).

Processes & Procedures Operational workflows, task redistribution, coordination mechanisms, and organisational structure adaptations	Foodstuffs North Island's RPA streamlines financial processes—processing ~3M invoices and 6,000 validations annually—saving 9,000 staff hours across 200 stores and shifting employee focus from routine data entry to analysis and complex issue resolution. (NZ Herald, 2023). Cardinal Logistics' fully automated 3PL warehouse demonstrates how technological advancements can restructure workflows, moving human roles from manual handling to oversight and system management (Innovatek, 2023).
Culture Organisational mindset, acceptance of automation, shared values regarding safety, innovation, and proactive readiness for change.	*"New Zealand may be on the edge of a wave of warehouse automation* [mindset].... *Where once it was feared robots would take jobs from humans, now they are filling chronic worker shortages*[acceptance]... *Automated solutions ... promise to deliver improvements in efficiency, safety and the ability to do more with a smaller footprint.* [shared value]*" (FTD Magazine, 2022).* "*Innovation fuels our transformation – futuristic tech and agile processes are reshaping our warehouse operations to be both efficient and competitive.*" (Grant, 2024)

Table 1: Examples from comprehensive data analyses for secondary data

.

4. Initial Results and Discussion

4.1 Hybrid Warehousing Models

Although "fully automated" may evoke images of so-called "dark warehouses" running without human involvement or lighting, powered solely by robotics, AI, and other automation technologies, our research shows that New Zealand's warehousing landscape is predominantly hybrid, characterized by two distinct operational models: Split Task and Collaborative. 1. Split Task model where technology handle repetitive, high-speed tasks, while human workers manage more complex or supervisory roles. For example, Fonterra's implementation of AGVs at their Edendale facility, where AGVs handle tasks, such as transporting raw materials and finished products, allowing human workers to focus on selecting and inspecting dairy products (Dematic, 2022). Foodstuffs, the country's largest grocery retailer, has deployed robotics to automate mundane tasks across its business, freeing up approximately 9,000 hours of staff time for other activities (NZ Herald, 2023). 2. Collaborative Model – Where technology enhances what humans do by making tasks more efficient, accurate, or safe. For example, wearable technologies, such as exoskeletons, help mitigate musculoskeletal injuries and promote ergonomic best practices in retail and logistics environments (Public, 2023). NZ Post's use of AI to detect potential safety incidents, enables human supervisors to respond promptly and maintain a safer work environment (Chapman, 2024).

4.2 New Zealand Workforce Evolution and Empowerment

Our key observation is that automation is not replacing jobs but reshaping them, necessitating workforce reskilling and role evolution. In New Zealand, policymakers and industry experts—such as those at the Future of Work Tripartite Forum—believe that while robotics and AI might disrupt certain sectors, they will also pave the way for new roles and opportunities. This dual impact not only helps balance out potential job losses but is also seen as a key solution to easing ongoing labour shortages (Allan & Sanderson, 2021). For example, adoption of AGVs at Fonterra

notably decreased the reliance on traditional forklift drivers, replacing manual transport with new roles like warehouse automation technicians, who manage AGV operations, troubleshoot issues, and oversee scheduling software (Dematic, 2022). The discussion on “Upskilling, and Retraining” highlights industry players' proactive efforts in supporting workers to adapt and acquire new skills, enabling them to thrive in the evolving job landscape. For example, Fonterra, as part of the Aotearoa New Zealand Skills Pledge, aims to double on-the-job training by 2025, likely incorporating skills for operating automated guided vehicles (AGVs) in its warehouses (Fonterra, n.d.b). Mainfreight enhances employee capabilities through training programs focused on new automation solutions, ensuring workers thrive alongside technology rather than being displaced (Mainfreight, n.d.). Foodstuff North Island supports workforce adaptability by reskilling employees to use advanced warehouse management systems, reflecting a shift toward digital proficiency (Foodstuffs, 2022). These examples demonstrate a commitment to evolving job roles alongside automation, a strategy Deloitte (2024) emphasises as crucial for sustaining employment and enhancing skills in New Zealand's changing labour market.

5. Conclusion

Our initial findings suggest that the adoption and implementation of smart warehouses in New Zealand’s logistics sector can be effectively understood through the lens of the Socio-Technical Systems Theory (STST) and that the landscape of smart warehousing in NZ spans multiple industries and features key players who are pioneering new ways of working. Our research illustrates that warehouses are embracing two key models of automation. The Split Task Model delegates repetitive, high-speed tasks to technology, while humans focus on complex or supervisory roles. The Collaborative Model enhances human work by improving efficiency, accuracy, and safety through technology. Findings also show in New Zealand, automation is not replacing jobs but reshaping them; driving the need for reskilling and evolving roles, while also helping to address labour shortages. Key industry players are responding by actively investing in on-the-job upskilling and retraining, ensuring the workforce remains adaptive, capable, and future-ready. Future interviews with managers are expected to deepen our understanding of these shifts, revealing how businesses navigate and integrate these socio-technical changes in practice. These conversations will also help identify common barriers to adoption and clarify how firms overcome them. By capturing these insights, we aim to build a clearer picture of both the opportunities and the challenges involved in embracing socio-technical transformation in Smart Warehouses.

Referencing

Allan, C., & Sanderson, L. (2021). *Labour market impacts of technology change: Evidence from linked employer–employee data* (CEU Working Paper No. 21/02). New Zealand Ministry of Business, Innovation & Employment. https://apo.org.au/sites/default/files/resource-files/2021-05/apo-nid313621.pdf

Barker, J. (2022). *Smart warehouses—a sociotechnical perspective* (pp. 47–60). Elsevier eBooks. https://doi.org/10.1016/b978-0-323-91614-1.00003-4

Chapman, T. (2024, October 2). How NZ Post is improving CX and warehouse safety using AI. Supply Chain Digital. https://supplychaindigital.com/digital-supply-chain/nz-post-improving-cx-warehouse-safety-with-a

Cherns, A. B. (1976). Principles of sociotechnical design. *Human Relations, 29*(8), 783–792. https://doi.org/10.1177/001872677602900806

Clark, P. (2024/02/15). Manufacturers urged to adopt more digital tech—'NZ is a decade behind Europe'. *Radio New Zealand.* https://www.rnz.co.nz/news/business/509229/manufacturers-urged-to-adopt-more-digital-tech-nz-is-a-decade-behind-europe

DataRobot. (2024). How a 180-year-old postal service reduces customer calls by 33% and improves warehouse safety with predictive and generative AI. https://www.datarobot.com/customers/nz-post/

Deloitte. (2024). *Aotearoa's changing workforce*. https://www.deloitte.com/nz/en/services/consulting/blogs/aotearoas-changing-workforce.html

Dematic. (2022, May 19). *Fonterra upgrades manufacturing facility with new automated guided vehicles from Dematic*. https://www.dematic.com/en-gb/newsroom/press-releases/2022/fonterra-upgrades-manufacturing-facility-with-new-automated-guided-vehicles-from-dematic/

Dematic. (2024). *AS Colour meets current and future growth with new automated Auckland DC*. https://www.dematic.com/en-us/insights/case-studies/as-colour-new-zealand/

Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods.* New York: McGraw-Hill.

Edouard, A., Sallez, Y., Fortineau, V., Lamouri, S., & Berger, A. (2022). Automated storage and retrieval systems: an attractive solution for an Urban warehouse's sustainable development. *Sustainability*, *14*(15), 9518.

Epe, M., Azmat, M., Islam, D. M. Z., & Khalid, R. (2024). Use of smart glasses for boosting warehouse efficiency: implications for change management. *Logistics*, *8*(4), 106. https://doi.org/10.3390/logistics8040106

Fonterra. (n.d.a). *Meet Fonterra's newest team member – Enid*. https://www.fonterra.com/nz/en/our-stories/articles/meet-fonterras-newest-team-member-enid.html

Fonterra. (n.d.b). *Why we've signed up to the Aotearoa New Zealand Skills Pledge*. https://www.fonterra.com/nz/en/our-stories/articles/why-weve-signed-up-to-the-aotearoa-new-zealand-skills-pledge.html

Foodstuffs. (2022, July 20). *Foodstuffs North Island accelerates digital transformation through workforce reskilling*. Retrieved from https://www.foodstuffs.co.nz/news/digital-transformation-reskilling

FTD Magazine. (February 22). *Is greater warehouse automation coming to New Zealand? Worker shortages and technological enhancements strengthen case for automation*. https://ftdmag.co.nz/2022/02/22/is-greater-warehouse-automation-coming-to-new-zealand-worker-shortages-and-technological-enhancements-strengthen-case-for-automation/

Grant, H. (2024, October 18). The future of warehousing: Predictions for 2025 and beyond. Scoop. https://www.scoop.co.nz/stories/BU2410/S00339/the-future-of-warehousing-predictions-for-2025-and-beyond.htm

Grosse, E. H., Sgarbossa, F., Berlin, C., & Neumann, W. P. (2023). Human-centric production and logistics system design and management: transitioning from Industry 4.0 to Industry 5.0. *International Journal of Production Research*, *61*(22), 7749-7759.

Grote, G. (2023). Shaping the development and use of Artificial Intelligence: how human factors and ergonomics expertise can become more pertinent. *Ergonomics*, *66*(11), 1702–1710. https://doi.org/10.1080/00140139.2023.2278408

Hermawati, S., Correa, R., Mohan, M., Lawson, G., & Houghton, R. (2024). Defining human-centricity in Industry 5.0 and assessing the readiness of ergonomics/human factors communities in UK. Ergonomics, 68(6), 857–876. https://doi.org/10.1080/00140139.2024.2343947

IBISWorld. (2024). *Warehousing and storage services in New Zealand*. IBISWorld Industry Report I5300NZ. https://www.ibisworld.com/new-zealand/industry/warehousing-and-storage-services/727

Innovatek. (2023, July 4). *NZ$450m fully automated warehouse investment*. Retrieved from https://innovatek.co.nz/nz450m-fully-automated-warehouse-investment-2/

Lee, C. K. M., Lv, Y., Ng, K. K. H., Ho, W., & Choy, K. L. (2017). Design and application of Internet of things-based warehouse management system for smart logistics. *International Journal of Production Research*, *56*(8), 2753–2768. https://doi.org/10.1080/00207543.2017.1394592

Ivanov, D. (2022). The Industry 5.0 framework: viability-based integration of the resilience, sustainability, and human-centricity perspectives. *International Journal of Production Research*, *61*(5), 1683–1695. https://doi.org/10.1080/00207543.2022.2118892

Mainfreight. (n.d.). *Warehousing*. https://www.mainfreight.com/new-zealand/en-nz/services-nz/warehousing

Mahroof, K. (2019). A human-centric perspective exploring the readiness towards smart warehousing: The case of a large retail distribution warehouse. International Journal of Information Management, 45, 176-190. https://doi.org/10.1016/j.ijinfomgt.2018.11.008.

Marcon, É., Marodin, G. A., & Frank, A. G. (2024). Combining organizational and social factors to support Industry 4.0 implementation: A sociotechnical and configurational analysis of technology adopters. *IEEE Transactions on Engineering Management, 72*(1), 146–160. https://doi.org/10.1109/TEM.2024.3506991

MicroStrategy. (2024). Catalyzing retail growth & agility with data insights & AI. https://www.strategysoftware.com/customer-stories/the-warehouse-group-catalyzing-growth-and-agility-in-retail-with-ai-analytics

Min, H. (2023). Smart warehousing as a wave of the future. *Logistics,* 7(2), 30. https://doi.org/10.3390/logistics7020030

Ministry of Business, Innovation & Employment (n.d.). Future of Work Tripartite Forum. Retrieved from https://www.mbie.govt.nz/business-and-employment/employment-and-skills/future-of-work-tripartite-forum

New Zealand Productivity Commission. (2020). *Technological change and the future of work: Final report*. https://www.treasury.govt.nz/sites/default/files/2024-05/pc-inq-tcfw-final-report-technological-change-and-the-future-of-work.pdf

NZ Herald. (2023, February 22). *How Foodstuffs is using robotics to free up 9000 hours of staff time*. https://www.nzherald.co.nz/business/how-foodstuffs-is-using-robotics-to-free-up-9000-hours-of-staff-time/E63H2ICKOFCWLAQSWJZYZ27ULE/

Public, J. (2023, March 16). *Kiwi business deploying exoskeletons and wearable tech to reduce injuries, future-proof workforces*. Infonews. https://www.infonews.co.nz/news.cfm?id=124350

Sarker, S., Chatterjee, S., Xiao, X., & Elbanna, A. (2019). The Sociotechnical Axis of Cohesion for the Is Discipline: Its Historical Legacy and its Continued Relevance. *MIS Quarterly*, *43*(3), 695–720, A1–A5. https://www.jstor.org/stable/26848052.

Sony, M., & Naik, S. (2020). Industry 4.0 integration with socio-technical systems theory: A systematic review and proposed theoretical model. *Technology in Society*, *61*. https://doi.org/10.1016/j.techsoc.2020.101248

Stoltz, M. H., Giannikas, V., McFarlane, D., Strachan, J., Um, J., & Srinivasan, R. (2017). Augmented reality in warehouse operations: opportunities and barriers. *IFAC-PapersOnLine*, *50*(1), 12979–12984. https://doi.org/10.1016/j.ifacol.2017.08.1807

Supermarket News. (2023). *New wearable tech to reduce workplace injury*. Retrieved from https://supermarketnews.co.nz/news/new-wearable-tech-to-reduce-workplace-injury/

Supermarket News. (2023, March 15). *Countdown launches micro-fulfilment tech*. https://supermarketnews.co.nz/news/countdown-launches-micro-fulfilment-tech/

Technology Magazine. (2020). *Foodstuffs North Island is embracing change with a customer driven digital transformation*. https://technologymagazine.com/company-reports/foodstuffs-north-island-embracing-change-customer-driven-digital-transformation

Van Geest, M., Tekinerdogan, B., & Catal, C. (2021). Design of a reference architecture for developing smart warehouses in industry 4.0. *Computers in industry*, *124*, 103343. https://doi.org/10.1016/j.compind.2020.103343

Van Geest, M., Tekinerdogan, B., & Catal, C. (2022). Smart warehouses: rationale, challenges and solution directions. *Applied sciences*, *12*(1), 219. https://doi.org/10.3390/app12010219

Vlachos, I., Pascazzi, R. M., Ntotis, M., Spanaki, K., Despoudi, S., & Repoussis, P. (2022). Smart and flexible manufacturing systems using autonomous guided vehicles (AGVs) and the internet of things (IoT). *International Journal of Production Research, 62*(15), 5574–5595. https://doi.org/10.1080/00207543.2022.2136282

Woolworths Group. (2024, Month Day). *Woolworths introduces micro-fulfilment centres with automated storage units for efficient inventory management*. https://www.woolworths.co.nz/info/news-and-media-releases/2021/countdown-launches-micro-fulfilment-tech-to-make-online-shopping-five-times-faster

Zhen, L., & Li, H. (2022). A literature review of smart warehouse operations management. *Frontiers of Engineering Management*, *9*(1), 31-55. https://doi.org/10.1007/s42524-021-0178-9

20R. WECHAT AS A DIGITAL HEALTH TOOL:

UNDERSTANDING THE ROLE OF DIGITAL HEALTH LITERACY IN ITS EFFECTIVE USE AMONG ELDERLY DIABETIC PATIENT IN CHINA

Ruilin Zhu
Lancaster University
ruilin.zhu@lancaster.ac.uk

Zhiqin Wu
The Second Affiliated Hospital of Chongqing Medical University
302955@cqmu.edu.cn

Abstract

Despite the plethora of literature on digital health literacy on effective use of digital health tool, research remains limited on the use of non-medical apps for healthcare purposes. Given the increasing prevalence of elderly diabetic patients in China, the study aims to uncover the impact of digital health literacy on such app as well as the effective use among users. Drawing on Health Empowerment Model, the research employs the quantitative approach with the data collected from the top-tier hospital in China over six months to examine the focal issue. The results will be expected to provide both academic and practical implications.

Keywords

Digital health literacy, WeChat, Elderly patient, Diabetes, China

1 Introduction

The digitalization of healthcare presents new possibilities for patient engagement, self-monitoring, and healthcare communication, particularly vital for elderly patients in need of long-term disease management (Andrade et al., 2021). It is thus reshaping the delivery of medical services and has become an essential component of chronic disease management (CDM) (Pong et al., 2024). To fully tap the advantage of digital healthcare, patients should possess sufficient digital health literacy (DHL), which is the ability to search, appraise, understand, and use health information within digital contexts and apply such knowledge to address health-related issues (Van Der Vaart & Drossaert, 2017; Yang et al., 2022). Extensive studies have examined the impacts of DHL on patients, primarily focusing on conventional health digital platforms, such as hospital or clinic webpages (Dadaczynski et al., 2021), mobile health app (Kim & Xie, 2015), and wearable devices (Nittayathammakul et al., 2021).

While these studies provide valuable insights to the DHL literature, there lacks research aiming at understanding WeChat, which has been widely used in the healthcare setting for CDM in China. Originally developed as an instant messaging app, WeChat has transformed into an all-in-one digital platform with features that are expanded far beyond its previous functionalities. In China specifically, WeChat is highly prevalent and popular, functioning as the "de facto platform" for communication, business activities, and health services (Zou et al, 2023). WeChat has been retrofitted in many hospitals as a formal channel for patient-doctor communication, appointments, and health education (Chen et al., 2024). In comparison to digital health applications specifically designed for the medical-context, WeChat provides a general function, prompting questions of whether and how elderly patients engage with health-related content accessed through a non-medical app. Despite the efforts from prior research that has focused on DHL in the context of

digital health apps for health-specific characteristics, it remains inconclusive or unknown that the DHL's association with patient engagement in a multipurpose, non-medical digital platform like WeChat. This research addresses an important gap through the investigation of the impact of DHLs on the effective use of WeChat for CDM amongst elderly diabetic patients in China.

Drawing on the Health Empowerment Model (HEM), this research aims at uncovering the important role of DHL as an enabler of health empowerment, focusing on the elderly patients' participation in their health management. HEM posits that empowerment occurs when people acquire knowledge, confidence, and self-efficacy, which allows them to make informed decisions about their health and participate in healthcare processes (Schulz & Nakamoto, 2013). This study moves beyond the traditional technology adoption models by incorporating the cognitive and behavioral mechanisms of engagement to explore the influences of DHL within the HEM framework. In doing so, this study employs a quantitative survey approach with data collected from elderly patients diagnosed with type 2 diabetes at a leading tertiary hospital in China. We have chosen diabetes as the focal chronic disease for this research is due to the rapidly increasing prevalence of diabetes in China. China now has the largest number of people (118 million) living with diabetes, accounting for approximately one-quarter of the global diabetes population, thus posing mounting medical, economic, societal pressures on the local healthcare of diabetic patients, especially the seniors (Xu et al., 2024). By leveraging validated DHL measurement scales and advanced statistical analyses, this study endeavors to provide theoretical and empirical insights into the relationship between DHL and WeChat-based CDM.

This research will make several key contributions. Academically, it extends DHL research beyond conventional health-specific platforms, providing new theoretical insights into how non-medical digital ecosystems shape health behaviors. Practically, it offers evidence-based recommendations for hospitals and policymakers on optimizing WeChat's role in CDM, ensuring accessible and effective digital health services for aging populations in China.

The remainder of the paper proceeds as follows. The next section presents a literature review that synthesizes research on digital health literacy, chronic disease management, and WeChat-based healthcare interventions. This is followed by the methodology section, detailing the quantitative approach, data collection methods, and analytical procedures. The expected outcomes section discusses the anticipated findings, including theoretical and practical implications.

2 Literature Review

2.1 Digital Health Literacy

Digital health literacy (DHL) is defined as an individual's ability to seek, understand, appraise, and apply digital health information to manage health-related decisions (Norman & Skinner, 2006). Over the years, DHL has evolved in response to the increasing integration of digital platforms into healthcare, shifting from basic information retrieval skills to more complex competencies such as critical appraisal of online health content and interactive engagement with digital tools (Van Der Vaart & Drossaert, 2017). Extant research underscores that DHL is a critical resource for assisting individuals navigate digital health landscapes, promoting informed health-related decision-making, and ultimately improving health outcomes (Jacobs et al., 2016). However, a large number of these studies were conducted with a more general or technologically savvy population, while only a few examined the application of DHL amongst the seniors,

especially those with chronic diseases, such as diabetes. Given the aging population and the burden of rising diabetes patients on healthcare, understanding DHL on elderly diabetic patients is necessary.

2.2 Digital Health Literacy and Chronic Disease Management

Chronic disease management (CDM) relates to the ongoing and long-term engagement with healthcare resources to manage chronic conditions, such as diabetes (Reynolds et al., 2018). The effectiveness of CDM is highly dependent upon consistent health monitoring, adherence to a treatment plan, and proactive participation from the patient, all of which, increasingly, can be driven by digital health tools (Matthew-Maich et al., 2016). DHL is an enabler for patients engaging in self-management for chronic conditions by providing access to credible information, telemedicine, or mobile apps (Yang et al., 2025). Research has indicated that higher levels of DHL positively correlate with chronic disease management outcomes (Lee & Nam, 2024).

Despite these endeavours, research on the interaction between DHL and CDM remains fragmented. Most studies were undertaken in the contexts where digital health interventions are available, acceptable, and affordable, while insights into different settings, such as the seniors with varying digital access, remain scarce. In addition, prior research often references younger or middle-aged adults and fails to examine aging populations who face cognitive decline, limited prior exposure to digital technologies, and different levels of social support. As a result, there is a lack of knowledge in the field about DHL and CDM among elderly patients.

2.3 WeChat as a Platform for Health Management

While Chinese seniors are currently not tapping into wearable health technology, mainly due to cost, accuracy, and data privacy concern (Li et al. 2020), WeChat serves as a primary platform of digital communication and services in China with more than 1.3 billion monthly active users (Sina, 2024). In other words, WeChat is significantly more utilized than wearable health technology and thus offers a neglected and relevant opportunity for digital health care management for senior patients with chronic conditions. Specifically, WeChat is widely used by seniors in China for everyday communication activities such as sending messages, making payments and scanning QR codes. The high acceptance of WeChat makes it a promising option for digital health interventions, especially targeting older adults living with chronic diseases.

Previous research has examined WeChat's role in health management in relation to patient education, communication with health professionals, and peer support groups (Wang et al., 2021), but they typically do not study the use of WeChat in the medical contexts under the influence of DHL, especially for older adults. Apart from the practical gap, several theoretical, conceptual, and methodological gaps persist. From the theoretical perspective, while DHL is acknowledged as a crucial factor in health technology utilization, its specific role in facilitating chronic disease self-management via mainstream digital platforms like WeChat remains under-theorized. Conceptually, existing frameworks do not sufficiently account for how seniors interact with general-purpose digital platforms for health-related activities, as most models focus on standalone health apps. Regarding methodology, many studies employ cross-sectional surveys or short-term interventions, limiting insights into long-term behavioral patterns and barriers to sustained engagement amongst seniors.

Given the increasing prevalence of diabetes among China's aging population and the established popularity of WeChat, this study is overdue. By examining how DHL influences the effective use of WeChat for diabetes management among the elderly patients, this research contributes to a more nuanced understanding of digital health engagement in aging populations and provides practical insights for designing more inclusive digital health interventions.

2.4 The Health Empowerment Model (HEM)

This study utilizes the Health Empowerment Model (HEM) to investigate the role of DHL in enhancing the utilization of WeChat for diabetes management among elderly patients in China. HEM suggests that health empowerment involves individuals acquiring information, skills, and confidence to actively manage their health. Unlike conventional models where patients passively receive medical care, HEM highlights an interactive process where individuals utilize health information to make informed decisions, interact with healthcare providers, and self-manage chronic conditions (Schulz & Nakamoto, 2013). By integrating HEM into this study, we depart from traditional technology adoption theories (e.g., TAM or UTAUT) to explore how DHL influences active and passive engagement in digital health through cognitive, behavioral, and social mechanisms. Thus, HEM offers a suitable framework for examining the central issue at hand.

3 Research Method

3.1 Research Design

This study employs a quantitative survey approach, which is well-suited for understanding behavioral patterns in large populations (Ghanad, 2023) and has been widely used in digital health research. Given the increasing importance of digital tools in healthcare, a structured and validated survey has been developed, which allows for rigorous data collection and statistical analysis, ensuring robust findings (Ju et al., 2006).

3.2 Survey Instrument

The instrument for the survey has been constructed by collectively integrating two widely-accepted DHL scales with adjustments to fit WeChat-based health management. We have integrated items from the eHealth Literacy Scale (eHEALS) (Norman & Skinner, 2006) and the Digital Health Literacy Instrument (DHLI) (Van Der Vaart & Drossaert, 2017). The use of these scales has demonstrated reliable measures of DHL in previous studies (Liu et al., 2020). In addition to these scales, the instrument also considered scenario-based items to capture specific interactions with WeChat health features. These include accessing hospital WeChat service accounts, following public health channels, engaging with diabetes-related group chats, and using mini-programs for appointment booking or health tracking. This operationalisation allows the study to contextualise DHL within WeChat's ecosystem and better capture how seniors enact digital health behaviours in real-world use. This adaptation process follows a systematic content validity approach, in which a panel of healthcare professionals, digital health experts, and senior WeChat users provided feedback for relevance and clarity of items. The final survey used for the study was pretested with 30 participants to establish the instrument's validity, reliability, and relevance to the cultural and professional context.

3.3 Data Collection

Data will have been collected over six months from senior patients diagnosed with Type 2 diabetes (more than one year) at a leading tertiary hospital in China with the targeted sample size of 300 participants. This hospital ranks among the top-tier medical institutions according to the National Health Commission of China, ensuring access to a diverse and representative patient population. A stratified random sampling strategy will be employed to ensure adequate representation of different age groups (60-69, 70-79, 80+) and education levels while excluding those with cognitive impairment and/or severe visual or motor impairments that prevent interaction with digital tools.

The recruitment process has been conducted in collaboration with the hospital outpatient departments and geriatric specialists. Participants will be approached by trained research assistants in follow-up visits to clarify the study purpose and gain informed consent to comply with ethical guidelines. Due to the participants' age and possible access challenges, subjects can choose to take the survey on a tablet device with self-administration or through an interviewer-assisted format to include all individuals. All data will be securely stored with additional password protections.

3.4 Data Analysis

This paper utilizes a multi-dimensional quantitative approach. First, we will perform exploratory factor analysis (EFA) to explore the underlying structure of the DHL constructs, as well as the validity for the scale. Second, we will conduct confirmatory factor analysis (CFA) to confirm the measuring model with structural equation modeling (SEM). Finally, we will apply hierarchical regression models to evaluate our hypotheses by investigating the impact of DHL on WeChat effective use, while controlling for sociodemographic and health and health concerns. The method aligns with our interest in exploring whether DHL interacts with various factors to influence WeChat engagement, consistent with the Health Empowerment Model’s emphasis on contextual and personal enablers of agency. Given the heterogeneity in engaging in digital interaction behaviors of seniors, latent class analysis (LCA) will identify subtle variation of groups engagement on DHL and WeChat interaction. This technique will offer deep understandings of digital health behavior and inform individual-tailored strategies to intervention. Analysis will be performed in Mplus 8.0 and SPSS 27, which is common practice in the digital health literature.

4 Expected outcomes

This research is anticipated to provide important insights regarding the role of DHL in the effective use of WeChat for diabetes management among elderly patients. Findings from this study will reveal that DHL influences both the extent and quality of WeChat engagement and that the higher levels of literacy correspond to higher levels of proactive health management behaviors. In contrast, seniors with lower levels of DHL may adopt digital avoidance behaviors, may experience information overload, or may be prey to misinformation in the digital space, resulting in health risks rather than benefits.

The study is also expected to uncover that the capabilities of WeChat influence health management practice in complicated ways, in the forms of enablers and constraints. For example, the platform provides a way to access healthcare providers and health information easily, but the fact that the platform is open also makes misinformation easy to access, which necessitates a critical level of DHL to distinguish credible sources from those that are not. Finally, the research expects to identify different user typologies through LCA, allowing for the discovery of subgroups of elderly

patients who interact with WeChat differently based on their DHL levels and influences from their social environment.

The potential implications are twofold. Academically, this research will expand current literature by addressing the gap between DHL and engagement in digital health settings for diabetes management. Existing frameworks of DHL often treat literacy as fixed quality tied to the individual and the study will provide empirical evidence to articulate the evolving nature of DHL in digital health environments. By employing HEM, this research will provide a conceptualization of DHL that focuses on effective use associated with contextual learning, social interaction, and adaptability to technology that is critically important but is often overlooked in digital health literature.

From a practical perspective, the findings will provide valuable guidance for healthcare providers, policymakers, and technology developers in designing more inclusive and supportive digital health interventions for elderly populations. For hospitals and public health organizations, the study will inform the development of tailored digital health literacy programs, ensuring that seniors, particularly those with diabetes, receive adequate training to navigate WeChat-based health resources effectively. Policymakers can leverage these insights to bridge the digital health divide, implementing targeted interventions to enhance the accessibility and usability of online health platforms. By addressing both theoretical and practical gaps, this research will provide a robust empirical foundation for future studies on digital health literacy, diabetes management, and the role of social media platforms in elderly healthcare.

References

Andrade, A. C., Corrado, M., & Bertolucci, P. H. F. (2021). Health care professionals and use of digital technologies in patients with dementia and related disorders. *Alzheimer's & Dementia*, 17, e056667.

Chen, S. L., Chen, W. L., Ding, W. W., Li, X. H., Deng, H. M., & Huang, Q. B. (2024). Remote health management via WeChat to improve the quality of life of patients after PCI. *Quality Management in Healthcare*, *33*(3), 160-165.

Dadaczynski, K., Okan, O., Messer, M., Leung, A. Y., Rosário, R., Darlington, E., & Rathmann, K. (2021). Digital health literacy and web-based information-seeking behaviors of university students in Germany during the COVID-19 pandemic: cross-sectional survey study. *Journal of Medical Internet Research*, 23(1), e24097.

Ghanad, A. (2023). An overview of quantitative research methods. *International Journal of Multidisciplinary Research and Analysis*, *6*(08), 3794-3803.

Jacobs, R. J., Lou, J. Q., Ownby, R. L., & Caballero, J. (2016). A systematic review of eHealth interventions to improve health literacy. *Health Informatics Journal*, *22*(2), 81-98.

Ju, T. L., Chen, Y. Y., Sun, S. Y., & Wu, C. Y. (2006). Rigor in MIS survey research: In search of ideal survey methodological attributes. *Journal of Computer Information Systems*, *47*(2), 112-123.

Kim, H., & Xie, B. (2015). Health literacy and internet‐and mobile app‐based health services: A systematic review of the literature. *Proceedings of the Association for Information Science and Technology*, 52(1), 1-4.

Lee, M., & Nam, S. (2024). Telehealth utilization among patients with chronic disease: Insights from the 2022 Health Information National Trends Survey. *Journal of Telemedicine and Telecare*, 1357633X241289158.

Li R., Tuo J., & Wang, Y. (2020) Cognition on wearable monitoring devices for health care and its influencing factors among the elderly in Lanzhou city[J]. *Chinese Journal of Public Health*, 36(5): 702-705.

Liu, P., Yeh, L. L., Wang, J. Y., & Lee, S. T. (2020). Relationship between levels of digital health literacy based on the Taiwan digital health literacy assessment and accurate assessment of online health information: cross-sectional questionnaire study. *Journal of Medical Internet Research*, *22*(12), e19767.

Matthew-Maich, N., Harris, L., Ploeg, J., Markle-Reid, M., Valaitis, R., Ibrahim, S., ... & Isaacs, S. (2016). Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. *JMIR mHealth and uHealth*, *4*(2), e5127.

Nittayathammakul, V., Chatwattana, P., & Piriyasurawong, P. (2021). Conceptual framework of crowd context-based learning via IoT wearable technology to promote digital health literacy. In 2021 *5th International Conference on E-Society, E-Education and E-Technology* (pp. 113-119).

Norman, C. D., & Skinner, H. A. (2006). eHealth literacy: essential skills for consumer health in a networked world. *Journal of Medical Internet Research*, *8*(2), e506.

Pong, C., Tseng, R. M. W. W., Tham, Y. C., & Lum, E. (2024). Current Implementation of Digital Health in Chronic Disease Management: Scoping Review. *Journal of Medical Internet Research*, 26, e53576.

Reynolds, R., Dennis, S., Hasan, I., Slewa, J., Chen, W., Tian, D., ... & Zwar, N. (2018). A systematic review of chronic disease management interventions in primary care. *BMC family practice*, *19*, 1-13.

Schulz, P. J., & Nakamoto, K. (2013). Health literacy and patient empowerment in health communication: the importance of separating conjoined twins. *Patient education and counseling*, *90*(1), 4-11.

Sina (2024) Tencent Q3 Report: Combined Monthly Active Users of WeChat and WeChat Reached 1.382 Billion (in Chinese), *Sina Finance*, accessed on 16th March https://finance.sina.com.cn/tech/shenji/2024-11-13/doc-incvxhxn5497553.shtml

Van Der Vaart, R., & Drossaert, C. (2017). Development of the digital health literacy instrument: measuring a broad spectrum of health 1.0 and health 2.0 skills. *Journal of Medical Internet Research*, 19(1), e27.

Wang, H., Yuan, X., Wang, J., Sun, C., & Wang, G. (2021). Telemedicine maybe an effective solution for management of chronic disease during the COVID-19 epidemic. *Primary Health Care Research & Development*, *22*, e48.

Xu, Y., Lu, J., Li, M., Wang, T., Wang, K., Cao, Q., ... & Ning, G. (2024). Diabetes in China part 1: epidemiology and risk factors. *The Lancet Public Health*.

Yang, J., Shen, Q., Tong, X., & Mukhopadhaya, P. (2025). The impact of digital literacy in enhancing individuals' health in China. *BMC Public Health*, *25*(1), 1-22.

Yang, K., Hu, Y., & Qi, H. (2022). Digital health literacy: bibliometric analysis. *Journal of Medical Internet Research*, 24(7), e35816.

Zou, P., Huang, A., Luo, Y., Tchakerian, N., Zhang, H., & Zhang, C. (2023). Effects of using WeChat/WhatsApp on physical and psychosocial health outcomes among oncology patients: a systematic review. *Health Informatics Journal*, *29*(1), 14604582231164697.

21 TOWARDS A HOLISTIC CONCEPTUAL FRAMEWORK FOR SUPPLY CHAIN AND THIRD-PARTY CYBERSECURITY RISK MANAGEMENT

Yusuf Akçakaya
Bogazici University
yusuf.akcakaya@bogazici.edu.tr

Meltem Mutlutürk
Bogazici University
meltem.mutluturk@bogazici.edu.tr

Bilgin Metin
Bogazici University
bilgin.metin@bogazici.ed

Abstract
As organizations deepen their reliance on external service providers and digital supply chains, third-party and supply chain cybersecurity risk management has become a strategic priority. While numerous international documents—such as standards, frameworks, and regulations—address aspects of third-party risk, their fragmented nature poses integration challenges, making implementation complex and inconsistent. To address this challenge, this study consolidates these documents related to third-party and supply chain cybersecurity. It then introduces a holistic conceptual framework to address this fragmentation by aligning common control themes across the vendor lifecycle. By synthesizing diverse sources into a unified and structured model, the study offers practical value for risk managers, internal auditors, and security leaders seeking to harmonize compliance obligations with effective third-party and supply chain risk management.

Keywords
Third-party risk, Third-party cybersecurity, supply chain risk, supply chain cybersecurity, cybersecurity, framework.

1 Introduction
Digitalization in recent years has greatly increased organizations' reliance on third-party services and digital supply chains (Metin *et al.,* 2024). These third-party providers play a critical role in enabling organizations to access advanced technologies, specialized expertise, and cost-effective services without the need to build and maintain these capabilities in-house (Akinrolabu & New, 2017; Benaroch, 2021). For instance, a business may hire a third-party supplier to provide raw materials or a third-party consultant to offer expertise (Tran, 2024). These third-party relationships can take many forms, from cloud computing and data center hosting to software development and cybersecurity services. For example, organizations may rely on cloud service providers like Amazon Web Services or Microsoft Azure to deliver scalable computing power, storage, and

software-as-a-service solutions, allowing them to focus on their core business functions (Jewhurst, 2022).

Recent reports show that around 60% of organizations work with more than 1,000 third parties (Deloitte, 2023). This number illustrates modern businesses' dependence on external providers to keep their operations running smoothly. Growing dependence on third parties means that companies must proactively manage these risks, particularly as cyber threats become more advanced and widespread (Cox, 2024; Shetty, 2022). KPMG survey showed that 73% of participants have faced at least one major disruption attributed to a third-party relationship during the previous three-year period (KPMG, 2022). As a result, many companies shifted their focus to developing more robust Third-Party Risk Management frameworks that address traditional risks, such as financial instability and operational failures, alongside newer cybersecurity-related risks (Gartner, 2023). Managing the digital and cybersecurity risks associated with third parties has become one of the greatest challenges for businesses (EY, 2023). Consequently, cybersecurity has become one of the most significant aspects of Third-Party Risk Management. This calls for a robust strategy examining third parties' financial and operational stability and their cybersecurity measures (Keskin *et al.*, 2021).

In this context, it is important to highlight the difference between supply-chain cybersecurity and third-party cybersecurity. Supply chain cybersecurity focuses specifically on risks within the product/service lifecycle (e.g., software vendors, component suppliers, manufacturers, logistics). On the other hand, third-party cybersecurity encompasses all external vendors, including those outside the supply chain (e.g., law firms, marketing agencies), making it broader in scope. This distinction is critical, as 75% of third-party breaches target the software supply chain, yet 29% of all breaches originate from non-supply-chain third parties (e.g., service providers) (SecurityScorecard, 2024).

Although third-party cybersecurity risks have gained increasing attention, with the development of various standards, guidelines, frameworks, and regulations, there remains a lack of an integrated and comparative overview of these documents that organizations use to manage third-party cybersecurity risks throughout the entire vendor lifecycle.
This study aims to synthesize and compare prominent international standards, guidelines, frameworks, and regulations to establish a unified reference point with a conceptual framework for third-party cybersecurity risk management by addressing a key research question:

RQ: Which internationally recognised documents collectively provide comprehensive guidance for managing third-party cybersecurity risk, and how can they be aligned for a holistic framework for a third-party lifecycle?

Due to the limitations of this research-in-progress format, this study focuses on supply-chain cybersecurity and Third-Party Cybersecurity Risk Management (TPCRM) documents, including standards, frameworks, guidelines, and regulations that are prominent in both academic discussions and industry practices. This paper provides risk managers, internal auditors, and IT professionals with a cohesive reference that connects operational needs with compliance requirements. Additionally, it culminates in a comprehensive conceptual TPCRM framework that integrates governance, risk assessment, contractual safeguards, continuous monitoring, and incident response planning into a single lifecycle model. This unified framework transforms fragmented guidance into an actionable roadmap, enabling organizations to implement scalable,

standards-aligned third-party risk programs while ensuring regulatory compliance and enhancing operational resilience.

2 Methodology

This study uses a two-stage method. First, a narrative literature search, guided by one author's 15 years of audit-consultancy practice, identified the standards, frameworks, best-practice guides, and regulations most often cited in professional and regulatory sources on third-party cybersecurity risk. Grey literature from major industry reports (Deloitte, EY, Gartner, BitSight, RiskRecon, SecurityScorecard) was screened. English documents updated 2018-Mar 2025. Setting the cut-off at 2018 captures the wave of post-GDPR supply-chain-focused revisions (General Data Protection Regulation, Regulation (EU) 2016/679) and excludes superseded versions that are no longer cited in practice. Items limited to a single technology or product category were excluded. In the second stage, a structured scoping-review procedure, adapted from Arksey and O'Malley (2005) and Levac et al. (2010), was conducted. Guided by our review question, the second stage applied a structured scoping-review procedure adapted from Arksey & O'Malley (2005) and Levac et al. (2010). Each retained document was decomposed into its control statements; two authors independently coded these statements to the nine life-cycle phases of our holistic TPCRM framework, reconciling differences by discussion. The resulting crosswalk of controls to phases is presented in Table II and Fig. 1.

3 Related Standards, Guidelines, Frameworks, And Regulations

In the literature, there have been several attempts to consolidate important cybersecurity standards frameworks (Cisternelli, 2024; Kosling, 2024; Ryerse, 2024). In this study, we bring together standards, frameworks, guidelines, and regulations and then examine them using a structured scoping review approach, as shown in Table I.

Name	Organization	Type
ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection - Information security management systems – Requirements	ISO/IEC	Standard
ISO/IEC 27036 Cybersecurity – Supplier relationships	ISO/IEC	Standard/Guideline
NIST SP 800-53 Rev. 5 Security and Privacy Controls for Information Systems and Organizations	NIST	Guideline
NIST SP 800-161 Rev. 1 Cybersecurity Supply Chain Risk Management Practices for Systems and Organizations	NIST	Guideline
NIST Cybersecurity Framework v2	NIST	Framework
NIS2 Implementation Regulation	EU	Directive
The Digital Operational Resilience Act	EU	Regulation
Good Practices for Supply Chain Cybersecurity	ENISA	Guideline
CIS Critical Security Controls	Center for Internet Security	Best Practices
Third-Party Cybersecurity Playbook	Riskrecon	Guideline (Playbook)

Table I. Third-Party Cybersecurity Standards, Guidelines, Frameworks, And Regulations

ISO/IEC 27001:2022 serves as an internationally recognized standard for establishing and improving an Information Security Management System (ISMS), with specific controls addressing

third-party risks through Annex A Controls 5.19-5.21. These controls mandate organizations to establish procedures for managing risks throughout supplier relationships, focusing on clear contractual agreements, continuous monitoring, and secure termination processes. ISO/IEC 27036 is a multi-part standard focused explicitly on managing information security risks in supplier relationships throughout their entire lifecycle. Divided into four parts addressing different aspects of supplier security (overview, requirements, supply chain security, and cloud services), it offers targeted guidance on everything from risk analysis before entering relationships to incident management and role definition. Its strength lies in its specialized focus on supplier relationships (ISO/IEC 27036, 2021). NIST SP 800-53 Rev. 5 provides a comprehensive catalog of security and privacy controls for information systems, with Revision 5 introducing a dedicated Supply Chain Risk Management control family comprising 12 distinct controls. These controls emphasize using formal risk management plans, ensuring transparency of systems throughout their lifecycle, and integrating privacy and security considerations collaboratively (NIST SP 800-53, 2020).

NIST SP 800-161 Rev. 1 (Boyens *et al.,* 2024) offers specialized guidance for Cybersecurity Supply Chain Risk Management, complementing NIST SP 800-53 Rev. 5 with more granular practices for both IT and OT environments and integrating into broader risk management activities and categorizes practices by maturity level (Foundational, Sustaining, and Enhancing). Its comprehensive approach to multi-tiered supply chains is valuable, though gaining visibility into complex supply chains remains challenging despite the framework's guidance. (Boyens et al., 2024)

The NIST Cybersecurity Framework v2 (NIST CSF V2) provides a flexible, risk-based approach with the addition of a Governance function that includes a dedicated Cybersecurity Supply Chain Risk Management category. This category encourages organizations to integrate supply chain security into broader risk management efforts, establish clear supplier risk priorities, and develop appropriate plans for the entire vendor relationship lifecycle. (National Institute of Standards and Technology, 2024).

The Network and Information Systems 2 (NIS 2) Directive (European Commission NIS2, 2022a) establishes a mandatory legal framework across the EU, requiring organizations in essential and important sectors to implement appropriate cybersecurity risk management measures, emphasizing third-party risks. Besides NIS 2, the NIS2 Implementing Regulation (2024) provides guidance through risk assessments of vendors, continuous monitoring, and implementation of contract security requirements. As a regulation with potential financial penalties for non-compliance, it establishes a strong baseline but may create challenges in coordinating with other regulatory requirements. The Digital Operational Resilience Act (DORA) focuses on enhancing operational resilience against digital disruptions in the EU financial sector (European Commission DORA, 2022b). Financial entities must assess their reliance on ICT providers, identify critical vendors, establish clear contractual obligations, monitor third parties, and develop exit strategies to ensure business continuity. While creating stringent requirements for this specific sector, its narrow focus limits its direct applicability to organizations in other industries (European Commission DORA 2022b). Furthermore, the "ENISA Guideline on Good Practices for Supply Chain Cybersecurity" provides comprehensive, non-binding recommendations for enhancing ICT supply chain security. It emphasizes establishing corporate-wide systems for third-party risk management, conducting thorough vendor assessments, implementing vulnerability handling procedures, and fostering

transparency between organizations and suppliers (European Union Agency for Cybersecurity, 2023). Moreover, the Center for Internet Security Critical Security Controls (CIS Controls) offers a prioritized set of best practices, with Control 15 specifically addressing service provider management. This control includes practical guidance on inventorying service providers, classifying them by risk, establishing security requirements in contracts, conducting due diligence, and monitoring for threats. Its straightforward, actionable approach makes it accessible even for organizations with limited cybersecurity expertise, though it may lack the comprehensiveness of more specialized frameworks (Center for Internet Security, 2024).

4 Presented Holistic Conceptual Third-Party Cybersecurity Risk Management Framework

The proposed holistic conceptual TPCRM framework, as illustrated in Figure 1 and detailed in Table II, provides a comprehensive and standards-aligned model for managing third-party cybersecurity risks. This framework is grounded in an integrated analysis of leading international standards and practices, including DORA, the NIS2 Implementing Regulation (2024), NIST CSF V2, NIST SP 800-53 Rev. 5, NIST SP 800-161 Rev.1, ISO/IEC 27036, ISO/IEC 27001:2022, CIS Controls v8, ENISA Good Practices, and practical insights from implementation playbooks (Riskrecon, 2018).

The framework is composed of nine interrelated lifecycle phases that together form a dynamic and risk-informed process model. Each phase consolidates recurring control themes identified across multiple sources. It begins with the establishment of governance and strategy, followed by third-party identification, and continues with inherent risk assessment, contractual engagement through contracting and procurement, and secure integration via implementation and onboarding mechanisms. Subsequent phases include continuous risk monitoring, incident management, and periodic review and termination planning, before concluding with reporting and assurance activities that support internal oversight and regulatory compliance.

Each lifecycle phase is linked to detailed key activities and mapped to authoritative sources from regulatory frameworks and technical standards, ensuring alignment with supervisory expectations and best practices across sectors. The framework enables organizations to operationalize third-party cybersecurity risk management in a scalable, repeatable, and end-to-end manner, supporting continuous adaptation to evolving threats and vendor ecosystem complexity. Through this lifecycle approach, it ensures both strategic resilience and operational maturity across the full third-party relationship lifecycle.

Fig.1: Holistic Conceptual Third-Party Cybersecurity Risk Management Framework

5 Conclusion

This research-in-progress study consolidates key standards, frameworks, best practices, and regulations relevant to third-party and supply chain cybersecurity. For internal auditors, risk managers, and IT leaders, engaging with these sources is no longer merely a technical obligation—it is a strategic imperative in navigating today's complex and interconnected threat landscape. Building on this foundation, the study proposes a holistic conceptual framework for third-party cybersecurity risk management that aligns diverse guidance into a unified model. Future research could explore the practical challenges of operationalizing this framework across different sectors and assess its real-world impact on reducing third-party cybersecurity risks. In addition, empirical validation through case studies or sector-specific implementations may yield actionable insights and enhance the framework's adaptability across varied organizational contexts.

Lifecycle	Key Activities	Mapped Sources
Governance & Strategy	Policy definition, roles & responsibilities, stakeholder engagement, multi-vendor strategy	NIST CSF 2.0 (GV Func.); ISO 27001:2022 (Cl. 5.2, 5.3, 4.2, A.5.19); DORA (Art. 5, 6, 28(1)); NIS 2 Imp Reg (Annex 1.1, 1.2); ENISA Good Pr. (Sec. 3.1); ISO 27036-1 (Sec. 5.3, 5.5); RiskRecon (Governance); NIST SP 800-161r1 (Ch. 2, Apdx C)
Third-Party Identification	Discovery of vendors, onboarding triggers, initial risk categorization	RiskRecon (Third Party Identification); DORA (Art. 28(2), 28(3)); NIST SP 800-161r1 (Sec. 2.3.2, Apdx C); ENISA Good Pr. (Sec. 3.2); ISO 27036-1 (Sec. 5.2, 6.2.2); NIST CSF 2.0 (ID.AM); ISO 27001:2022 (A.5.19); NIS 2 Imp Reg (Annex 2.1); CIS (15.1)
Risk Assessment	Inherent risk analysis, profiling, due diligence, questionnaires, continuous risk analysis	RiskRecon (Risk Assessment); DORA (Art. 8, 28(4), 29, 30); NIST CSF 2.0 (ID.RA, GV.SC, GV.RM); ISO 27001:2022 (Cl. 6.1.2, 8.1, A.5.20, A.5.21); ENISA Good Pr. (Sec. 3.2); NIST SP 800-161r1 (Sec. 2.3.3, Apdx D); NIST SP 800-53r5 (RA, SA families); CIS (Ctrl. 15.3, 15.5); NIS 2 Imp Reg (Annex 2.1)
Contracting & Procurement	Security clauses, right to audit, exit strategies, Service Level Agreement/ KPI formalization	DORA (Art. 30, 28); RiskRecon (Legal and Procurement); CIS (Ctrl. 15.4); ISO 27001:2022 (A.5.20, A.5.21, A.5.22); ENISA Good Pr. (Sec. 3.3); NIS 2 Imp Reg (Annex 5.1); ISO 27036-1 (Sec. 6.2); NIST SP 800-161r1 (Apdx C)
Implementation & Integration	Secure onboarding, access provisioning, secure development integration (e.g. CI/CD), architecture mapping	ISO 27001:2022 (Cl. 8, A.5.23, A.8.2, A.8.25); NIST CSF 2.0 (PR.AA, PR.PS, ID.AM); CIS (Ctrl. 4.6, 16.1); NIST SP 800-53r5 (AC, SA, CM families); ENISA Good Pr. (Sec. 3.3); RiskRecon (Risk Treat.); NIST SP 800-161r1 (Sec. 2.3.4, Apdx C - Integrating); NIS 2 Imp Reg (Annex 6.1)
Continuous Monitoring	Surface risk assessments, threat intelligence, patching, vulnerability handling, fourth-party awareness	RiskRecon (Monit. & Resp.); NIST CSF 2.0 (DE.CM); DORA (Art. 9, 10, 16); ISO 27001:2022 (Cl. 9.1, A.5.7, A.8.8, A.5.22); ENISA Good Pr. (Sec. 3.4); CIS (Ctrl. 7, 8.12, 13); NIST SP 800-161r1 (Sec. 2.3.5, Apdx C - Monitoring); NIST SP 800-53r5 (RA-5, SI-2, SR-5); NIS 2 Imp Reg (Annex 2.2, 3.2, 5.1.6, 5.1.7)
Incident Management	Reporting mechanisms, coordination, response plans, communication, post-incident review	DORA (Art. 17-21); NIST CSF 2.0 (RS Func., RC Func.); ISO 27001:2022 (A.5.24-A.5.27, Cl. 10.2); CIS (Ctrl. 17); NIS 2 Imp Reg (Annex 3); ENISA Good Pr. (Sec. 3.3); NIST SP 800-53r5 (IR Family); RiskRecon (Dangerous Condition Hunting)
Review & Termination	Periodic reassessment, offboarding, data destruction, risk reclassification	RiskRecon (Risk Resource Management); DORA (Art. 28(7), Art. 30(2); ISO 27001:2022 (Cl. 9.1, 9.3, A.5.22, A.8.10); CIS (Ctrl. 15.7); NIST CSF 2.0 (GV.OV, GV.SC); ENISA Good Pr. (Sec. 3.3); ISO 27036-1 (Sec. 6.2 - implies review/termination); NIS 2 Imp Reg (Annex 5.1.4, 12.5)
Reporting & Assurance	Reporting to regulators, risk dashboards, tiering updates, audit trail maintenance	DORA (Art. 5, 6, 8, 19); ISO 27001:2022 (Cl. 9, 7.5); NIST CSF 2.0 (Profiles & Tiers); RiskRecon (Ch. 7); NIS 2 Imp Reg (Annex 2.2, 2.3); CIS (Ctrl. 8); ENISA Good Pr. (Sec. 3.5)

Table II. The Holistic Framework: Phases, Key Activities, and Mapped Sources

References

Akinrolabu, O., & New, S. (2017). Can Improved Transparency Reduce Supply Chain Risks in Cloud Computing? *Operations and Supply Chain Management: An International Journal*, 130–140. https://doi.org/10.31387/oscm0280185

Arksey, H. & O'Malley, L. (2005). Scoping studies: towards a methodological framework. International Journal of Social Research Methodology, 8(1), 19-32.

Benaroch, M. (2021). Third-party induced cyber incidents—Much ado about nothing? *Journal of Cybersecurity*, *7*(1), tyab020. https://doi.org/10.1093/cybsec/tyab020

Boyens, J., Smith, A., Bartol, N., Winkler, K., Holbrook, A., & Fallon, M. (2024). *Cybersecurity Supply Chain Risk Management Practices for Systems and Organizations*. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-161r1-upd1

Center for Internet Security (2024), *CIS Critical Security Controls v8.1*. https://www.cisecurity.org/controls (Accessed 2025-06-01)

Cisternelli, E. (2022, August 15). *7 cybersecurity frameworks that help reduce cyber risk*. *Bitsight Blog*. https://www.bitsight.com/blog/7-cybersecurity-frameworks-to-reduce-cyber-risk (Accessed 2025-06-01)

Cox, O. (2024, June 3). *Exploring the benefits and risks of third-party data solutions*. *Darktrace Blog*. https://www.darktrace.com/blog/exploring-the-benefits-and-risks-of-third-party-data-solutions (Accessed 2025-06-01)

Deloitte. (2023). *The rising importance of third party risk management (TPRM)*. Deloitte. https://www2.deloitte.com/us/en/pages/risk/articles/supply-chain-risk-management-strategies.html (Accessed 2025-06-01)

European Parliament & Council. (2016). *Regulation (EU) 2016/679: General Data Protection Regulation*. *Official Journal of the European Union*, L 119.

European Commission NIS2 (2022a). *Directive (EU) 2022/2555 of the European Parliament*. European Union. https://eur-lex.europa.eu/eli/dir/2022/2555/oj

European Commission DORA (2022b). *Regulation (EU) 2022/2554 of the European Parliament*. European Union. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022R2554

European Union Agency for Cybersecurity. (2023). *Good practices for supply chain cybersecurity.* Publications Office. https://data.europa.eu/doi/10.2824/805268

EY. (2023). *2023 EY Global Third-party Risk Management Survey*, https://www.ey.com/en_gl/insights/risk/2023-ey-global-third-party-risk-management-survey (Accessed 2025-06-01)

Gartner. (2023). *Third-Party Risk Management Benchmarking Report 2023*. Gartner. https://emt.gartnerweb.com/ngw/globalassets/en/legal-compliance/documents/trends/third-party-risk-management-benchmarking-report-2023.pdf (Accessed 2025-06-01)

ISO/IEC 27036 (2021). *ISO/IEC 27036-1:2021 Cybersecurity – Supplier relationships – Part 1: Overview and concepts*.

ISO/IEC 27001 (2022). *ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection – Information security management systems – Requirements. International Organization for Standardization.*

Jewhurst, H. (2022). *Who Is Considered a Third Party or Vendor?, Venminder Blog,* https://www.venminder.com/blog/who-considered-third-party-vendor (Accessed 2025-06-01)

Keskin, O. F., Caramancion, K. M., Tatar, I., Raza, O., & Tatar, U. (2021). Cyber Third-Party Risk Management: A Comparison of Non-Intrusive Risk Scoring Reports. *Electronics*, *10*(10), 1168. https://doi.org/10.3390/electronics10101168

Kosling, K. (2024). *What Are 5 Top Cybersecurity Frameworks?* https://www.itgovernanceusa.com/blog/top-4-cybersecurity-frameworks (Accessed 2025-06-01)

KPMG. (2022). *Third-Party Risk Management Outlook 2022*. KPMG. https://assets.kpmg.com/content/dam/kpmg/cn/pdf/en/2022/03/third-party-risk-management-outlook-2022.pdf (Accessed 2025-06-01)

Levac, D., Colquhoun, H., & O'brien, K. K. (2010). Scoping studies: advancing the methodology. *Implementation science*, *5*, 1-9.

Metin, B., Özhan, F. G., & Wynn, M. (2024). Digitalisation and Cybersecurity: Towards an Operational Framework. *Electronics*, *13*(21), 4226. https://doi.org/10.3390/electronics13214226

NIST CSF V2. (2024). *The NIST Cybersecurity Framework (CSF) 2.0* (NIST CSWP 29; p. NIST CSWP 29). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.CSWP.29

NIS2 Implementing Regulation (2024). Commission Implementing Regulation (EU) 2024/2690 laying down rules for the application of Directive (EU) 2022/2555 as regards technical and

methodological requirements of cybersecurity risk-management measures and further specification of the cases in which an incident is considered to be significant . https://eur-lex.europa.eu/eli/reg_impl/2024/2690/oj

NIST SP 800-53r5 (2020). *Security and Privacy Controls for Information Systems and Organizations* (Revision 5). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-53r5

RiskRecon (2018). Third-party security risk management playbook. https://www.riskrecon.com/third-party-security-risk-management-playbook (Accessed 2025-06-01)

Ryerse, J. (2024, May 29). *Top 11 cybersecurity frameworks*. *ConnectWise Blog*. R, (Accessed 1 June 2025) https://www.connectwise.com/blog/11-best-cybersecurity-frameworks

Shetty, P. (2022). Third Party Risks and their Impact to the Overall Risk Posture, how to Improve it? *Journal of Mathematical & Computer Applications*, *1*(4), 1–3. https://doi.org/10.47363/JMCA/2022(1)127

SecurityScorecard. (2024, February 28). Global third-party cybersecurity breach report. https://securityscorecard.com/company/press/global-third-party-risk-report/

Tran, T. (2024). *What Is A Third-Party Contractor? Definition And Significance*. https://eraworkforcesolutions.com/third-party-contractor/ (Accessed 2025-06-01)

22R. AI'S LIMITS IN GENERATING AUDIO-VISUAL MEDIA FOR COMMUNICATION OF SERIOUS CONTENT

Ines Janusch
Johannes Kepler University Linz
ines.janusch@jku.at

Simon Currle
Johannes Kepler University Linz
simon.currle@jku.at

Abstract

Generative AI, in particular large language models such as ChatGPT, have enabled broad application of AI in everyday life. While text-generating AI approaches are already widely used by the general public, the application of audio-visual generative AI is lagging behind. Audio-visual generative AI seems particularly suitable for knowledge sharing purposes, i.e., to provide serious content. Thus, this paper evaluates state-of-the-art audio-visual generative AI approaches, i.e., text-to-video, and their suitability for communicating serious content effectively. In a small proof-of-concept, several self-proclaimed AI-based video generation tools have been tested. Based on a sample video, we also assessed the uncanniness of AI-generated videos, to clarify their applicability for serious content, i.e., training and education, as we assume this might influence the perception of the content.

Keywords

AI-generated videos, communication of serious content, virtual instructors

1 Introduction

AI has changed the current discussion on what is possible and what is not. With the public release of OpenAI's large language model (LLM) ChatGPT in late 2022, generative AI (GenAI) found its way into everyday life. While this was the starting point for many people to "play around" with AI, in particular LLMs, audio-visual applications (e.g., for generating videos) have mainly been discussed in the context of their harmful effects such as deep fakes (Gregory, 2023; Roy & Raval, 2023) or their impact on the creative industry, whether economical (i.e., artists being replaced by AI), ethical (i.e., AI mimicking an artist's style without consent and for nefarious purposes), or socio-cultural (i.e., biases in generative art) (Jiang et al., 2023; Srinivasan & Uchino, 2021). At the same time, video has become one of the most powerful learning media (Sablić et al., 2021) and audio-visual GenAI seems to have the potential for widespread adoption for knowledge sharing purposes e.g., for training and education. For instance, Hossain et al. (2025) have explored the application of GenAI to improve student engagement and prevent teachers' burnout.

In this research in progress paper, we aim to identify the possibilities and limitations of state-of-the-art audio-visual GenAI to create videos that provide serious content based on text-to-video (T2V) approaches and their applicability for e.g., training and education. To provide a proof-of-concept, we tested several self-proclaimed AI-based video generation tools and developed a sample video. Our research shows a lack of GenAI tools and approaches that are suitable for the creation of audio-visual serious content of a particular length beyond a few seconds. In the context of the communication of serious content, the available tools either over-promise or under-deliver: solutions that seem suitable for T2V creation for educational or training purposes market their videos as AI-generated although they do not employ GenAI approaches, while the current

limitations of true GenAI solutions (i.e., prompt length and video duration) prohibit their application in the context of knowledge sharing. Moreover, we determined the need for further evaluation regarding the impact of GenAI videos on the viewers' emotional responses, engagement, and learning outcomes (i.e., retention and transfer) when applied in an education context.

The remainder of this paper is structured as follows: in Section 2 we provide an overview of existing T2V approaches. Next, we show the methodological approach for our proof-of-concept and its results, a discussion and limitations leading to an outlook towards future work.

2 Video Generation and Text-to-Video Approaches

Based on a review of existing tools and approaches, we identified two very different directions of T2V generation for the purpose of transferring serious content: on the one hand, 3D-model-driven video synthesis and on the other GenAI approaches.

2.1 3D-Modeling-Based Approaches

3D modelling-based approaches to create videos of humans in motion predate GenAI-based approaches by more than a decade (Yu et al., 2015) and provide fewer degrees of freedom in terms of the appearance of the character or the overall scene. However, they may ensure other requirements of structural correctness. While these approaches have originally required complex multi-camera 3D capture studios, non-rigid 3D modelling from monocular video allows to capture 3D models of humans in motion from videos captured with a single commodity camera (Tome et al., 2017). 3D-modeling-based approaches are particularly well suited for the modelling of 3D faces in speech and are thus in use in available commercial T2V solutions for the creation of photo-realistic avatars.

2.2 Generative AI Approaches

Generative T2V models require two main components, a vision generator, and a language interpreter. Throughout this paper, we will focus on vision generation architectures only to discuss how different architectures affect performance. The most popular architectures employed for video generation purposes are generative adversarial networks (GANs), autoregressive-based models, and most recently diffusion models. GANs have been around for quite some time and are used, e.g., in the tool StoryGAN (Li et al., 2019). GANs consist of two networks, a discriminator and a generator that are competing against each other. The discriminator is trained to correctly identify training samples as genuine samples and those produced by the generator as fake while the generator is trained to minimize the probability of its samples being identified as fake (Bond-Taylor et al., 2022). Despite their challenges (vanishing gradients, mode collapse, mode jumping, slow convergence, prone to oversampling (Bond-Taylor et al., 2022; Gulrajani et al., 2017)), GANs are capable of producing vivid, realistic images and videos and allow for efficient sampling of high-resolution images with good perceptual quality (Cho et al., 2024; Karras et al., 2020). Nevertheless, due to the limitations of GANs, the research focus has shifted to autoregressive architectures, used e.g., in CogVideo (Hong et al., 2022). The basic idea of the architecture is to decompose the joint probability distribution into a product of conditional probability distributions which can be modelled as a one-dimensional sequencing problem, where the value of the next pixel is determined by all the previously generated pixels and is fed back into the network (van den Oord et al., 2016). However, due to the sequential sampling process and the computationally demanding architecture, autoregressive models are limited to low-resolution images and have long

training times (Rombach et al., 2022). Nevertheless, state-of-the-art results in text-to-image (T2I) and T2V generation were achieved (Wu et al., 2022). Diffusion models are a family of probabilistic generative models that are e.g., used in the video generating tools Imagen Video (J. Ho et al., 2022), and runway (Esser et al., 2023). The fundamental element of diffusion models is the gradual perturbation of the data by the addition of noise, referred to as the "diffusion" process (Yang et al., 2023). After a certain number of steps, the process is reversed, with noise being gradually removed to generate new data.

2.3 Video Generation for Serious Content

Videos are used in various ways, in particular, to provide content of different types. Based on the classification from serious gaming, we differentiate between entertaining (or leisure) and serious content (Mildner & Mueller, 2016) e.g., content for training and education. Guo et al. (2014) analysed how video production decisions affect student engagement in online educational videos and provided recommendations for instructional design:

- Shorter videos (max. 6 minutes) are much more engaging.
- Videos that intersperse an instructor's talking head with slides are more engaging.
- Videos where instructors speak fairly fast and with high enthusiasm are more engaging.
- Tutorials should be optimized for re-watching and skimming.

What remains unclear though, is the question, how AI-generated videos are perceived by the recipients. As it has been widely discussed in the context of robots, there is a specific point - the so-called uncanny valley - that determines whether or not people feel comfortable with artificially generated human-like artifacts (Mori et al., 2012). The same seems to apply to AI-generated videos meant to transfer knowledge.

3 Proof-of-Concept

To simulate an end-to-end use of AI for sharing serious content via audio-visual media, we generated a short instructional video showing a virtual person who explains serious content in an audio-visual way. We also assessed how the final sample video is perceived by students.

3.1 Development of the Video

Based on the suggestions by Guo et al. (2014), we developed a scenario, i.e., the process of enrolment to courses at a university, using a particular campus management tool that sets the conditions for the final video. This scenario was selected as it fulfils several conditions: it does not need any intervention, the boundaries (start and end) are clear, and a process can be derived and modelled beforehand. In addition, the content follows a certain logic and does not need a lot of mimic and gesture from the virtual instructor. The enrolment process to be explained was outlined in the form of a process model using the business process modelling and notation (BPMN) syntax. ChatGPT4 was used to prepare a video script explaining the software process based on the BPMN model and screenshots of the software's user interface. The video script and screenshots were then used to generate a talking head educational video.

Tool	Video Library	Avatars	GenAI
invideo	✗	✗	
Fliki	✗		
Visla	✗		
HourOne		✗	
Synthesia		✗	
runway			✗

Table 16: Tools categorized according to their methodology

For this purpose, the following tools were analysed and tested: invideo (https://invideo.io), Fliki (https://fliki.ai), Visla (https://www.visla.us), HourOne (https://hourone.ai), Synthesia (https://www.synthesia.io/), and runway (https://runwayml.com). Based on their underlying methodology (see Table 16), these tools can be grouped into three categories: 1) Tools building upon stock media libraries. Although they are marketed as AI video creation, they do not employ visual GenAI. AI support is limited to the suggestion of suitable stock media, voice overs using AI-generated voices, or video editing. 2) Tools focusing on avatars created based on human models using pre-recorded footage (images or videos) of the human models. Their T2V approach does not include GenAI but rather synchronizes lip movements and/or gestures to the input video script and the content to be presented by the avatar. 3) Tools using GenAI approaches (e.g., diffusion models in the cases of runway (Esser et al., 2023) to generate short videos based on a text prompt.
The final sample video was created in January 2024 applying the tool provided by Synthesia using one of their existing avatars. The video is almost three minutes long and shows the talking head avatar (modelled after a male human actor) with a male-perceived voice in the right half of the video frame. screenshots visualizing the software process are shown simultaneously (left half and background of the frames). A single frame of the video is shown in Figure 9.

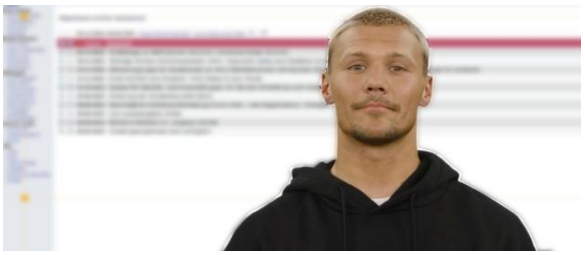

Figure 9: Example frame from the proof-of-concept video developed.

It is important to note that Synthesia describes its videos as AI video-from-text. However, the tool does not employ GenAI approaches as such for generating avatars. Synthesia's avatars are rather modelled based on human actors and actresses using non-rigid 3D modelling and 3D human pose estimation (Tome et al., 2017). Nevertheless, Synthesia was used in this experiment as it provided avatars as well as the option to use a complete video script as input. True GenAI approaches (e.g., models provided by Sora or Runway) could not be used due to their limited prompt length and video duration. We also investigated other approaches and tools to overcome human interaction, aiming at directly developing AI-generated videos from text or a process model, however, it was impossible to reach the same results.

3.2 Perception of the Video

We exposed three cohorts of students to the video (cohort 1: 53 people, 03/2024; cohort 2: 27 people, 04/2024; cohort 3: 42 people, 04/2024; in total: n = 122). People from cohort 1 were attending an informational lecture before they started their studies and thus have not had any experience with the campus management system described. Cohorts 2/3 were attending courses at the university where the software in the video is used. We showed the video and afterwards asked questions regarding the subjects' emotional responses: four questions taken from a scale developed for measuring uncanniness (C.-C. Ho & MacDorman, 2010) (machine-like - human-like, unpleasant - pleasant, unfriendly - friendly, incompetent - competent; using the five-point Likert scale); two questions about the acceptance of AI-generated videos (Would you accept AI-generated videos instead of a human-made videos? - yes / no / undecided) and (How long did you feel uncomfortable with the video? - not at all / only at the beginning / for a while / until the end of the video); one open question (In general, how do you feel about the video?).

4 Results

The four questions from the uncanniness measuring scale (C.-C. Ho & MacDorman, 2010) showed, that 78% of the 122 participants perceived the video to be "rather machine-like". Only 7% of the participants perceived the video to be "rather human-like", none of them experienced it as human-like, and the rest identified it as "neither - nor". For the scale "unpleasant - pleasant" most students (85%) answered "neither - nor", only 5% identified it as "rather unpleasant", and the rest of the students experienced it as "rather pleasant". Interestingly, almost all participants found the video to be "rather friendly" (88%) or even "friendly" (7%), and 5% answered "neither - nor". What is even more interesting is the perception regarding competence, i.e., "incompetent - competent". Most of the students found the AI-generated talking head to be "rather competent" (75%) and only 3% found it "rather incompetent" (the remaining 22% answered "neither - nor"). Regarding acceptance, the picture remains unclear, as 42% of the students "would accept an AI-generated video", and 38% "would not accept such a video" (20% remain "undecided"). Most students felt uncomfortable "for a while" (88%), and only 3% felt "uncomfortable the whole time", 9 % "only at the beginning" (one student answered "not at all"). Only 37 students answered the open question. However, in analysing these question, we found some interesting statements showing that students expect AI-generated videos to be an upcoming tool in teaching ("We for sure will have AI-generated videos in classes soon."), are not afraid of the technology ("It is just a video, it can't do any harm.") or were impressed by the quality ("I did not expect it to be that good."). Most students answering this question mentioned that the avatar had "a weird mimic", "too long pauses between sentences" and "he did not move, e.g., his hands". Some answers stated that the competence of the talking head depends on who created it. Interestingly, all answers referred to the talking head as "he". Correlation and T-tests revealed no causal relationships and no differences between the cohorts.

5 Discussion

Current T2V approaches have achieved strong performance regarding the generation of general objects in an artistic way or for entertainment. However, current AI-based T2V approaches (e.g., the newest models of Sora and Runway) only generate videos of a few seconds, up to one minute (Runway, 2025; *Sora*, 2025), and are not suitable for instructional videos due to their limited video duration. Moreover, the available T2V approaches are limited in their input accepting text prompts

of only a few lines (e.g., single sentences). In contrast, educational videos require more extensive input such as a video script to ensure completeness and correctness of the content presented. Regarding the perception of AI-generated videos for serious content, the main focus in this area was on fictional characters or personae (Hudson & Hurter, 2016; Pataranutaporn et al., 2021). The impact of human-like avatars on a learning and training situation has - to the best of our knowledge - not been researched in detail. Our results show that there is a kind of familiarization effect, as most participants reported that after some time the uncanniness of the video disappeared. These effects have already been discussed (Cheetham et al., 2014), in particular, the relationship between human-likeliness and pleasantness. From our data, however, this effect can only be assumed but not confirmed, as there were no significant relationships identified. This may be because most participants identified the talking head as rather machine-like. Regarding competence - most of the participants perceive the talking head to be competent. Avatars in AI-generated content are perceived to be competent when the source of the video seems to be trustworthy (Cooke et al., 2024), which is in line with what was mentioned in the open question.

6 Limitations and Future Research

Limitations of the proof-of-concept include the use of a single AI model as well as a single video explaining a relatively simple process which constrain the generalizability of the findings. Future research will need to examine in more detail (e.g., using multiple videos that were generated using different AI tools and presenting content of varying degree of complexity) the effect of AI-based instructors on the emotional response of viewers to investigate how communication of content via AI-generated videos is affected by their uncanniness. Such an evaluation is of particular importance considering that large online-first universities such as the German IU International University of Applied Sciences (more than 100,000 enrolled students) have already implemented AI video production engines for the creation of their educational videos and video lectures (Schiessl, 2024). Moreover, research is needed to investigate, whether the uncanny valley, once experienced, creates a constant feeling of unease, or can be overcome again (e.g., by continued exposure and getting accustomed to the uncanniness of AI-generated content). Moreover, the current study did not analyse the effect of AI-generated, human-like avatars on learning outcomes. Previous studies have primarily focused on the effect of fictional characters used as instructors (Hudson & Hurter, 2016; Pataranutaporn et al., 2021). Thus, these results may not directly translate to instructional videos using human-like avatars. A first study conducted on the effect of virtual, AI-generated, human-like instructors on learning outcomes, showed improved transfer for video lectures using an AI-generated voice and an AI-generated instructor modelled after the human instructor as compared to a recording of a lecture given by the human instructor (Pi et al., 2022). Nevertheless, further research regarding the impact AI-generated educational videos on learning outcomes is needed before considering a large-scale adoption replacing established and proven teaching methods.

References

Bond-Taylor, S., Leach, A., Long, Y., & Willcocks, C. G. (2022). Deep Generative Modelling: A Comparative Review of VAEs, GANs, Normalizing Flows, Energy-Based and Autoregressive Models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(11), 7327–7347.

Cheetham, M., Suter, P., & Jancke, L. (2014). Perceptual discrimination difficulty and familiarity in the Uncanny Valley: More like a “Happy Valley.” *Frontiers in Psychology*, *5*, 1219.

Cho, J., Puspitasari, F. D., Zheng, S., Zheng, J., Lee, L.-H., Kim, T.-H., Hong, C. S., & Zhang, C. (2024). *Sora as an AGI World Model? A Complete Survey on Text-to-Video Generation* (arXiv:2403.05131). arXiv.

Cooke, D., Edwards, A., Barkoff, S., & Kelly, K. (2024). *As Good As A Coin Toss: Human detection of AI-generated images, videos, audio, and audiovisual stimuli*.

Esser, P., Chiu, J., Atighehchian, P., Granskog, J., & Germanidis, A. (2023). Structure and Content-Guided Video Synthesis with Diffusion Models. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, 7312–7322.

Gregory, S. (2023). Fortify the Truth: How to Defend Human Rights in an Age of Deepfakes and Generative AI. *Journal of Human Rights Practice*, *15*(3), 702–714.

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. (2017). Improved training of wasserstein GANs. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 5769–5779.

Guo, P., Kim, J., & Rubin, R. (2014). *How video production affects student engagement: An empirical study of MOOC videos*.

Ho, C.-C., & MacDorman, K. F. (2010). Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Computers in Human Behavior*, *26*(6), 1508–1518.

Ho, J., Chan, W., Saharia, C., Whang, J., Gao, R., Gritsenko, A., Kingma, D. P., Poole, B., Norouzi, M., Fleet, D. J., & Salimans, T. (2022). *Imagen Video: High Definition Video Generation with Diffusion Models* (arXiv:2210.02303). arXiv.

Hong, W., Ding, M., Zheng, W., Liu, X., & Tang, J. (2022). *CogVideo: Large-scale Pretraining for Text-to-Video Generation via Transformers* (arXiv:2205.15868). arXiv.

Hossain, M., Amin, S. Z., & Hossain, F. (2025). Empowering Education Harnessing AI Tools for Innovative Content Development. In *Strengthening Human Relations in Organizations With AI* (pp. 295–326). IGI Global Scientific Publishing.

Hudson, I., & Hurter, J. (2016). Avatar Types Matter: Review of Avatar Literature for Performance Purposes. In S. Lackey & R. Shumaker (Eds.), *Virtual, Augmented and Mixed Reality* (pp. 14–21). Springer International Publishing.

Jiang, H. H., Brown, L., Cheng, J., Khan, M., Gupta, A., Workman, D., Hanna, A., Flowers, J., & Gebru, T. (2023). AI Art and its Impact on Artists. *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 363–374.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and Improving the Image Quality of StyleGAN. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8107–8116.

Li, Y., Gan, Z., Shen, Y., Liu, J., Cheng, Y., Wu, Y., Carin, L., Carlson, D., & Gao, J. (2019). *StoryGAN: A Sequential Conditional GAN for Story Visualization*. 6322–6331.

Mildner, P., & Mueller, F. "Floyd." (2016). Design of Serious Games. In *Serious Games: Foundations, Concepts and Practice* (pp. 57–82). Springer.

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, *19*(2), 98–100. IEEE Robotics & Automation Magazine.

Pataranutaporn, P., Danry, V., Leong, J., Punpongsanon, P., Novy, D., Maes, P., & Sra, M. (2021). AI-generated characters for supporting personalized learning and well-being. *Nature Machine Intelligence*, *3*(12), 1013–1022.

Pi, Z., Deng, L., Wang, X., Guo, P., Xu, T., & Zhou, Y. (2022). The influences of a virtual instructor's voice and appearance on learning from video lectures. *Journal of Computer Assisted Learning*, *38*(6), 1703–1713.

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). *High-Resolution Image Synthesis with Latent Diffusion Models*. 10674–10685.

Roy, M., & Raval, M. S. (2023). Unmasking DeepFake Visual Content with Generative AI. *2023 IEEE 11th Region 10 Humanitarian Technology Conference (R10-HTC)*, 169–176.

Runway. (2025). *Gen-3 Alpha FAQs*. Runway. https://help.runwayml.com/hc/en-us/articles/30266495761171-Gen-3-Alpha-FAQs

Sablić, M., Mirosavljević, A., & Škugor, A. (2021). Video-Based Learning (VBL)—Past, Present and Future: An Overview of the Research Published from 2008 to 2019. *Technology, Knowledge and Learning*, *26*(4), 1061–1077.

Schiessl, F. (2024). *How IU teaches 100,000 students with AI generated videos*. https://www.synthesia.io/case-studies/iu-university

Sora. (2025). https://openai.com/sora/

Srinivasan, R., & Uchino, K. (2021). Biases in Generative Art: A Causal Look from the Lens of Art History. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 41–51.

Tome, D., Russell, C., & Agapito, L. (2017). Lifting from the Deep: Convolutional 3D Pose Estimation from a Single Image. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5689–5698.

van den Oord, A., Kalchbrenner, N., & Kavukcuoglu, K. (2016). Pixel recurrent neural networks. *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48*, 1747–1756.

Wu, C., Liang, J., Ji, L., Yang, F., Fang, Y., Jiang, D., & Duan, N. (2022). NÜWA: Visual Synthesis Pre-training for Neural visUal World creAtion. In S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, & T. Hassner (Eds.), *Computer Vision – ECCV 2022* (pp. 720–736). Springer Nature Switzerland.

Yang, L., Zhang, Z., Song, Y., Hong, S., Xu, R., Zhao, Y., Zhang, W., Cui, B., & Yang, M.-H. (2023). Diffusion Models: A Comprehensive Survey of Methods and Applications. *ACM Comput. Surv.*, *56*(4), 105:1-105:39.

Yu, R., Russell, C., Campbell, N. D. F., & Agapito, L. (2015). Direct, Dense, and Deformable: Template-Based Non-rigid 3D Reconstruction from RGB Video. *2015 IEEE International Conference on Computer Vision (ICCV)*, 918–926.