Proceedings of 2024 International Conference on Information Resources Management

26-28 May 2024

Conference theme:

Humanizing the Digital Age

Hosted by:

The American University in Cairo

Editors:
Lech Janczewski
Galal H. Galal-Edeen
Barbara Krumay

ISBN 978-0-473-71035-4 ISSN: 2744-6220

Blank page

Table of Content

Blank page

1. Welcome Message from Conf-IRM 2024 Conference Co-Chairs

Dear conference attendees,

Hello—(اهلاً وسهلاً)

We are thrilled that you will join us at the 2024 International Conference on Information Resources Management (Conf-IRM). Since its inaugural edition in 2008, the conference series has traveled to several countries across the globe, including Canada, the United Arab Emirates, Jamaica, South Korea, Austria, Brazil, Vietnam, South Africa, Chile, China, and New Zealand. This year's edition takes place in the vibrant city of Cairo (Egypt), the city founded in 969 AD and home to 20+ million people, making it one of the largest cities in the world.

This year marks the 16th edition of Conf-IRM and will be covering one of the most talked about subjects—artificial intelligence. The theme is "Humanizing the Digital Age." There has been increasing discourse in recent years relating to AI and its impact on individuals, organizations, and society as a whole. As such, the conference will provide further opportunities to engage with this subject and consider the humanization of AI.

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 speakers will delve into the latest trends, from artificial intelligence and cybersecurity to cloud computing and blockchain. 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 business cards, and forge lasting relationships.

Join us in celebrating the spirit of innovation that drives the IT community forward. We look forward to seeing you at the event taking place at The American University in Cairo School of Business. Follow the conference news on social media using #Conf-IRM24 for updates, behind-the-scenes glimpses, and snapshots from different sessions.

Let us together make this conference an unforgettable experience! See you in Cairo soon.

Sherif Kamel

Felix Tan

2. Welcome Message from Conf-IRM 2024 Conference Program Co-Chairs

Welcome to Egypt and the 2024 International Conference on Information Resources Management (Conf-IRM). It is an honor and pleasure to chair this year's conference program. Again, we are going back to being fully on-site at the wonderful American University of Cairo. It is a well-established tradition of the conference to particularly involve participants from the area, and this also shows in the contributions and tracks of Conf-IRM 2024.

This year, the conference topic "Humanizing the Digital Age" has attracted many interesting submissions; however, artificial intelligence (AI) has also found a lot of attention lately, which is reflected in the conference program. Thus, the three conference days feature a broad range of topics, covering different topics in the area of information management. Conf-IRM 2024 received submissions from 21 countries in response to our Call for Papers. Each submission was subject to a rigorous double-blind review process that involved evaluation by at least two reviewers, assessment by the track chairs, and quality assurance of the revised submission, leading to 27 papers, 19 of which are full papers. To allow all participants to be heard, we again managed to have a single-stream conference. Besides papers covering the conference topic and many papers investigating AI in different ways, topics like security and privacy or information system strategy – to name just two - have also been covered. That Conf-IRM always wants to encourage and discuss new research is reflected by the highly relevant research-in-progress papers that have been submitted and accepted for presentation.

Conf-IRM 2024 includes a panel on "Understanding and Mitigating the Dark Side of Information Technology Use" and a workshop on "AI, Automation of Work, and the Role of Trust and Authentic Leadership During Times of Digital Transformation." On day three, Conf-IRM 2024 again conducts a Doctoral Consortium for students that have submitted their research beforehand, but also open to Doctoral students from local universities. Also, on day three, the conference integrates a local IT Symposium open to all participants to evoke vivid discussion among practitioners and researchers. Finally, outstanding keynote speakers will contribute to the conference: Associate Professor Jonathan Liebenau, University of London, and to honor the 30th anniversary of AIS, Professor Claudia Loebbecke, University of Cologne.

On behalf of the Conf-IRM 2024 Conference Committee, we would like to thank all our authors for their submissions. We would also like to extend our thanks to the track chairs, reviewers, Doctoral consortium mentors, and keynote speakers for their invaluable contributions to the conference's success. Finally, we expect and trust that you will enjoy Conf-IRM 2024, and we are looking forward to meeting you in Cairo.

Galal H. Galal-Edeen

Barbara Krumay

3. Conference Committees

Conference Co-Chairs

Felix B Tan
Auckland University of Technology
New Zealand

Sherif Kamel
The American University in Cairo
Egypt

Programme Co-Chairs

Galal Galal Edeen
The American University in Cairo
Egypt

Barbara Krumay
Johannes Kepler University Linz
Austria

Publications and Proceedings Chair

Lech Janczewski
University of Auckland
New Zealand

Executive Committee and International Chairs

Gerald Grant
Carleton University
Canada

G. 'Hari' Harindranath
Royal Holloway University of London
United Kingdom

Jairo Gutierrez
Auckland University of Technology
New Zealand

4. Keynote Speakers

Dr Jonathan Liebenau, Associate Professor (Reader) of Technology Management, Department of Management, London School of Economics and Political Science, University of London, UK

Jonathan Liebenau is an Associate Professor (Reader) in Technology Management. He specialises in two areas: fundamental concepts of information, and the problems and prospects of information and communication technology in economic development. He has previously worked in academic administration, technology policy, and the economic history of science-based industry, all positions in which he has emphasised the use of information in organisations. He is the author or editor of several books and over 70 other major publications and has provided consultancy services to leading companies and strategic government agencies, including: Dell, BT, IBM, Microsoft, Tata Consultancy Services, Nortel, EDS, Lloyd Thompson; and in the UK Government, the Office of Science and Innovation, the Department of Trade and Industry and the Home Office. He also has worked with the Royal Society for the Arts and DigitalEurope, and is on the advisory boards of Istanbul Bilgi University and American University in Cairo, School of Business.

Dr Jonathan Liebenau leads LSE Tech, a research group at LSE's Department of Management active in the area of Internet and Communications Economics, Policy and Strategy.

Professor Claudia Loebbecke, Department Media and Technology Management University of Cologne, Germany

Claudia Loebbecke holds the Chair of Media and Technology Management at the University of Cologne and serves as a member of the Administrative Board of the Regional Public Broadcaster WDR (WDR Verwaltungsrat). Since 2023, she has been a Distinguished Professor at Woxsen University in Hyderabad, India, a member of the Global Diplomacy Lab (GDL), and since 2015, she has been appointed as a member of the Scientific Advisory Board of the Scientific Institute for Infrastructure and Communication Services (WIK) and WIK Consult (Board Member 2018-2022).

She received a Master's Degree (1990) and a Ph.D. (1995) in Business Administration, both from the University of Cologne, Germany, and an M.B.A. from Indiana University, Bloomington, Indiana, USA (1991). In 2013, she received the Executive Certificate Management and Leadership from the Massachusetts Institute of Technology, and in 2018, she successfully completed MIT's Executive Education Course Artificial Intelligence: Implications for Business Strategy. Also, since 2018, she has held the title Certified Board Member (Zertifizierter Aufsichtsrat) awarded by Steinbeis Business Academy in Berlin.

In 2005-2006, she was elected President of the global Association for Information Systems (AIS), which named her AIS Fellow in 2012, and a Distinguished Member cum laude in 2019. Since 1991, she has been a member of the Honor Society Beta Gamma Sigma.

5. Panel and Workshop

Panel on Understanding and Mitigating the Dark Side of Information Technology Use

- **G. Harindranath**, School of Business and Management, Royal Holloway, University of London
- **Khaled Hassanein**, DeGroote School of Business, McMaster University, Canada
- **Milena Head**, DeGroote School of Business, McMaster University, Canada
- **Ravishankar Sharma**, Zayed University, United Arab Emirates

Digital communication technologies have great potential to enrich and transform our personal, professional, and educational lives. As we are increasingly connected through technology and the Internet, we can experience great benefits through instant access to all kinds of information and connection to anyone, anytime from anywhere. However, these benefits or bright sides come with dark sides where IT can be misused, become a threat, or cause harm. The rapid emergence of AI has brought these challenges to the fore with ethical as well as regulatory considerations around what constitutes "responsible" use. This panel discussion will present some examples of the dark sides of IT use with a focus on understanding how to mitigate their negative effects.

A Workshop on AI, Automation of Work, and the Role of Trust and Authentic Leadership During Times of Digital Transformation

- **Rami Al-Sharif**, University of Glasgow, United Kingdom
- Aleksandra **Bavdaz**, University of Glasgow, United Kingdom
- **Novika Grasiaswaty Kamal**, University of Glasgow, United Kingdom

The workshop will involve current trends for AI, insights from academic research, activities, and a case study. Further, the workshop will give emphasis on hands-on activities and discussions while using examples and excerpts from real business situations, where trust is achieved through authentic leading. Imagining an AI-based workplace, workshop participants will tentatively step in the shoes of both a transformational leader and an uncertain employee in the era of transformation. The target audience includes academics, practitioners, business owners/managers, and students.

6. Tracks and Track Chairs

Track 01 - Humanizing the Digital Age – Technology, Sustainability, and Development
Malcolm Garbutt, University of Western Cape, South Africa
G. Harindranath, Royal Holloway, University of London, United Kingdom

Track 02 - Artificial Intelligence Applications
Ahmed Doha, School of Business at Carleton University, Canada
Gunjan Mansingh, The University of the West Indies, Jamaica
Cesar Alexandre Souza, University of São Paulo, Brazil

Track 03 - Digital Transformation, Entrepreneurship, and Innovation
Maria Alexandra Cunha, Fundação Getulio Vargas, Brazil
Manuel Muehlburger, Johannes Kepler University Linz, Austria

Track 04 - Digital Business Platforms, Blockchain, Social Networking, and the IoT
Eusebio Scornavacca, University of Baltimore, USA
Alex Sims, The University of Auckland, New Zealand

Track 05 - Cybersecurity Applications / Cybersecurity Issues and Architectures
Lynn Futcher, Nelson Mandela University, South Africa
Lech Janczewski, University of Auckland, New Zealand

Track 06 - Digital Application in Health, Education, and Government
Khaled Emara, Ain Shams University, Egypt
Samaneh Madanian, Auckland University of Technology, New Zealand

Track 07 - Designing and Implementing Digital Systems – an Industry Perspective
Hanan Moussa, The American University in Cairo, Egypt
Hisham Arafat Shehata, Nile University and Innovay, Egypt

Track 08 - Digitalization in the Middle East and Africa
Kamarul Faizal Hashim, University of Dubai, United Arab Emirates
Gamel Wiredu, Ghana Institute of Management and Public Administration (GIMPA), Ghana

Track 09 - Workshops, Tutorials, and Panels
Nehal Afifi, NOVA University Cairo, Egypt
Jairo A Gutierrez, Auckland University of Technology, New Zealand
Ghada Refaat, Future University, Egypt

Track 10 - Doctoral Consortium
Sherif Aly, The American University in Cairo, Egypt
Anol Bhattacherjee, University of South Florida, USA

7. Conference Reviewers

- Zahidah Abdulkadir, Higher Colleges of Technology, United Arab Emirates
- Aakash Alurkar, Zoom, USA
- Kofi Arhin, Ghana Institute of Management and Public Administration, Ghana
- Nana Assyne, Ghana Institute of Management and Public Administration, Ghana
- Evronia Azer, Coventry University, United Kingdom
- Abhijeet Bhattacharya, SSENSE, Canada
- Anol Bhattacherjee, University of South Florida, USA
- Alton Bodley, University of West Indies, Jamaica
- Luis Hernan Contreras Pinochet, University of São Paulo, Brazil
- Adele Da Veiga, UNISA, South Africa
- Nuraan Davids-Latief, University of the Western Cape, South Africa
- Ahmed Doha, School of Business at Carleton University, Canada
- Khaled Emara, Ain Shams University, Egypt
- Lynn Futcher, Nelson Mandela University, South Africa
- Galal. H. Galal-Edeen, The American University in Cairo, Egypt
- Gustavo Galegale, Universidade de São Paulo, Brazil
- Malcolm Garbutt, University of Western Cape, South Africa
- Sinfree Gono, Royal Holloway, University of London, United Kingdom
- Iris Groher, Johannes Kepler University Linz, Austria
- Jairo A Gutierrez, Auckland University of Technology, New Zealand
- Waleed Khalid Hadban, Charmo University, Iraq
- Najmeh Hafezieh, Royal Holloway University of London, United Kingdom
- G. Harindranath, Royal Holloway, University of London, United Kingdom
- Kamarul Faizal Hashim, University of Dubai, United Arab Emirates
- Daniel Hasler, Johannes Kepler University Linz, Austria
- Anna Hermes, University of Applied Sciences Wr. Neustadt, Austria
- Andreas Hutterer, Johannes Kepler University Linz, Austria
- Lech Janczewski, University of Auckland, New Zealand
- Ines Janusch, Johannes Kepler University Linz, Austria
- Ayanda Kondlo, University of the Western Cape, South Africa
- Barbara Krumay, Johannes Kepler University, Austria
- Yuanyuan Lai, Royal Holloway University of London, United Kingdom
- Maria R Lorini, Royal Holloway University of London, United Kingdom
- Samaneh Madanian, Auckland University of Technology, New Zealand
- Sive Mange, University of the Western Cape, South Africa
- Gunjan Mansingh, The University of the West Indies, Jamaica
- Thembinkosi Maphosa, University of the Western Cape, South Africa
- Ntswaki Matlala, University of the Western Cape, South Africa
- Mziwoxolo Mayedwa, University of the Western Cape, South Africa
- Sami Miniaoui, University of Dubai, United Arab Emirates

- Maryam Mirzai, AUT, New Zealand
- Nicky Moster, Nelson Mandela University, South Africa
- Valeria Feitosa Moura, University of São Paulo, Brazil
- Sudad Muammar, University of Dubai, United Arab Emirates
- Manuel Muehlburger, Johannes Kepler University Linz, Austria
- Avuyile Naki, University of the Western Cape, South Africa
- Stavros Polykarpou, Royal Holloway University of London, United Kingdom
- Ammar Rashid, Ajman University, United Arab Emirates
- Ghada Refaat, Future University, Egypt
- Alexander Reppel, Royal Holloway University of London, United Kingdom
- Lumka Salamnu, Rhodes University, South Africa
- Lukas Schnell, Johannes Kepler University Linz, Austria
- Chuene Semono, University of the Western Cape, South Africa
- Alex Sims, The University of Auckland, New Zealand
- Cesar Alexandre Souza, University of São Paulo, Brazil
- Felix B Tan, Auckland University of Technology, New Zealand
- Lakshminarasimhan Vedanthachari, Royal Holloway University of London, United Kingdom
- Kamala Venigandla, Osmania University, USA
- Sarel J Viljoen, University of the Western Cape, South Africa
- Gamel Wiredu, Ghana Institute of Management and Public Administration (GIMPA), Ghana
- Ruchen Wyngaard, University of Western Cape, South Africa
- Tina Yazdizadeh, Carleton University and Ciena, Canada

8. Best Paper Award Nominees - Full Papers

The Short, Medium, and Long-Term Effects of ICT Use on Corruption: A Panel Data Model (Winner)
Anol Bhattacherjee, Avijit Senguptak, and Sumantra Sarkar

An Exploration of Online Instructor Performance Prediction Model Using LMS Data with Machine Learning Techniques
Gary Yu Zhao, Cindy Zhiling Tu, Yufei Yuan

Research-in-Progress Papers

The Snowball Effect of Multiple Value Co-Creation (VCC) Interactions Leading to Value Co-Destruction (VCD)
Mohammed A Ateia, and Nermeen Fawzy

The Impact of Seeing Human Involvement in Sensor-Based Journalism on Reader Appreciation (Winner)
Irina Boboschko

AI Tool for Cleaning Up Lewd Music: A Jamaican Perspective
Matthew B Stone, and Gunjan Mansingh

9. Doctoral Consortium

Participants and Topics

- **Nur Sena Tanriverdi** - Development of a Framework for AI Auditing Focusing on Challenges.
- **Handan Ercan** - The Effectiveness of Artificial Intelligence-Enabled Advertisements: A Regulatory Focus Perspective.
- **Busra Sebin** - Leveraging Large Language Models for Enhanced Classification in CRM Data.

Mentors

- **G. "Hari" Harindranath**, Royal Holloway College, UK, G.Harindranath@rhul.ac.uk
- **Milena Head**, McMaster University, Canada, headm@mcmaster.ca
- **Khaled Hassanein**, McMaster University, Canada, hassank@mcmaster.ca
- **Ravishankar Sharma**, Zayed University, UAE, Ravishankar.Sharma@zu.ac.ae

10. IT Symposium: Business Change in the Age of AI

Moderated by Galal Galal-Edeen, School of Business, The American University in Cairo and Mostafa Abou Gamrah, MediaSci

Overview

Artificial intelligence (AI) is rapidly transforming the business landscape. From automating tasks to personalizing customer experiences, AI offers a wealth of opportunities to gain a competitive edge and to serve humankind. Although ideas about "Artificial Intelligence" and research into it have been in existence as far back as the early 40s, the world has recently experienced an exceptional surge in interest in AI, fuelled by the stunning experiences with Generative AI systems such as ChatGPT and Google's Gemini. Developments of other competing systems and technologies are happening at an unprecedented speed, with research labs in academia and industry spawning new developments on an almost weekly basis. Since information and communication technologies, which we consider AI to be part of, have been vital to modern day businesses and organisations, we need to think about ways in which collaboration among academia, business, government, and others, can lead to fruitful, safe, and ethical ways of leveraging such developments.

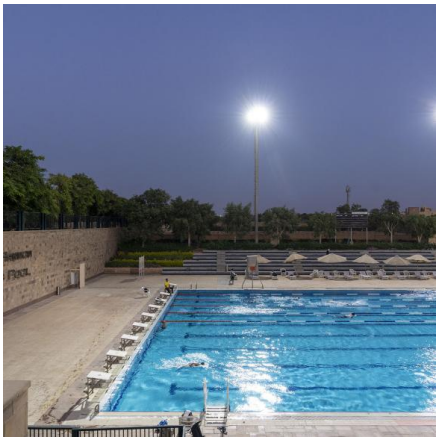
Our symposium brings together industry leaders and experts to explore the exciting possibilities of AI. It is proposed to discuss:

1. Business AI adoption use cases and challenges in adopting AI in various businesses and industries.
2. What Industry expect from Academia, and vice-versa?
3. What collaboration modalities can enhance the dialogue and support progress in AI exploitation?
4. How can the technical issues of Gen AI types (e.g. tailored vs off-the shelf, data acquisition and sources) be approached?
5. Strategies for overcoming challenges and ethical considerations.
6. The future of AI and its potentially disruptive impact on business models.
7. What skills should the next generation of leaders have to be able to steer businesses and organisation into a favourable state through AI exploitation?

IT Symposium Participants

- Ahmed Bahgat, Head of Technology Innovation, Ministry of Justice.
- Ahmed Salama, Solutions Director and CTO, Microsoft Egypt.
- Ayman Ismail, Abdul Latif Jameel Endowed Chair of Entrepreneurship and Director, The Venture Lab, The American University in Cairo.
- Hanan Moussa, The American University in Cairo, School of Business.
- Heba Attia, Senior Director Business Intelligence & Data Analysis, The American University in Cairo

- Hisham Arafat, Digital Transformation Consulting Director, EMEA, Innovay and Associate Professor of Practice, Nile University.
- Maysa Bedair, Corporate Sustainability Lead, PwC.
- Mohab ElHoseiny, Product Manager AI and Messaging, Vodafone
- Mohamed Abaza, Head of Growth, Synapse Analytics.
- Mohamed Khattab, CEO, Mylo.
- Nihad Shelbaya, President and Founder, Nada Foundation for Safer Egyptian Roads
- Radwan Hassan, Director of Medical Informatics Unit, Kasr El-Einy Hospitals and Strategic Planning Officer, Faculty of Medicine, Cairo University.
- Ravishankar Sharma, Professor of Information Systems & Technology and Management, Zayed University.
- Rehab Abdel Hai, Professor of Public Health & Community Manager,Cairo University
- Sherif Aly, Chair, Department of Computer Science, The American University in Cairo.
- Sherif Kamel, Professor and Dean, AUC School of Business
- Wafaa Mohey, General Manager, SaudiSoft Cairo.
- Youssef Youssef, Data Analytics Manager, ITWorx.



11. Conference Program

Sunday 26 May 2024

09:00 - 09:30 **Registration**

09:30 - 09:40 **Welcome Remarks**
Sherif Kamel, The American University in Cairo, Egypt

09:40 – 10:30 **Keynote: There Is No Need to Humanize the Digital Age**
Jonathan Liebenau, London School of Economics, United Kingdom

10:30 - 12:00 **Session 1**
Chair: Gerald Grant, Carlton University, Canada
1. A Systematic Literature Review for Artificial Intelligence in Advertising
Handan Derya Ercan, Bogazici University, Turkey
Nur Sena Tanriverdi, Bogazici University, Turkey
Nazim Taskin, Bogazici University, Turkey
2. Artificial Intelligence Meets Seafood Supply Chain Management
Behrang Ijadi, University of Canterbury, New Zealand
Stephen Wingreen, University of Canterbury, New Zealand
Pedro Reyes, Baylor University
8. Demystifying the Maze: An Investigation of Barriers to AI Marketing Adoption in SMEs
Mariya Rai, University of Bedfordshire, UK
Yanqing Duan, University of Bedfordshire, UK
Annie Danbury, University of Bedfordshire, UK
10. Exploring the Intersection of Large Language Models (LLMs) and Explainable AI (XAI): A Systematic Literature Review (Research-in-Progress)
Busra Sebin, Bogazici University, Turkey
Nazım Taskın, Bogazici University, Turkey
Nijat Mehdiyev, Bogazici University, Turkey
3. AI Tool For Cleaning Up Lewd Music: A Jamaican Perspective (Research-in-Progress)
Matthew B Stone, University of the West Indies, Jamaica
Gunjan Mansingh, University of the West Indies, Jamaica

13:10 - 13:55 **AIS 30 Years Anniversary**
Claudia Loebbecke, University of Cologne, Germany

13:55 - 14:55 **Session 2**
Chair: G "Hari" Harindranath, Royal Holloway, University of London
7. Consumers' View and Discourse on ChatGPT through the Lens of YouTube Comments
Omar H Fares, Toronto Metropolitan University, Canada
Seung Hwan Mark Lee, Toronto Metropolitan University, Canada
Frances Gunn, Toronto Metropolitan University, Canada
Joseph Aversa, Toronto Metropolitan University, Canada
11. Investigating Higher-Education Students' ICT Skills Learning Experiences: Extending the Technology
Charlette Donalds, University of the West Indies, Jamaica
Kweku-Muata Osei-Bryson, Virginia Commonwealth University, USA

20. The Short, Medium, and Long-Term Effects of ICT Use on Corruption: A Panel Data Model
Anol Bhattacherjee, University of South Florida, USA
Avijit Senguptak, University of Queensland, Australia
Sumantra Sarkar, State University of New York, Binghampton, NY, USA

16:00 - 17:10 **Session 3**
Chair: Malcolm Garbutt, University of Cape Town, South Africa
27. The Role of Information Systems Strategy in Driving Successful Digitalization: An IoT Perspective
Hany Abdelghaffar, German University in Cairo, Egypt
Mona A Ali, German University in Cairo, Egypt
17. The Impact of Seeing Human Involvement in Sensor-Based Journalism on Reader Appreciation (Research-in-Progress)
Irina Boboschko, University of Cologne, Germany
26. The Snowball Effect of Multiple Value Co-Creation (VCC) Interactions Leading to Value Co-Destruction (VCD) (Research-in-Progress)
Mohammed A Ateia, The German University in Cairo
Nermeen Fawzy, The German University in Cairo
28. A Student-Dominant View of the Readiness to Use Metaverse in Education: The TRI-F Framework
Malcolm Garbutt, University of Western Cape, South Africa
Ilhaam Ismail, University of Western Cape, South Africa
Raeez Adams, University of Western Cape, South Africa
Calvineo Juries, University of Western Cape, South Africa

Monday 27 May 2024

09:30 - 11:00 **Session 4**
Chair: Jairo Gutierrez, Auckland University of Technology, New Zealand
4. An Examination of Industry Privacy Statements in Top New Zealand Websites
Sumedha Mukherjee, Auckland University of Technology, New Zealand
Jairo A Gutierrez, Auckland University of Technology, New Zealand
19. The Role of SOCs for Critical Infrastructure Providers
Michaela K. Trierweiler, Johannes Kepler University Linz, Austria
Barbara Krumay, Johannes Kepler University Linz, Austria
13. IT Executive Dream Teams
Ariel I La Paz, University of Chile, Chile
David Lopez, University of Chile, Chile
14. Sprinting Toward Digitalization: Transition To Industry 4.0 in the Context of Industrial SMEs (Research-in-Progress)
Redouane El-Amrani, Audencia Business School, France
18. The Mediating Role of Government Support in E-wallet Acceptance Intention in Malaysia (Research-in-Progress)
Kamarul Faizal Hashim, University of Dubai, United Arab Emirates
Felix B Tan, Auckland University of Technology, New Zealand
Jairo A Gutierrez, Auckland University of Technology, New Zealand

11:15 - 12:45 **Panel on Understanding and Mitigating the Dark Side of Information Technology Use**
G. Harindranath, Royal Holloway, University of London

Khaled Hassanein, McMaster University, Canada
Milena Head, McMaster University, Canada
Ravishankar Sharma, Zayed University, United Arab Emirates

13:45 - 15:05 **Session 5**
Chair: Barbara Krumay Johannes Kepler University Linz, Austria
9. Effective Churn Prediction in the Online Payment Sector: A Survival Analysis
Gözde Özalpay, Boğaziçi University, Turkey
Nazim Taskin, Bogazici University, Turkey
23. Deep Learning Methods for Hepatic Tumor Detection
Manal Makram, Cairo University, Egypt
Hana Abdel Malek, Cairo University, Egypt
24. Examining the Influence of Explainable Artificial Intelligence on Stakeholder Decision-Making Processes in an Audit Firm
Sinthuri Konar, University of Cape Town, South Africa
Grant Oosterwyk, University of Cape Town, South Africa
Pitso Tsibolane, University of Cape Town, South Africa
25. Optimization of Call Center Agent Resources Using Various Machine Learning Methods: A Systematic Review
Can Yüzüak, Bogazici University, Turkey
Furkan Nurdag,
Mesut Tartuk,

15:15 - 16:35 **Session 6**
Chair: Myriam Raymond, Institut D'administration Des Entreprises, France
16. Impact of Generative Artificial Intelligence on Employees' Creativity in the Workplace: A Literature Review and Research Agenda
Kyrillous George, The German University in Cairo
Nada Elsehemy, The German University in Cairo
2. Accessing and Navigating Reliable Health-Related Information for African SMEs During the Pandemic Insights from an Empirical Study in 5 African Countries
Myriam Raymond, Institut D'administration Des Entreprises, France
5. An Exploration of Online Instructor Performance Prediction Model Using LMS Data with Machine Learning Techniques
Gary Yu Zhao, Northwest Missouri State University, USA
Cindy Zhiling Tu, Northwest Missouri State University, USA
Yufei Yuan, McMaster University
22. Unrevealing The Digital Thread: Exploring Students' LMS Digital Behavior and Its Impact on Academic Performance in Kuwait Higher Education
Ibtisam L F H Almutairi, Public Authority for Applied Education and Training, Kuwait
Brad McKenna, University of East Anglia, UK
Adrian Benfell, University of East Anglia,UK

16:35 – 16:50 **Closing Remarks**
Sherif Kamel, The American University in Cairo, Egypt

18:15 - 21:00 **Conference Dinner and Best Paper Award Ceremony**
Liberty Boat Yacht

Tuesday 28 May 2024

09:00-09:15 **Doctoral Consortium**

Anol Bhattacherjee, University of South Florida, USA
Sherif Aly, The American University in Cairo, Egypt

09:15 – 11:15 Student Presentations and Feedback

11:15 – 12:15 Panel Discussion with Mentors
How to be successful in the academic profession.

12:15 – 12:30 Concluding remarks

14:00 – 14:15 IT Symposium
Co-Chairs: Galal Galal Edeen, The American University in Cairo, Egypt
Mostafa Abou Gamrah, CEO MediaSci, Egypt

14:15 – 15:45 Roundtable Conversation

15:45 – 16:00 Concluding remarks

12. Conference Papers Index

Notes:

- Number in brackets: page number of the paper text.
- All these papers are also uploaded to https://aisel.aisnet.org/confirm2024/

13. Conference Papers Index

14. Conference Papers

1. A Systematic literature review for Artificial Intelligence in Advertising

Handan Derya Ercan
Bogazici University
handan.ercan@std.bogazici.edu.tr

Nur Sena Tanrıverdi
Bogazici University
nur.tanriverdi@std.bogazici.edu.tr

Nazım Taşkın
Bogazici University
nazim.taskin@bogazici.edu.tr

Abstract

The advertising ecosystem has grown and adapted to the advancements in technology over time. Unprecedented developments in artificial intelligence (AI) technology have huge impacts on advertising. Four main processes of advertising, capturing consumer insights, designing new advertisements, planning and buying media, and measuring the impact and effectiveness have benefited from advents in AI. Hence, it is crucial to acknowledge the influence and application of AI in advertising, as well as keep up-to-date with the latest trends and issues in this field. To this end, a thorough systematic literature review (SLR) was conducted to explore studies published between 2005 and 2023 that delve into the role and evolution of AI in advertising. The review specifically targeted programmatic, computational, and intelligent advertising in the literature. The study findings not only highlight current research topics in AI advertising, but also emerging areas such as manipulated advertising, Generative AI, and consumer concerns, including ethical issues.

Keywords: Programmatic advertising, computational advertising, systematic literature review, generative artificial intelligence, deepfake.

1. Introduction

Web 1.0 (1990 - 2005) started the Internet age and web pages have become the new source of information. Since the usage of the first banner in 1994, digital advertising has evolved exponentially (Li, 2019). E-commerce sites open the doors of a new advertising medium. In the web 2.0 era, through global access to the Internet with high-speed communication and the introduction of YouTube and social media networks, user-generated content has become an essential source of information, generating vast amounts of data. Thus, the share of digital advertising increased year by year transformed the entire advertising industry, and gave rise to a new field of research over the past thirty years.

Web 3.0 is associated with technologies that might also influence the advertising industry, such as the use of semantic web (Halpin, 2023), virtual and augmented reality in meta-marketing

engagement (Kotler, Sertijawa, & Setiawan, 2023), blockchains (Joo, Kim, Ghose, & Wilbur, 2023), and unprecedented development of artificial intelligence (AI) (Huh & Malthouse, 2020; Tobaccowala & Jones, 2018). The availability of big data in various forms, advances in computational power and information technology, and the refocus of investments on AI techniques and technological enablers have accelerated AI penetration (Overgoor, Chica, Rand, & Weishampel, 2019; Vlacic, Corbo, eSilva, & Dabic, 2021).

Considering shifts in the advertising ecosystem with the unprecedented developments in AI technology, this research aims to review the literature about artificial intelligence's impact on advertising processes. Accordingly, our research objective is to explore and synthesize the usage of Artificial Intelligence in the advertising ecosystem and explore the latest themes and concerns in AI advertising.

The research questions are as follows:

RQ-1: What are the developments in programmatic advertising, computational advertising, and intelligent advertising?

RQ-2: How does AI technology enable/drive the main processes of advertising?

RQ-3: What are the latest themes in AI advertising?

2. Background Information

2.1. AI in Advertising

Rodgers (2021, p.2) defined artificial intelligence advertising (AI advertising) as "brand communication that uses a range of machine functions that learn to carry out tasks with intent to persuade with input by humans, machines, or both". It is a distinct subdiscipline of advertising that is positioned at the intersection of cognitive science, computer science, and advertising (Rogers, 2021).

AI has benefits in four main processes of advertising: (1) Capturing consumer insights, (2) designing new advertisements, (2) planning and buying media, and finally (4) measuring the impact and effectiveness (Qin & Jiang, 2019). Machine learning (ML) algorithms are used in capturing consumer behavioral and personal data, needs, and wants through pattern detection of users' tracks. Customer insights can be collected and analyzed through other new AI technologies such as computer vision, speech recognition, natural language processing (NLP), and other integrated systems (Rodgers & Nguyen, 2022). Designing new advertisements tailored to each consumer is possible by utilizing artificial intelligence technology. Examples of other uses may include studies such as ML combined with computer vision for detecting text and image mismatch (Ha et al., 2020), deep learning for prediction model of click-through rate for display advertising (Wang & Zhao, 2022). Planning and buying media became automated by programmatic advertising. Thus, it is possible to tailor a vast number of personalized ads in a very short time, efficiently, and with low cost (Chen, Xie, Dong & Wang 2019). Finally, measuring the impact of advertising, both attitudes and behavioral intentions of consumers, click-through, and convergence rates are predicted by machine learning algorithms.

Li (2019) explained the evolutionary stages of digital advertising, which evolved from "interactive" to "programmatic" and finally morphed into "intelligent" advertising. Intelligent advertising is "consumer-centered, data-driven, algorithm-mediated brand communication" (Li, 2019, p.334). Three key elements of AI, which are data, algorithms, and computing power, facilitate intelligent advertising (Li, 2019).

Huh and Malthouse (2020) define computational advertising "as a broad, data-driven advertising approach relying on or facilitated by enhanced computing capabilities, mathematical models/algorithms, and the technology infrastructure to create and deliver messages and monitor/surveil an individual's behaviors" (p. 367). Computational advertising could be aligned with the narrowest or weakest form of AI (Kaplan & Haenlein, 2019), describing models and algorithms that can perform only specific tasks or one task at a time and must be continually directed by humans to be effective.

Programmatic advertising (PA) is part of computational advertising and refers to the automation of advertisement execution (Li, 2019; Lee & Cho, 2020). Advertisers and online ad platforms agree on the price and place of the ads within an automated real-time bidding system (RTB) (Chen et al., 2019). Unlike purchasing ad space in media, advertisers buy only the exposure to the audience who are most likely related to a particular ad. For example, readers of a news portal are not served the same ad but are displayed different banner ads depending on the reader's demography and online behavior (Lee & Cho, 2020).

Programmatic media buying penetrated widely with the advent of machine learning and the utilization of big data. AI has been able to aid in the creation of programmatic advertisements (Chen et al., 2019).

Further advances in artificial intelligence enable synthetic advertisements. Content and design changes can be "autonomously" manipulated by AI algorithms. Deep neural network technologies enable hyper-personalized or even individualized ads by changing voice, facial characteristics, background, and emotional expressions (Campbell, Plangger, Sands, & Kietzmann, 2022a).

3. Research Methodology

This study adopts Systematic Literature Review (SLR) as a methodology to accomplish a comprehensive review of AI advertising literature. Results obtained by SLR are more rigorous than ad-hoc literature review considering the evidence-based guidelines provided for researchers (Kitchenham et al., 2009; Tranfield, Denyer, & Smart, 2003). Moher et al.'s (2010) PRISMA (preferred reporting items for systematic reviews and meta-analysis) protocol, broadly used in the marketing literature (Ford, Jain, Wadhwani, & Gupta, 2023; Ter Huurne, Ronteltap, Corten, & Buskens, 2017; Lim & Rasul, 2022) is applied to this SLR. According to this protocol, identification, screening, eligibility, and reporting are the main steps to follow in a rigorous review process (Moher et al., 2010). Several tasks within the review process, such as creating a review protocol and synthesizing and reporting articles suggested by Tranfield et al.'s (2003) SLR methodology are also applied to this review.

As Tranfield et al. (2003) mentioned, in order to provide an objective review, the review strategy is determined. Thus, the literature review protocol, inclusion, and exclusion criteria are established according to the needs of this SLR. Table 1 shows the review protocol. Inclusion and exclusion criteria are given in the conducting stage.
In the literature review protocol Web of Science (WoS) is determined as the search database because of its coverage and the quality of papers included (Birkle et al., 2020). Journal articles, including early access articles and review articles, are searched in WoS and evaluated for this SLR because of their contributions to the literature (Hult, Reimann, & Schicke, 2009). Moreover, the time period for the review is chosen between 2005 and 2023. The emergence of YouTube and other social media platforms in 2005 marks the starting point of this review. (Wedel & Kannan, 2016). These platforms have led to a massive amount of user-generated data created, accumulated and used for analysis in advertising. This development has opened a new age in advertising toward personalization and online behavioral targeting

Unit of analysis	Journal articles
Type of analysis	Qualitative
Time period	2005 - 2023
Search fields	Title, abstract, keywords
Databases	Web of Science
Total number of articles used in this study	62

Table 1: Review Protocol

Upon defining the literature review protocol, the next step is to execute it, which involves identifying, assessing, extracting, and synthesizing articles (Tranfield et al., 2003). First of all, keywords were identified to find related articles to AI advertising according to the PRISMA protocol. The main aim is to understand the role of AI in advertising mainly through the lens of AI-enabled technologies. Within this area, digital advertisement is an umbrella term (Li, 2019).

At the beginning of this SLR study, a limited number of searches with the “digital adverti*” keyword was conducted in order to understand the relevance of papers including this keyword to the aim of this review. However, this selection provided articles from a broader area, including studies such as social media advertisements, search engines, mobile advertisements, etc., that do not fit the scope and aim of this study. More specified keywords were determined in order to reach AI-specific advertising articles. These are intelligent, programmatic, computational, and AI advertising. Furthermore, keywords like manipulative, synthetic, or smart advertising are also used in a few recent studies; yet they are explained under AI-driven, AI-enabled, or AI-created advertising. Therefore, the latter three keywords are included as they provide a better and broader fit, instead of the former three. In the query, “adverti*” is used to include all possible uses such as “advertisement”, “advertising”, “advertizement”, “advertizing” words in the search. The search query is executed in WoS database with “topic” search rule, which includes “title”, “abstract” and “keywords” of papers. The final search query is formulated with the identified keywords in Table 2.

"programmatic adverti*" (Topic) or **"AI adverti*"** (Topic) or **"computational adverti*"** (Topic) or **"artificial" AND "intelligence" AND "adverti*"** (Topic) or **"intelligent adverti*"** (Topic) or **"AI-enabled" AND "adverti*"** (Topic) or **"AI-driven" AND "adverti*"** (Topic) or **"AI-created" AND "adverti*"** (Topic)

Table 2: Database Search Query

As a result of the execution of the search query, 659 papers were extracted from the database. After a thorough screening process, 265 papers were removed from the list due to their lack of a rigorous review process or because they were still in progress. These papers included book chapters, conference proceedings, retracted publications, editorial materials, and film reviews. Next, 22 non-English articles were excluded. 126 articles published in unrelated disciplines, such as biotechnology, biomedicine, agriculture, and anthropology, were excluded from the sample. These articles were removed, and 248 were selected for abstract reading.

After the screening of the full list of extracted articles, the eligibility of articles was evaluated by reviewing their title, abstracts, and main text. The use of AI in advertising was identified as a key focus for eligibility criteria. The articles included in the study were filtered by their focus on the evolution, importance, usage, and examination of AI advertising. During the screening phase, articles unrelated to advertising or irrelevant to the research scope were excluded, even if the identified keywords were present. For instance, studies solely concentrated on the method of mathematical or algorithmic modeling of AI tools or systems, and studies that AI-enabled advertising was mentioned as secondary, the ones focusing on specific categories or legal issues (e.g., ads targeting children, AI use on elections). According to the eligibility criteria, 182 articles were assessed as irrelevant and removed from the review. An additional two articles were eliminated due to the inaccessibility of their full copies. Figure 1 shows the flowchart created based on PRISMA methodology.

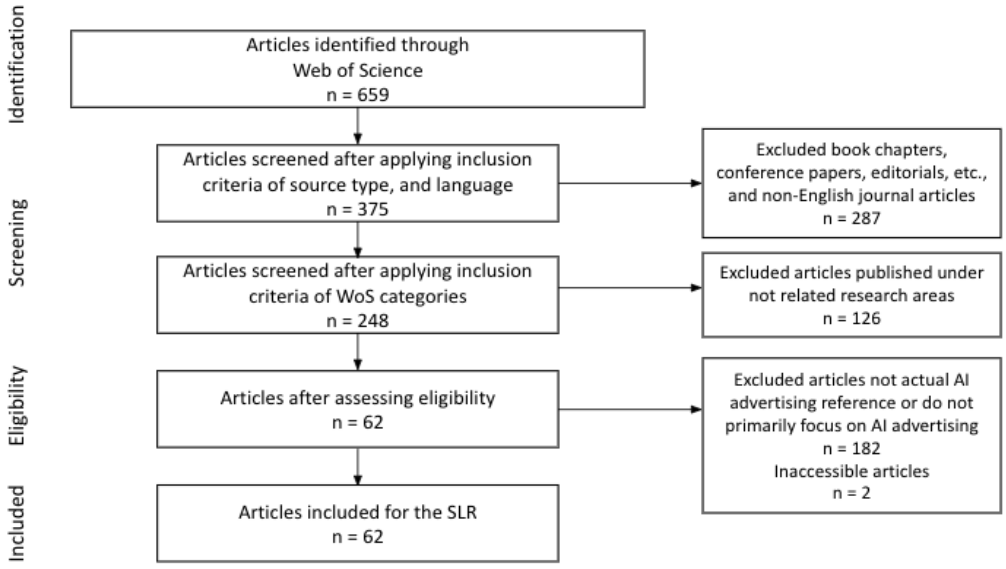


Figure 1: Flowchart of The Literature Review

Descriptive analysis and thematic analysis should be presented according to Tranfield et al. (2003) in the reporting stage. Therefore, the rest of the paper includes reporting on the review.

Descriptive analysis, a mainly simpler set of analyses to provide a broad understanding, is provided in the Results Section. Thematic analysis, which outlines key themes regarding research questions, is given in the Discussion Section.

4. Results

Among 62 articles in the final set[1], the oldest study was published in 2013. From 2019 onwards, there was a significant increase in the number of articles published. By 2023, this number had reached the highest point, with 21 publications. The methodology of reviewed articles is also evaluated. Most articles conducted a conceptual study (27 articles). These articles mainly focus on model or framework development conceptually. Quantitative methods, such as experiments and surveys, and qualitative approaches, such as in-depth interviews, case studies, and literature reviews are mostly used in studies, respectively 18 and 12 times. Four articles use mixed-method approaches. There is also one article that adopts a data-driven approach. Moreover, programmatic advertising is the most studied AI advertising type. The second most studied type is computational advertising (13 articles). At the same time, 14 articles focused on AI advertising from a broad perspective. Manipulated advertising is discussed within five articles, and Generative AI (GenAI) in advertising is the main subject of only two articles till the end of 2023. On the other hand, intelligent AI is the least studied topic in the review list.

5. Discussion

5.1 Computational Advertising

The principles of CA apply to all processes of advertising, from collecting personal data, segmenting and tailoring customers, creating ad messages, placing them in the right channels, and evaluating the results of the campaigns (Huh & Malthouse, 2020; Qin & Jiang, 2019). Customer behavioral data tracking, insight generation, and segmentation increasingly rely on sophisticated machine-learning algorithms. Salminen, Jansen, and Mustak (2023) extracting 172 articles, found researchers employ 46 different algorithms and 14 different evaluation metrics for algorithmic segmentation. Computational advertising is the convergence of marketing, advertising, and information technology (Huh & Malthouse, 2020).

After a multidisciplinary collaboration among scholars to examine advances in CA, five conceptual articles were published in the special issue of Journal of Advertising. Huh, and Malthouse (2020) introduced the editorial with definitions, historical background, key characteristics, and a summary of the following studies. Helberger, Huh, Milne, Strycharz, and Sundaram (2020) provided conceptual development, CA ecosystem, and macro issues, and Liu-Thompkins, Maslowska, Ren, and Kim (2020) delivered the main issues on user-generated content and discussed the roles of users. Van Noort, Himelboim, Martin, and Collinger (2020) proposed a model of automated brand-generated content. It relies on an iterative process in which both consumer and brand data feed optimization of this content. The latter two articles are related to the advertising creation process. The other critical process is selecting the right media where the messages will be conveyed to consumers. Next to these articles, Araujo,

[1] The final list of articles obtained as a result of execution of review protocol have not been included in the References Section. The list of articles will be provided upon request.

Copulsky, Hayes, Kim, and Srivastava (2020) discussed media planning decisions that engage consumers and focused on the centrality of measurement in the digital ecosystem with the implications of CA for interaction planning. An example of an advertising effectiveness process is by Yun et al. (2020), which delved deeper into the CA measurement system and discussed the challenges and future directions of CA.

In this cluster, Wang, Xu, and Lin (2023), using the theoretical model of UTAUT, investigated influencing factors of using behavior for CA. One of the early studies, by Soriano, Au, and Banks (2013), focused on text mining in the context of CA and proposed topic modeling to persuade readers. The authors shared their comments on the limits of text mining and NLP besides the computational burden and difficulty of creating statistical models that accurately reflect the relevant structure. Just a decade after this study, both text mining and NLP technologies developed in unprecedented ways, enabling breakthrough outcomes.

Huang and Liu (2022), relying on their field study results, found that the style of advertisement text (assertive, expressive, directive) in CA affects consumer behavior. Other studies selected in this cluster are a systematic literature review on personalized advertising computational techniques (Viktorados & Tsadiras, 2021), optimization of marketing communication strategies with AI (Wen, Lin, & Guo, 2022), and machine learning techniques used to improve targeted advertising (Choi & Lim, 2020).

The goal of computational advertising is to ensure effective usage of advertising resources through accurate targeting and enhanced ad relevance. Marketing and data analysis experts in firms should work arm in arm and marketers align their strategies according to the results of extracted advertising performance data.

5.2 Programmatic Advertising

Most of the selected articles are clustered in this domain. Chen, Xia, and Wang (2019b) discussed the evolution and position of PA from a broad perspective on digital advertising. Samuel, White, and Thomas (2021) and Diwanji, Lee, and Cortese (2022) explore consumer concerns and the role of AI in programmatic advertising. In another sub-cluster of PA, studies on ethical perspectives in programmatic buying shed light on consumer concerns regarding the usage of their personal and behavioral data (Martinez-Martinez, Aguado, & Boeykens, 2017; Rodgers & Nguyen, 2022). Palos-Sanchez, Saura, and Martin-Velicia (2019) examined how users' privacy concerns are influenced by their perception of the usefulness of programmatic advertising. Programmatic advertising can benefit from neuromarketing by understanding customers' preferences and capturing their attention (Bajaj, Syed, & Singh, 2023).

5.2.1 Programmatic buying

Programmatic buying is the automation of advertisement purchasing “optimizes the effect of correcting and tracking the way the advertisement is executed by exposing the right message to the right people at the right time in the most cost-effective manner through data-based real-time bidding (RTB)” (Lee & Cho, 2020, p.336)

Rodgers and Nguyen (2022) drew the attention of platform developers to the decline of the hype in programmatic advertising due to the real return on investment concerns of advertisers.

Advertisers aim to target consumers who respond positively to advertising. Ad platforms earn higher revenue when they target consumers with a high probability of making a purchase. Frick, Belo, and Telang (2023) put forth that firms and ad platforms agree on which consumers to target critically because showing ads to high purchase probability consumers with higher rates may not be feasible for the company. In a similar vein, Lambrecht and Tucker (2019) found that as targeting younger women requires higher ad prices in the bidding system, yet automation built optimum ad buying plans, younger women might see ads less than men. Shehu, Abou Nabout and Clement (2021) argued website quality influence on advertising effectiveness in the context of search engine advertising.

5.2.2 Programmatic ad creative

Giant e-commerce companies, such as Amazon, can produce millions of ad alternatives for each and every customer for special occasions such as the black Friday shopping season or the Christmas period. This data-driven approach utilizes programmatic ad creative algorithms that add speed, scale, and automation to the creative process. It enables digital advertisers to produce dynamic ads and optimize creative message design and testing.

Chen et al. (2019b) delved into the concept of programmatic creative and expanded programmatic advertisement. The authors also discussed the technological and legal challenges faced by programmatic creative. Current programmatic advertising focuses on performance through informative and rational ad content which is easier to track and measure (Diwanji et al., 2022; Qin & Jiang, 2019).

Consumers' attitude may differ according to the type of advertised product and ad appeal. For example, programmatic creative ads are found to be less effective for emotional appeals and hedonic products but more effective for rational ad appeals and utilitarian products (Bakpayev, Baek, van Esch, & Yoon, 2022). Moreover, consumers' attitudes toward human-created and Al-created advertisements are equal when ads are not emotional but cognitive. When an ad message is more emotional, consumers favor human-created ads over AI-created ones (Bakpayev et al., 2022). Wu and Wen (2021) examined the factors affecting consumers' acceptance of AI-generated ads and found that consumers perceive machines to be more reliable and safer than humans, thus perceiving AI-driven ad creation as more objective. This perception positively influences users' attitude towards AI-generated ads.

5.3 Intelligent Advertising

Li (2019), introduced the concept of “intelligent advertising” as a subsequent stage of programmatic advertising. Only 2 articles use the term “intelligent advertising” in our list, yet, research in AI advertising has incremental development since 2019 (Ford et al., 2023). Rodgers and Nguyen (2022) presented an intelligent advertising theme by examining its present and future as well as identifying the key phases of intelligent advertising.

5.4 Improving Advertising Efficiency

Numerous studies offer AI systems to improve advertising effectiveness. Wang et al. (2019) developed an AI system to produce personalized ads and examined the effect of these intelligent ads. Deng, Tan, Wang, and Pan (2019) presented a smart personalized advertising copy system enabling testing the effect of advertising. Malthouse, Hessary, Vakeel, Burke, and

Fuduric (2019) recommended two-sided consumer and advertiser platforms to personalize products. Vakratsas and Dang (2020) proposed a creative advertising system to produce and test creative ideas. Shumanov, Cooper, and Ewing (2022) conducted a mixed-method study to evaluate advertisement effectiveness of ads based on AI-predicted personality. Those systems are designed to improve the system effectiveness of more cognitive tasks. Recently with the advent of AI technology, AI solutions allow advertisers to assist for affective tasks such as identifying the patterns in the ad viewer's response to a particular video content. Sharakhina, Ilyina, Kaplun, Teor, and Kulibanova (2023) focussed on AI-based face-tracking technologies in the analysis of visual advertising messages. Calibrating the length and attractiveness of ads enables advertising effectiveness with respect to media buying budgets and click-through intentions.

5.5 The Latest Themes in AI Advertising

5.5.1 Manipulative Advertising

Synthetic and manipulated advertising using deepfakes and generative adversarial networks (GANs) are newly investigated concepts in the AI advertising literature (Campbell et al., 2022a; Campbell, Plangger, Sands, Kietzmann, & Bates, 2022b; Kietzmann, Mills, & Plangger, 2021; Sivathanu, Pillai, & Metri, 2023). Synthetic advertising belongs to the larger category of manipulative advertising.

Benefits and risks of manipulated advertisements according to AI advertisement stakeholders (Campbell et al., 2022b), companies (Kietzmann et al., 2021), and consumers (Campbell et al., 2022a; Sivathanu et al., 2023) have also been analyzed. Benefits and risks for companies rise at the same time (Campbell et al., 2022b; Kietzmann et al., 2021). For example, while production costs decrease, costs for digital security and brand protection increase. Additionally, idea creation is no longer limited by barriers such as cost, time and space, and even talent (Kietzmann et al., 2021). However, the abundance of advertisements can hinder creativity (Campbell et al., 2022b).

Tracking and controlling advertising campaigns becomes more challenging with increasing abilities in targeting and personalization of advertisements, leading to the emergence of numerous micro-segments. Moreover, increased abilities require more computing power.

Furthermore, consumers mostly show a positive attitude toward manipulated advertising and increased intentions to purchase (Sivathanu et al., 2023). Besides, serious concerns emerge about privacy and security-related issues of manipulated advertisements (Campbell et al., 2022a).

In the literature, the need for regulations for synthetic content generation and using them in advertising are also emphasized. It is important to maintain the privacy of consumers and the rights of companies concerning their brands, and advertisements (Campbell et al., 2022a; Kietzmann et al., 2021).

5.5.2 Generative Artificial Intelligence in Advertising

Generative artificial intelligence (GenAI) "refers to a class of AI models that generate seemingly new content in the form of text, images, or other media" (Susarla et al., 2023, p.

399). GenAI technologies are expected to have a huge impact on digital advertising (Baek, 2023) and have a transformative impact on AI advertising business and research (Huh, Nelson, & Russell, 2023) with more effective and productive outcomes than previous generations of digital technologies (Kshetri et al., 2023). For example, Generative AI technology might transform search advertising (Huh et al., 2023) and open a new era of personalized recommendations and campaign planning. In this SLR, we have 2 studies related to these subjects.

Remountakis, Kotis, Kourtzis and Tsekouras (2023) explored the capabilities of ChatGPT, and conducted pilot experiments in the context of a hotel recommender system. ChatGPT driven hotel recommender systems can effectively influence user's decisions and encourage them to take targeted actions such as booking a specific hotel or upgrading one's room. This result is obtained by utilizing persuasive techniques. The study highlights the importance of ethical guidelines to prevent the manipulation of ChatGPT users. Gołąb-Andrzejak (2023) investigated advertising campaign creation by utilizing GenAI. The author conducted secondary research and one in-depth interview with industry professionals. Accordingly, the study summarizes the role of GenAI at the stage of the process of creating digital advertising campaigns.

5.5.3 Concerns of Consumers in the age of AI advertising

Intelligent advertisement algorithms generate hyper-personalized ads which might manipulate consumers' choices and influence their autonomy (Coffin, 2022). On the other hand, when consumers find an online ad relevant to their needs and interests, they are more likely to have a positive attitude (Ciuchita et al., 2023). Consumers' perceived usefulness of personalized ads mitigates their privacy concerns (Nunez-Barriopedro et al., 2023).

Higher levels of privacy concerns generate strong feelings of intrusiveness (Zarouali et al., 2017), feelings of vulnerability (Aquirre et al., 2015) and eeriness (Wu & Wen, 2021). Advertising technology companies have been developing alternative solutions for consumer privacy controls in order to mitigate their concerns and reactions. Cooper et al. (2023) conducted a multi-method study of privacy concerns with respect to regulatory legislations. Authors suggest ad platforms using contextual targeting to serve ads tailored to external factors such as the consumer's time and location. Users might still feel tracked but would not be annoyed due to the usage of their personal data. Companies should seek a balance between their business objectives and consumers' interests.

Ethical concerns have emerged during interactions of consumers with brands. A vast amount of studies in the SLR attract attention to ethical perspective in advertising (Arango, Singaraju, & Niininen, 2023; Coffin, 2022; Dwivedi et al., 2021; Gao et al., 2023; Martinez-Martinez et al., 2017; Rodgers & Nguyen, 2022; White & Samuel, 2019) and in GenAI advertising (Remountakis et al., 2023). Besides benefits, using AI-enabled technologies raises ethical challenges such as fairness and bias, inaccuracy, autonomy of users, data sensitivity/privacy, security, and transparency.

6. Conclusion

The main objective of this SLR is to explore and synthesize the usage of AI in the advertising processes and explore the latest themes and concerns in AI advertising. As AI influence in advertising is observed through AI algorithms, models, technology, or specific programs, a variety of keywords were covered to understand AI's contribution to advertising in articles published between 2005 and 2023. The main findings of the study show that programmatic advertising is widely penetrated in the industry and creates timely, cost-effective, hyper-relevant ad delivery to the most responsive audience which is far more effective compared to traditional ad-slot buying. It is possible that some inefficiencies may arise due to the complexity or unexplainable nature of automation. For example, programmatic targeting of already loyal customers or brand's website visitors for ad exposure may generate exaggerated ad effectiveness rates. Retargeting campaigns may not always lead to the desired results due to various factors, such as ad fraud, bots collecting cookies, and excessive retargeting with the wrong strategy (e.g., A user was shown a sneaker advertisement several days after they had completed the purchasing process.). Consequently, advertisers and agencies collaboratively set audience selection strategies to determine the most efficient group of users to achieve higher budget efficiency. Inhouse digital marketing teams consisting of data engineers and marketers need to work very closely. Integrating blockchain systems into the advertising ecosystem might enhance transparency and build trust between the supply and demand side.

Main studies in the "Intelligent advertising" category support cognitive tasks in developing efficient AI-enabled advertising systems or theoretical models. Developments in AI advertising are fast and irrevocable. Hyper-personalized advertising, besides benefits, includes costs, such as consumer privacy concerns and unwillingness to disclose personal information. Online behavioral ads, following users' actions on the web through machine learning algorithms, may evoke strong feelings of intrusiveness, feelings of vulnerability, and eeriness. Transparency about the actions taken would support mitigating those concerns. Further multidisciplinary studies with the latest models would add value to the AI advertising literature and practitioners.

In this study, emerging topics of synthetic advertising, GenAI, and consumer concerns are briefly discussed as they all have been identified as critical and important themes recently. Research on manipulative ads and GenAI advertising falls short of the hype in practice. AI-powered content generation has become increasingly popular in creative fields, such as advertising, media, content, and campaign management starting from November 2022, with the publicity of ChatGPT. Conversational chatbots and virtual assistants resemble humans much more than ever. As AI-generated image and video content reality develops, image-generator tools will have an increasing share in emotional tasks.

Further exploration and future research are needed due to the potential impact of manipulative synthetic ads and GenAI on consumers. If AI chatbots become an advertising platform, are there any risks involved for existing ad platforms and search engine advertising? Shall consumers trust recommendations of conversational chatbots more than brands' own and paid media communication? If the answer is yes, it has the potential to disrupt the advertising ecosystem. Moreover, how will user-generated content, such as product and service claims or brand ambassadors' praises shared on the web influence the "fuel" of GenAI? Uploaded data to the system might not only influence the advertisements on search engines but also the

effectiveness of brand-generated ads. Therefore, the contribution of studies that focus on the data management process within the AI life cycle might be a future research opportunity and add value to the advertising ecosystem.

The ethical perspective of AI is an emerging topic. Additionally, upcoming GenAI hype in advertising might transform players, their functions, and interlinks within the advertising ecosystem. Brands may face issues with consumer autonomy when chatbots like ChatGPT affect consumer preferences.

In further studies, the effects of AI on each advertisement process might be further analyzed from economic, cultural, social, and ethical perspectives within a multidisciplinary approach. The advertising industry might face an accelerated paradigm shift. Regulations would be critical to provide guidance in conflicting conditions.

References

Baek, T. H. (2023). Digital advertising in the age of generative AI. Journal of Current Issues & Research in Advertising, 44(3), 249-251.

Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, *1*(1), 363-376.

Campbell, C., Plangger, K., Sands, S., & Kietzmann, J. (2022a). Preparing for an era of deepfakes and AI-generated ads: A framework for understanding responses to manipulated advertising. *Journal of Advertising*, *51*(1), 22-38.

Campbell, C., Plangger, K., Sands, S., & Kietzmann, J. (2022). Preparing for an era of deepfakes and AI-generated ads:A framework for understanding responses to manipulated advertising. Journal of Advertising,51(1),22-38.

Campbell, C., Plangger, K., Sands, S., Kietzmann, J., & Bates, K. (2022). How deepfakes and artificial intelligence could reshape the advertising industry: The coming reality of AI fakes and their potential impact on consumer behavior. Journal of Advertising Research, 62(3), 241-251.

Chen, G., Xie, P., Dong, J., & Wang, T. (2019). Understanding programmatic creative: The role of AI. Journal of Advertising, 48(4), 347-355.

Ford, J., Jain, V., Wadhwani, K., & Gupta, D. G. (2023). AI advertising: An overview and guidelines. Journal of Business Research, 166, 114124.

Gao, B., Wang, Y., Xie, H., Hu, Y., & Hu, Y. (2023). Artificial Intelligence in Advertising: Advancements, Challenges, and Ethical Considerations in Targeting, Personalization, Content Creation, and Ad Optimization. SAGE Open, 13(4), https://doi.org/10.1177/21582440231210759

Halpin, H. (2023). The Hidden History of the Like Button: From Decentralized Data to Semantic Enclosure. Social Media+ Society, 9(3), 20563051231195542.

Huh, J., & Malthouse, E. C. (2020). Advancing computational advertising: Conceptualization of the field and future directions. Journal of Advertising, 49(4), 367-376.

Huh, J., Nelson M.R., & Russell, C.A., (2023) ChatGPT, AI Advertising, and Advertising Research and Education, Journal of Advertising, 52:4, 477-482, DOI: 10.1080/00913367.2023.2227013

Huh, S. (2023). Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: a descriptive study. Journal of educational evaluation for health professions, 20.

Hult, G. T. M., Reimann, M., & Schilke, O. (2009). Worldwide faculty perceptions of marketing journals: rankings, trends, comparisons, and segmentations. Global Edge Business Review, 3(3), 1-23.

Joo, M., Kim, S. H., Ghose, A., & Wilbur, K. C. (2023). Designing Distributed Ledger technologies, like Blockchain, for advertising markets. International Journal of Research in Marketing, 40(1), 12-21.

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Business horizons, 62(1), 15-25.

Kitchenham, B., Pearl Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering - A systematic literature review. Information and Software Technology, 51(1), 7–15. https://doi.org/10.1016/j.infsof.2008.09.009

Kotler, P., Kartajaya, H., & Setiawan, I. (2023). Marketing 6.0: The Future Is Immersive. John Wiley & Sons.

Lee, H., & Cho, C. H. (2020). Digital advertising: present and future prospects. International Journal of Advertising, 39(3), 332-341.

Li, H. (2019). Special section introduction: Artificial intelligence and advertising. Journal of advertising, 48(4), 333-337.

Lim, W. M., & Rasul, T. (2022). Customer engagement and social media: Revisiting the past to inform the future. Journal of Business Research, 148, 325–342.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2010). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. International journal of surgery, 8(5), 336-341.

Overgoor, G., Chica, M., Rand, W., & Weishampel, A. (2019). Letting the computers take over: Using AI to solve marketing problems. California Management Review, 61(4), 156-185.

Rodgers, S. (2021). Themed issue introduction: Promises and perils of artificial intelligence and advertising. Journal of Advertising, 50(1), 1-10.

Rodgers, W., & Nguyen, T. (2022). Advertising benefits from ethical artificial intelligence algorithmic purchase decision pathways. Journal of business ethics, 178(4), 1043-1061.

Qin, X., & Jiang, Z. (2019). The impact of AI on the advertising process: The Chinese experience. Journal of Advertising, 48(4), 338-346.

Song, H., Kim, J., Nguyen, T. P., Lee, K. M., & Park, N. (2021). Virtual reality advertising with brand experiences: the effects of media devices, virtual representation of the self, and self-presence. International Journal of Advertising, 40(7), 1096-1114.

Ter Huurne, M., Ronteltap, A., Corten, R., & Buskens, V. (2017). Antecedents of trust in the sharing economy: A systematic review. Journal of Consumer Behaviour, 16(6), 485–498.

Tobaccowala, R., & Jones, V. K. (2018). To Thrive in Today's Marketing Landscape, Embrace Schizophrenia!. Journal of Current Issues & Research in Advertising, 39(3), 266-271.

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. British Journal of Management, 14(3), 207–222. https://doi.org/10.1111/1467-8551.00375

Vlačić, B., Corbo, L., e Silva, S. C., & Dabić, M. (2021). The evolving role of artificial intelligence in marketing: A review and research agenda. Journal of Business Research, 128, 187-203.

Wang, F., & Zhao, L. (2022). A Hybrid Model for Commercial Brand Marketing Prediction Based on Multiple Features with Image Processing. Security and Communication Networks, 2022.

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. Journal of marketing, 80(6), 97-121.

2. Accessing and navigating reliable health-related information for African SMEs during the pandemic Insights from an empirical study in 5 African countries

Raymond Myriam
IAE Angers – Université d'Angers - GRANEM
myriam.raymond@univ-angers.fr

Abstract

The COVID-19 pandemic posed formidable challenges for small and medium-sized enterprises (SMEs) globally, with distinct impacts on those in Africa. This empirical study, conducted across five African countries, delves into the critical question of how African SMEs perceived, navigated, and implemented health-related information amidst the complexities of the pandemic. The study reveals multifaceted challenges faced by SMEs, emphasizing the importance of reliable information for their growth and resilience. Drawing from a diverse set of survey respondents, the research explores SMEs' responses to public measures, the clarity of messages, and the implications for public policy. Findings indicate that while technology plays a crucial role, challenges persist in accessing business information. The study also evaluates SMEs' perceptions of government communication, their responses to pandemic measures, and their trust in information sources. The nuanced insights contribute to discussions on effective governance, communication strategies, and technology use, providing valuable considerations for supporting SMEs in navigating crises. Despite limitations stemming from aggregated responses across diverse regions, this study offers key empirical evidence specific to COVID-19, shedding light on the intricate challenges faced by African SMEs and underscoring the need for targeted interventions to enhance their resilience.

Keywords: Information Governance, Access to Information, Covid-19 pandemic, Health-Information Reliability, African SMEs.

1. Introduction

The COVID-19 pandemic has presented unprecedented challenges to the global economy, with significant implications for small and medium-sized enterprises (SMEs) worldwide, including those in Africa. SMEs have faced a myriad of challenges during this crisis, including reduced sales turnover, increased bankruptcy rates, and substantial revenue declines (Sudarmiatin et al., 2023; Van Tien, 2022). Statistical analyses have demonstrated the marked difference in SME performance before and during the pandemic (Deyshappriya & Padmakanthi, 2022).

Amidst these challenges, the role of innovation practices and external support, particularly through information technology and infrastructure, has become increasingly crucial in assisting SMEs to navigate the complexities of the COVID-19 crisis (Adam & Alarifi, 2021). Access to reliable information emerges as paramount for the sustainability and resilience of African SMEs. However, the pandemic has exacerbated the lack of reliable information, underscoring the need for support in managing economic disruptions (Fubah & Moos, 2022).

This paper addresses the fundamental question of how African SMEs have perceived, navigated, and implemented health-related information amidst the challenges of the COVID-19 pandemic. Specifically, we explore their responses to public health measures and restrictions, as well as the clarity of public health messages and dashboards. In doing so, we aim to fill a critical gap in the literature by examining the implications of these dynamics for public policy and future actions.

The subsequent sections will review existing literature on the impact of the COVID-19 pandemic on SMEs in Africa and their interactions with COVID-19-related information. We follow this by a theoretical framing anchored in TMIM (Theory of Motivated Information Management) (Afifi and Weiner's, 2004). Then we explicit the research methodology, followed by the empirical study results and interpretation. We conclude with a discussion highlighting main research contributions and limitations.

2. Literature Review

SMEs encounter challenges in accessing reliable information globally, as evidenced by studies in Spain and the UK (Pedauga et al., 2022; Pavic et al., 2007). Emphasizing the vital role of SMEs in economies, these studies stress the need for reliable information to spur growth and resilience. In Africa, the lack of dependable information impedes access to finance, hinders economic navigation, and impacts competitiveness. The study of Bunyasi et al. (2014) for instance show the influence of business information on SME growth in Kenya.

2.1 SMEs and the COVID-19 pandemic

Louis–Jean et al. (2020) shed light on the detrimental impact of limited access to quality healthcare and resources during the COVID-19 pandemic, especially affecting communities with unreliable health information, like African American communities.

2.2 African SMEs

Fubah & Moos (2022) explored challenges and coping mechanisms for SMEs in the South African entrepreneurial ecosystem during the pandemic, revealing a research gap in addressing COVID-19 challenges in the broader African entrepreneurial ecosystem. Acquaach et al. (2021) contributed a managerial perspective on the pandemic's implications for African SMEs.

2.3 eHealth and Healthcare information

The internet is crucial for healthcare information dissemination, yet challenges persist due to the complexity of the topics and the numerous unfamiliar sources (Chi et al., 2020). This complexity is compounded in the context of COVID-19, with challenges including novelty, lack of expertise, delayed misinformation removal, and algorithmic bias favoring sensationalist content (Fan et al., 2020; Gisondi et al., 2022; Morrow et al., 2022). Seytre (2020) highlighted erroneous COVID-19 communication messages in Africa, emphasizing the importance of adherence to prevention recommendations and insights into health communication challenges during the pandemic.

While valuable insights exist on SME challenges during the pandemic, there's a gap in research relating to the African SMEs' specific need for reliable information, particularly regarding COVID-19 (Kuckertz and Brändle, 2021). Limited research has explored how African SMEs access health-related information during the pandemic, and empirical evidence on their specific health information needs is lacking. Regional studies can offer nuanced insights into how SMEs in different African contexts navigate health information disparities, crucial for informing policy recommendations to improve information dissemination and support SMEs in crises. Addressing these gaps through empirical studies can enhance SME resilience in navigating health crises effectively.

3. Theoretical Framing

Drawing on Afifi and Weiner's (2004) Theory of Motivated Information Management (TMIM), which addresses the relationship between uncertainty and the information seeking process, this study seeks to theoretically frame the management of health-related information seeking by African SMEs during the COVID-19 pandemic. The theory proposes a three-phase process of information management, highlighting the role of efficacy and the interactive nature of information management. In the context of the COVID-19 pandemic, characterized by widespread uncertainty regarding health risks, government measures, and economic impacts, the theory offers insights into how African SMEs may strategically seek, interpret, and utilize information to mitigate uncertainty and make informed decisions. Although not interpersonal in nature, by applying this theoretical framework, the study aims to elucidate the motivations and strategies underlying African SMEs' management of health-related information during times of crisis, aligning with Afifi and Weiner's call for a comprehensive understanding of information management processes in uncertain contexts.

4. Methodology

During the Coronavirus pandemic, The African Union, UNDP, and AfriLabs collaborated to launch the Rollo Africa program for African SMEs. Initiated in August 2020 and concluded in January 2021, the program focused on capacity building and gathering information on SME challenges through surveys. With a goal to offer evidence-based policy guidance for impactful stimulus packages, the Rollo Africa Program selected 120 enterprises with 360 participants across 5 countries. Engaging in 10 online bootcamps, the program addressed diverse challenges during the COVID-19 pandemic, positioning itself as a valuable resource for shaping post-pandemic public policy pathways.

This report centers on data from SMEs in the Rollo Africa Program. The associated research aimed to comprehend evolving challenges faced by African SMEs during the pandemic. Beginning with a literature review identifying 12 specific challenges, the research developed a comprehensive survey tool to assess their topicality and explore SMEs' strategies in addressing them. Administered on three separate days, the survey covered various topics (see Table 1 for survey topics), unveiling persisting challenges and strategies employed by SMEs during the COVID-19 pandemic. The surveys' findings and recommendations intend to guide policy supporting SME businesses in Africa.

In this paper, we focus on results obtained from Surveys 5 and 6 relating to accessing reliable information and particularly those pertaining to health and safety during COVID-19 pandemic.

Bootcamp	Survey	Topics
Day 1	Survey 1	• Access to Finance
	Survey 2	• Electricity & Power. • Internet usage, bandwidth and stability
Day 2	Survey 3	• Cybersecurity and security of digital tools
	Survey 4	• Human resources • Taxes and government duties • Business continuity plans
Day 3	Survey 5	• Lockdown and pandemic control measures • Access to health and safety information & services
	Survey 6	• Negative Perception • Access to reliable information • Government Support

Table 1: Surveys' Topics

4.1 Lack of Adequate Business Information

A primary challenge for African SMEs is the insufficient availability of business information from governmental and service-provider sources (Olawale & Garwe, 2010; Muriithi, 2017). Rooted in underdeveloped technological and communication infrastructures, coupled with inadequate business support systems, this issue impacts SMEs' ability to make informed decisions, especially during events like the COVID-19 pandemic.

4.2. Supplementing Literature with Real-time Challenges

Augmenting existing literature with insights from texts and articles during the pandemic, we aimed to capture recent and pandemic-specific challenges faced by SMEs. This approach provides a comprehensive view of challenges that emerged or intensified due to the unique circumstances of the ongoing global health crisis.

In the upcoming sections, we detail challenges identified through this combined approach, shedding light on their nature and implications. Our objective is to unravel how African SMEs grappled with accessing crucial health-related information during the COVID-19 pandemic and how these challenges influenced their perceptions, navigation, and responses.

4.3 Lack of consistency and clarity in confinement measures

Initially, the lack of consistency and clarity in confinement measures, operational hours, curfews, and criteria for closing and reopening establishments were prominent challenges during the COVID-19 pandemic. Numerous changes introduced by governments stirred controversy, with examples like the selective closure of businesses. Our surveys ultimately assess the impact of these lockdown and confinement measures on African SMEs.

Recent research has formalized the view that the level of rationale clarity, or shared understanding of a decision's basis, affects project performance (Haymaker et al., 2011). While the Rationale Clarity Framework (RCF) was mainly used in architecture projects, it could apply to governments' measures and decisions and aid SMEs in predicting their actions.

4.4 Accessing health and safety information.

Accessing health and safety information and services emerged as a significant challenge during the COVID-19 pandemic. Recommended measures, coupled with region-specific restrictions,

raised questions about information dissemination for African SMEs. Our research aimed to understand what information and health services were available to them and how these resources benefited or hindered SMEs during the outbreak.

4.5 Reliability of health-related information

Concerns arose regarding the reliability of data in some African countries. Factors like limited detection capacity and concerns about testing equipment added complexity to the information landscape. Additionally, attitudes toward COVID-19 best practices and public trust were shaped along partisan lines by the news media (Zhao et al., 2020).

4.6 Survey Builder: Questionnaires developed around different challenges were peer-checked for validity and comprehension by academia colleagues and experts from UNDP and the African Union. Translated into French, the surveys were built using the online tool Alchemer (www.alchemer.com), allowing gathering feedback from respondents in various formats.

5. Empirical Study Results and interpretation

Table 2 details the respondents' numbers across the 5 countries. It also presents how many completed responses per survey were collected. The data used for analysis is the data collected from completed surveys 5 and 6.

5.1 Access to information and the information infrastructure

In this survey, we try to establish whether the access to information is a remaining challenge as indicated by previous literature reviews on African SMEs, and how the challenge is routed. Table 3 shows results obtained from SME respondents across the 5 participating countries.

Country / Survey	1	2	3	4	5	6
Egypt	66	42	45	52	51	44
Sudan	62	57	57	57	55	57
Cameroon	49	36	38	33	37	34
Senegal	57	60	48	35	53	34
Zimbabwe	61	57	54	49	52	53
Completed	295	252	242	226	248	242
Partial Responses	72	42	9	40	20	21

Table 2: Survey Respondents' Numbers per country

The survey's first question on information challenges among African SMEs yielded nuanced results. Responses to "My business lacks access to adequate business information" were evenly split, indicating a lack of unanimous awareness of this challenge. Notably, 43.8% agree, and 30.2% strongly agree that technology and support systems contribute to efficiency, emphasizing technology's recognition by 74% of respondents.

Concerning technological and communication infrastructures, only 36.6% express concerns about a poor information environment due to underdeveloped technological infrastructures. Similarly, for communication infrastructure, 32.2% agree that their business has a poor information environment, suggesting some concern but not serious malfunctioning.

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
My business lacks access to adequate business information	Count	16	72	65	71	18
	%	6.6%	29.8%	26.9%	29.3%	7.4%
Presence of appropriate technology and associated support systems like hardware and software make it easier for businesses like mine to be efficient and effective which in turn lowers costs of production and operations	Count	4	12	47	106	73
	%	1.7%	5.0%	19.4%	43.8%	30.2%
My business has a poor information environment resulting from underdeveloped technological infrastructures.	Count	12	69	70	75	16
	%	5.0%	28.5%	28.9%	31.0%	6.6%
My business has a poor information environment resulting from underdeveloped communication infrastructure	Count	10	69	85	61	17
	%	4.1%	28.5%	35.1%	25.2%	7.0%
My business has a poor information environment resulting from inadequate business support systems	Count	8	55	86	69	24
	%	3.3%	22.7%	35.5%	28.5%	9.9%
I believe that it is the role of the official authorities to facilitate access to reliable information for businesses like mine	Count	4	19	62	109	48
	%	1.7%	7.9%	25.6%	45.0%	19.8%

Table 3: Access to Information (N = 242)

Regarding the business support system, the survey results show a balanced perspective. Notably, 22.7% disagree, and 3.3% strongly disagree, while 28.5% agree, and 9.9% strongly agree that their business has a poor information environment due to inadequate business support systems. This distribution suggests that overall, surveyed SMEs do not perceive significant issues with their business support systems.

A consensus emerges regarding the role of authorities in supporting information access, with an overwhelming 66% of participants agreeing that it is the role of official authorities to facilitate access to reliable information for businesses like theirs, underscoring the perceived responsibility of authorities in enhancing information access for SMEs.

Regarding information sources, respondents predominantly rely on social media and digital press for obtaining information about their sector or activity, indicating a significant trend toward digital channels. Reliable sources mentioned include FAO OP, SocialnetLink, Euro News, Jeune Afrique, France 24, 3FPT, Seneweb.com, leral.net, and ANSD.

The split responses regarding access to adequate business information may stem from variations in SMEs' reliance on different sources of information and their perceptions of the sufficiency of available data. Those who agree or strongly agree that technology and support systems contribute to efficacy likely have access to robust technological infrastructure and support networks, enabling them to gather and utilize information effectively. Concerns about poor information environments due to underdeveloped technological and communication infrastructures may reflect challenges faced by SMEs operating in regions with limited access to technology and reliable communication networks. The balanced perspective on the role of authorities in supporting information access suggests that while SMEs recognize the importance of government support, they may also rely on alternative sources for information due to perceived limitations in official channels.

5.2 Perception of the government's Covid-19 related communication and ensuing pandemic measures

In this section we wanted to investigate how the policies and measures to combat the Covid-19 were perceived by African SMEs. Table 4 shows results obtained on this subject.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
The measures taken in response to the COVID-19 pandemic have been clear to me	15	29	41	119	44
Count	6.0%	11.7%	16.5%	48.0%	17.7%
%					
The communication undertaken about the measures taken in response to the COVID-19 pandemic have been effective	17	41	63	100	27
Count	6.9%	16.5%	25.4%	40.3%	10.9%
%					
In general, in my company, there was consensus about the measures taken in response to the COVID-19 pandemic	6	24	56	112	50
Count	2.4%	9.7%	22.6%	45.2%	20.2%
%					
The authorities changed often the measures taken in reaction to the COVID-19 pandemic	10	30	59	117	32
Count	4.0%	12.1%	23.8%	47.2%	12.9%
%					
Announcements regarding measures taken by the authorities in reaction to the COVID-19 pandemic were announced relatively late.	9	68	74	76	21
Count	3.6%	27.4%	29.8%	30.6%	8.5%
%					
The measures taken in response to the COVID-19 pandemic have been reasonable	9	39	65	113	22
Count	3.6%	15.7%	26.2%	45.6%	8.9%
%					
The measures taken by the authorities in response to the COVID-19 pandemic have taken into account social factors	16	47	70	90	25
Count	6.5%	19.0%	28.2%	36.3%	10.1%
%					
The measures taken in response to the COVID-19 pandemic have taken into account economic factors	24	65	62	79	18
Count	9.7%	26.2%	25.0%	31.9%	7.3%
%					
The measures taken by the authorities in response to the COVID-19 pandemic have taken into account health and safety factors	14	21	57	120	36
Count	5.6%	8.5%	23.0%	48.4%	14.5%
%					
In general, I believe there was a lack of consistency in the measures taken by the authorities in response to the COVID-19 pandemic	9	42	75	100	22
Count	3.6%	16.9%	30.2%	40.3%	8.9%
%					
In general, I believe the authorities will not impose very strict measures in case of a second COVID-19 or pandemic wave	19	46	83	79	21
Count	7.7%	18.5%	33.5%	31.9%	8.5%
%					

Table 4: Governments' policies and measures to combat the Covid-19 (N = 248)

Analyzing the survey data reveals a notable consensus on perceptions of clarity in response measures to the Covid-19 pandemic. A substantial 65.7% of respondents agree (48%) or strongly agree (17.7%) that these measures have been clear. Similarly, the majority, comprising 51.2%, either agree or strongly agree that communication about these measures has been effective, indicating a positive reception to pandemic-related announcements.

The timing of these announcements remains inconclusive from the data, as responses to the statement "Announcements regarding measures taken by the authorities in reaction to the Covid-19 pandemic were announced relatively late" do not provide a decisive trend.

Acceptance of the measures taken by government officials is widespread, with 65.4% expressing agreement (45.2%) or strong agreement (20.2%) that there was consensus within their companies. The perceived reasonableness of these measures is echoed by 64.8% of respondents, who agree (45.7%) or strongly agree (19.1%) with the statement "The measures taken in response to the Covid-19 pandemic have been reasonable."

Concerns arise about the consistency of these measures, with nearly half of the respondents (49.2%) indicating agreement (40.3%) or strong agreement (8.9%) that there was a lack of consistency in the authorities' responses. Notably, only 20.5% did not feel this inconsistency.

In terms of considerations for social, economic, and health factors, a mixed sentiment prevails. While 46.4% agree (and strongly agree) that social factors were considered, 39.1% hold a similar view for economic factors. The majority (62.9%), however, concurs that health and safety factors were taken into account by authorities.

Despite the perceived consideration of health and safety, there is an indication of relaxed implementation, as suggested by 47.2% who disagree (35.5%) or strongly disagree (11.7%) that the measures were strictly followed.

The high percentage of respondents agreeing or strongly agreeing that measures taken by authorities have been clear indicates a positive perception of government communication efforts. This clarity likely contributes to SMEs' understanding of pandemic-related regulations and their ability to comply with them. Similarly, positive perceptions of the effectiveness of communication about these measures suggest that government announcements and campaigns have effectively reached SMEs and conveyed relevant information. However, the inconclusive data regarding the timing of announcements may reflect varying experiences among SMEs regarding the timeliness of government communication, indicating potential room for improvement in the speed and frequency of information dissemination.

5.3 African SMEs’ response and ability to navigate the Covid-19’ measures.

In this survey segment, we aim to assess how African SMEs navigated and responded to the challenges posed by Covid-19 measures. Table 5 presents the data obtained on this section.

Most SMEs (71.7%) displayed a commendable understanding of Covid-19 safety protocols, ensuring their companies were well-prepared to address safety hazards related to Covid-19.

The data showing a lack of strict adherence to authorities' measures by nearly half of SMEs suggests various factors influencing compliance. Potential explanations include challenges in implementing measures due to operational constraints or conflicting priorities, differing interpretations of the necessity of certain restrictions, or perceptions of government measures as overly burdensome or ineffective. Satisfaction with operating hours may stem from SMEs' ability to adjust schedules to meet changing demands while balancing health and safety

considerations. The high percentage of respondents believing they have gained valuable insights for dealing with future health crises suggests a proactive approach to learning from current challenges and adapting strategies accordingly.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Implementation of the measures taken by the authorities in response to the COVID-19 pandemic was strictly followed					
Count	29	88	69	51	11
%	11.7%	35.5%	27.8%	20.6%	4.4%
In general, the hours of functioning for my business were reasonable and allowed us to continue offering our service / product					
Count	14	43	64	109	18
%	5.6%	17.3%	25.8%	44.0%	7.3%
Personally, I feel I have gained a learning that would allow me to function reasonably in case of rising COVID-19 numbers and / or a second wave					
Count	5	11	67	133	32
%	2.0%	4.4%	27.0%	53.6%	12.9%
In my company, we knew exactly how to deal with the safety hazards related to Covid-19					
Count	9	7	54	129	49
%	3.7%	2.9%	21.7%	52.0%	19.7%

Table 5: African SMEs response and ability to navigate COVID-19 measures (N=248)

5.4 Covid-19 information circulation

In this section we needed to know how effective the authorities were in circulating information about Covid-19 and raising awareness against the virus and what could be done as precautionary measures. Table 6 shows data collected on the subject of COVID-19's information circulation.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
The information about the physical distancing, barrier gestures and other Covid-19 prevention actions were circulated regularly					
%	3.6%	8.5%	24.7%	42.9%	20.2%
There were systematic awareness campaigns about the actions needed to prevent the spread of the Covid-19					
%	4.5%	6.1%	23.5%	47%	19%

Table 6: COVID-19 Information Circulation (N=248)

Authorities effectively communicated preventive actions, with only 3.6% strongly disagreeing and 8.5% disagreeing. Awareness campaigns about Covid-19 prevention received positive responses "The information about the physical distancing, barrier gestures and other Covid-19 prevention actions were circulated regularly". The majority acknowledged systematic awareness campaigns by officials.

Additionally, most respondents confirmed the use of a daily dashboard by their country's officials for COVID-19 information as indicated in Table 7.

Surprisingly, 16.6% are unsure about dashboards, and 10.5% believe there are none. Effective communication of preventive actions by authorities likely boosts SMEs' compliance with

Covid-19 measures. Regular dissemination of information about preventive actions shows proactive efforts. However, uncertainty about dashboards indicates potential awareness gaps. Trust in health-related information may hinge on messaging consistency and transparency. Conflicting information underscores the need for clear communication to build trust.

		Yes	No	I don't know
There was a daily dashboard of new COVID-19 contracted cases, deaths and other vital information	Count	180	26	42
	%	72.9%	10.5%	16.6%

Table 7: Daily dashboard of vital COVID-19 Information (N=248)

5.5 Trust in Government's COVID-19 circulated messages.

To dig further on the reliability and the usage of these dashboards, we probed our 72.9% (180 respondents) to answer a few more questions on their confidence and trust in the daily published Covid-19 cases". Table 8 hereunder shows gathered responses on that topic.

		Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
We trust the veracity of the numbers reported on the daily circulated dashboard of new COVID-19 contracted cases, deaths and other vital information	Count	21	43	63	41	12
	%	11.7%	23.9%	35.0%	22.8%	6.7%
I believe the numbers reported on the daily circulated dashboard of new COVID-19 contracted cases, deaths and other vital information reflected in a transparent way the real situation	Count	22	47	57	41	13
	%	12.2%	26.1%	31.7%	22.8%	7.2%
There were contradicting information released from different sources regarding new COVID-19 contracted cases, deaths and other vital information	Count	12	33	64	52	19
	%	6.7%	18.3%	35.6%	28.9%	10.6%

Table 8: Trust and reliability of daily Covid-19 related information and dashboard (N=180)

Unfortunately, the collected dashboard information is inconclusive regarding its reflection of reality and public trust in the communicated numbers (39.5% agree/strongly agree on contradicting information). This may stem from misinformation and conflicting reports, fostering skepticism and confusion among SMEs. Factors such as government transparency, communication effectiveness, and past experiences may influence trust levels. The prevalence of contradicting information highlights challenges in ensuring consistency during a crisis, warranting further exploration to develop strategies for enhancing trustworthiness and transparency in authorities' information dissemination.

6. Discussion

In light of the Theory of Motivated Information Management, our discussion highlights challenges for African SMEs as active digital information seekers, with government authorities as providers. However, our study didn't specify SMEs' information-seeking behavior, aligning with TMIM's acknowledgment of varying cognitive engagement.

Our survey echoes global challenges, emphasizing reliable information's crucial role for SME growth (Bunyasi et al., 2014). The reliance on digital channels mirrors challenges in navigating

online health information (Chi et al., 2020), particularly significant for African SMEs and authorities.

Perceptions of government communication during Covid-19 were generally positive, but concerns about inconsistency suggest the need for clearer strategies (Louis–Jean et al., 2020). Mixed sentiments on policymaking factors mirror decision-making complexities during crises. Regarding SME responses to Covid-19 measures, our survey aligns with previous findings, emphasizing adaptability and resilience (Fubah & Moos, 2022). Examining authorities' effectiveness in circulating Covid-19 information suggests a positive reception, but dashboard uncertainty aligns with literature on misinformation (Fan et al., 2020).

Survey results on trust and reliability reflect challenges of contradicting information, highlighting the need for transparent communication (Seytre, 2020). These findings contribute to understanding SME challenges, emphasizing targeted interventions. Distrust in government response isn't unique to Africa, reflecting global trends (Rowe et al., 2020). Trust in health information may stem from combined local and international initiatives (Braa, et al., 2023).

7. Conclusion

The survey offers insights into African SMEs' experiences during the COVID-19 pandemic, providing a modest contribution to scarce empirical evidence. Positive trends in information dissemination and response measures are evident, alongside challenges such as inconsistency in measures and concerns about data accuracy. These findings inform discussions on effective governance, communication strategies, and technology use for SMEs in Africa.

Study limitations include aggregating responses from diverse geographical locations, introducing complexity due to varying contextual factors and unique pandemic measures. Interpretation should consider nuances in regional responses and governmental strategies. Future research could benefit from a more granular examination, considering specificities of individual countries. Another limitation involves potential methodological bias associated with sample composition, influenced by SMEs' affiliation with the AfriLabs network, introducing selection bias towards high-tech SMEs.

Implications guide policymakers in formulating effective strategies for supporting African SMEs during pandemics. Insights into SMEs' access to information, perceptions of government communication, response to measures, and trust in information sources provide a comprehensive understanding of challenges. Policymakers can leverage these findings to tailor interventions conducive to SME growth and resilience, prioritizing robust information systems, transparent communication, and technology adoption.

Contemplating future research, the utilization of mixed methods emerges as promising. Triangulation with secondary sources such as the Rollo Africa expert panels would enrich analysis, highlighting the multifaceted nature of challenges. Future publications could benefit from a comprehensive mixed-methods framework, integrating quantitative survey data with qualitative insights, contributing to a more comprehensive discourse on supporting SMEs in times of crisis and humanizing the digital age.

Bibliography

Acquaah, M., Namatovu, R., & Kiggundu, M. N. (2021). Introduction to the special issue: preliminary investigations into the COVID-19 pandemic and management in Africa. *Africa Journal of Management*, *7*(1), 1-12.

Adam, N. A., & Alarifi, G. (2021). Innovation practices for survival of small and medium enterprises (SMEs) in the COVID-19 times: the role of external support. *Journal of innovation and entrepreneurship*, *10*(1), 15.

Afifi, W. A. and Weiner, J. (2004). Toward a theory of motivated information management. Communication Theory, 14(2), 167-190. https://doi.org/10.1111/j.1468-2885.2004.tb00310.x

Braa J;, Sahay S., Monteiro E. (2023) Design theory for Societal Digital Transformation: The Case of Digital Global Health, *Journal of the Association for Information Systems*, 24(6), 1645-1669. _DOI: 10.17705/1jais.00816

Bunyasi, G. N. W., Bwisa, H., & Namusonge, G. (2014). Effect of access to business information on the growth of small and medium enterprises in Kenya. *International Journal of Business and Social Science*, *5*(10).

Chi, Y., He, D., & Jeng, W. (2020). Laypeople's source selection in online health information-seeking process. Journal of the Association for Information Science and Technology, 71(12), 1484-1499. https://doi.org/10.1002/asi.24343

Deyshappriya, N. P., & Padmakanthi, N. P. D. (2022). Modelling the Impact of COVID-19 on SMEs' Performance Econometrics Evidence from Sri Lanka. *NPD, Modelling the Impact of COVID-19 on SMEs' Performance Econometrics Evidence from Sri Lanka.*

Eggers, F. (2020). Masters of disasters? challenges and opportunities for SMEs in times of crisis. Journal of Business Research, 116, 199-208. https://doi.org/10.1016/j.jbusres.2020.05.025

Fan, J., Wang, Z., Xie, Y., & Yang, Z. (2020, July). A theoretical analysis of deep Q-learning. In *Learning for dynamics and control* (pp. 486-489). PMLR.

Fubah, C. and Moos, M. (2022). Exploring covid-19 challenges and coping mechanisms for SMEs in the south african entrepreneurial ecosystem. Sustainability, 14(4), 1944. https://doi.org/10.3390/su14041944

Gisondi, M. A., Barber, R., Faust, J. S., Raja, A., Strehlow, M. C., Westafer, L. M., & Gottlieb, M. (2022). A deadly infodemic: social media and the power of COVID-19 misinformation. *Journal of medical Internet research*, *24*(2), e35552.

Haymaker, J.,Chachere, J., & Senescu, R. (2011). measuring and improving rationale clarity in a university office building design process. Journal of Architectural Engineering, 17(3), 97-111. https://doi.org/10.1061/(asce)ae.1943-5568.0000041

Kuckertz, A. and Brändle, L. (2021). Creative reconstruction: a structured literature review of the early empirical research on the covid-19 crisis and entrepreneurship. Management Review Quarterly, 72(2), 281-307. https://doi.org/10.1007/s11301-021-00221-0

Louis-Jean, J., Cenat, K., Njoku, C. V., Angelo, J., & Sanon, D. (2020). Coronavirus (COVID-19) and racial disparities: a perspective analysis. *Journal of racial and ethnic health disparities*, *7*, 1039-1045.

Morrow, R. L., Mintzes, B., Gray, G., Law, M. R., Garrison, S., & Dormuth, C. R. (2022). Industry sponsor influence in clinical trial reporting in canada: a qualitative interview study. *Clinical therapeutics*, *44*(3), 374-388.

Muriithi, S. M. (2017). African small and medium enterprises (SMEs) contributions, challenges and solutions.

Olawale, F., & Garwe, D. (2010). Obstacles to the growth of new SMEs in South Africa: A principal component analysis approach. *African journal of Business management*, *4*(5), 729.

Pavic, S. C. L. K., Koh, S. C. L., Simpson, M., & Padmore, J. (2007). Could e-business create a competitive advantage in UK SMEs?. *Benchmarking: An International Journal*, *14*(3), 320-351.

Pedauga, L., Sáez, F., & Delgado-Márquez, B. L. (2022). Macroeconomic lockdown and SMEs: the impact of the COVID-19 pandemic in Spain. *Small business economics*, *58*(2), 665-688.

Rowe F. Ngwenyama O., Richet J.L., Contact-tracing apps and alienation in the age of COVID-19, *European Journal of Information Systems*, vol 29, n°5, 2020, pp.545–562.

Seytre, B. (2020). Erroneous communication messages on COVID-19 in Africa. *The American Journal of Tropical Medicine and Hygiene*, *103*(2), 587.

Sudarmiatin, S., Wang, L., & Hidayati, N. (2023). The implementation of digital marketing in sme's: opportunities and challenges during the covid-19 pandemic. Indonesian Journal of Business and Entrepreneurship. https://doi.org/10.17358/ijbe.9.1.72

Van Tien, N. (2022). Recommendations Enhancing the Competitive Capacity of Small and Medium Enterprises after the Covid-19 Pandemic. *Emerg. Sci. J*, *6*, 122-135.

Zhao, E., Wu, Q., Crimmins, E. M., & Ailshire, J. (2020). Media trust and infection mitigating behaviours during the covid-19 pandemic in the usa. BMJ Global Health, 5(10), e003323. https://doi.org/10.1136/bmjgh-2020-003323

3. AI tool for cleaning up lewd music: A Jamaican Perspective

(Research- in-Progress)

Matthew Stone
University of the West Indies
matthew.stone@mymona.uwi.edu

Gunjan Mansingh
University of the West Indies
gunjan.mansingh@uwimona.edu.jm

Abstract

We propose the development of an innovative AI tool specifically designed to transform Jamaican music and other genres into clean radio-edit versions suitable for airplay. Utilizing advanced generative AI technology, our tool aims to sanitize songs while retaining their intrinsic essence. In this paper, we leverage modern generative AI techniques to recreate airplay-friendly versions within the same musical genre, focusing on Jamaican music's unique cultural significance. Our evaluation, conducted across multiple songs, reveals that the tool demonstrates commendable efficacy in its transformative task. However, ongoing refinements are necessary to ensure the tool consistently maintains the distinctive style and mood of the original songs.

1. Introduction

In an era where music serves as a ubiquitous form of expression, the diversity of lyrical content spans a broad spectrum, from the profound to the provocative. While artistic freedom remains a cornerstone of the music industry, the need for more family-friendly and socially acceptable versions of songs has become increasingly apparent. In response to this demand, we present a groundbreaking tool that leverages the power of generative AI technology to transform lewd songs into clean radio edits—a transformative endeavor at the intersection of creativity, technology, and responsible content consumption.

This paper delves into the development and methodology of our innovative tool, aiming to address the challenges posed by explicit content in music. Through a meticulous process involving background music removal, vocal-to-text conversion, and the application of advanced AI, our tool reimagines explicit lyrics with a focus on creating radio-friendly versions. The integration of GPT-4, a state-of-the-art generative language model, plays a pivotal role in generating clean and coherent renditions of the original lyrics.

As we navigate through the intricacies of our methodology, from the deconstruction of the song's components to the recreation of vocals in the style of the original singer, this paper sheds light on the technological intricacies underpinning our tool. Additionally, we explore the implications of our tool in mitigating explicit content, contributing to the ongoing dialogue regarding responsible content creation and consumption in the music industry.

The development of this tool not only showcases the potential of AI in addressing societal concerns related to explicit content but also raises questions about the ethical dimensions of content transformation. By providing a comprehensive overview of our methodology, implementation details, and evaluation results, this paper aims to contribute to the evolving discourse on the intersection of technology and creative expression in the musical landscape.

2. Related Work

Typically, in the music industry, making a clean radio edit of a song can involve many methods. One method involves replacing the explicit words with ad libs. Other methods include muting and reversing the segment of the audio that contains the explicit content. These methods come with its own issues, which include not knowing which words could be considered explicit. There are some curse words that are easily identifiable but other words in a different context to the wrong ear could be deemed as very inappropriate. There has been some work in using AI to automatically detect explicit lyrics in songs. One such method makes use of a tree classifier in combination with an LSTM to detect explicit content in English lyrics. This method was able to censor offensive content for children with a 96% accuracy (Chen et al., 2023). Another body of work investigated the automatic detection of explicit lyrics for Italian songs. In this paper they evaluated the use of neural language models on this task and they found that neural language models outperformed non neural language models (Rospocher, 2023). Another paper used DistilBert to detect objectionable song lyrics (Bolla et al., 2023).

A low resource language is one where there is relatively less data to train conversational AI systems such as Jamaican creole, compared to high resource languages such as English, French and Spanish. Several models have been proposed to improve on speech to text transcription for low resource languages Wav2Vec (Baevski et al., 2020) utilized a self-supervised approach and achieved impressive results on low resource languages. This was further improved with the introduction of Massively Multilingual Speech (MMS) (Pratap et al., 2023) from Meta. Whisper (Radfored et al., 2022) is a popular model from OpenAI for text to speech. Whisper is an automatic speech recognition (ASR) system trained on 680,000 hours of multilingual and multitask supervised data collected from the web. Approximately one third of Whisper's database is non-English, which is one of the reasons we decided to go with this model in our research.

The introduction of transformers (Vaswani et al., 2017) has led to an explosion of what are known today as large language models (LLMs). LLMs are large transformers that can consist of an encoder or decoder or both and are trained in an unsupervised way on a large corpus of text. One of the first examples of this is BERT (Devlin et al., 2018) which is an encoder transformer trained in an unsupervised way trained to predict masked tokens. Decoder style transformers such as GPT made use of next token prediction. These models have made massive improvements in downstream tasks such as question and answering, code completion and summarization. Several improvements on this model led to models such as GPT-3 (Brown et al., 2020), GPT-3.5 and most recently GPT-4 (OpenAI 2023). There was a question as to whether these models could interpret Jamaican language. This was investigated by researchers (Bonadio et al., 2022) and it was found that the state-of-the-art LLM at the time GPT-3 had some basic abilities to interpret the Jamaican creole, even managing to produce Jamaican style poems with it.

3. Methodology

We make use of the design science research methodology. The following outlines the steps we took to create the tool outlined above.

3.1. Problem Identification and Motivation

- Objective: Recognizing the need for family-friendly versions of songs in the music industry, our research addresses the challenge of transforming lewd content into clean radio edits.

3.2. Objectives of the Design

- Objective: Develop an innovative tool that utilizes advanced AI techniques to systematically convert explicit lyrics into culturally sensitive and radio-friendly versions, preserving the artistic intent.

3.3 Design and Development

- **Step 1: Song Loading:**
  - Description: Load the song into memory using the librosa Python library, setting the foundation for subsequent processing.
  - Objective: Establish a starting point for the transformation process.
- **Step 2: Background Music Removal:**
  - Description: Utilize the Demucs (Defossez et al., 2019) model for vocal extraction which is based on the U-Net convolutional architecture.
  - Objective: Isolate vocals for further analysis and modification.
- **Step 3: Speech to Text:**
  - Description: Transform extracted vocals into text using the whisper model.
  - Objective: Create a textual representation of the vocal content for subsequent modifications.
- **Step 4: Sanitize Lyrics with GPT-4:**
  - Description: Deploy GPT-4 with a customized prompt to sanitize lyrics for airplay.
  - Objective: Ensure the content aligns with cultural and broadcasting standards
- **Step 5: Text to Speech with GenerSpeech Model**:
  - Description: Use the GenerSpeech Model (Huang et al., 2022) to convert sanitized text back to speech while maintaining the original vocal style.
  - Objective: Reconstruct vocals in a style coherent with the original singer.
- **Step 6: Background Audio Reattachment:**
  - Description: Reattach the original background music to the transformed vocals using the librosa library.
  - Objective: Create a seamless, final version of the clean edited song.

3.4 Evaluation

- A panel of experts will be assembled with diverse backgrounds in music production, linguistics and AI technology. Each expert will be tasked with evaluating the transformed lyrics based on predefined criteria, including lyrical coherence, cultural appropriateness, and adherence to the original song's style and mood.

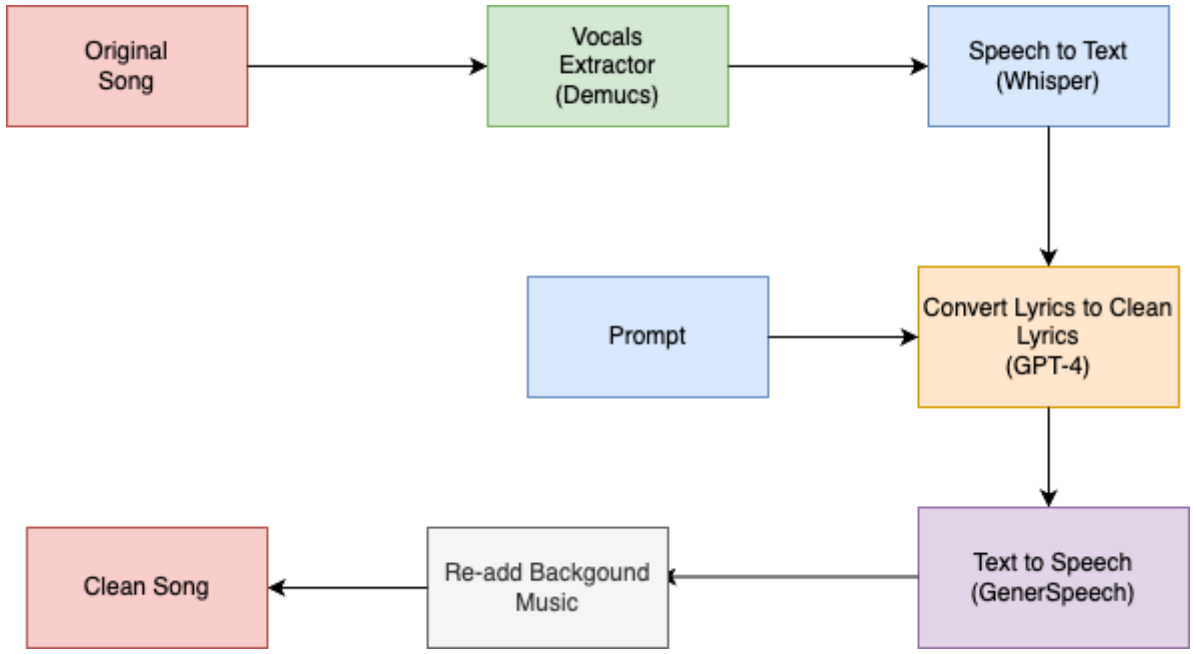


Figure 1: Workflow of Song Transformation Process

4. Results

In our preliminary evaluation of the AI tool, we conducted a self-evaluation using a sample of songs. Each transformed song was subjected to a thorough review by our team, focusing on aspects such as lyrical coherence, cultural appropriateness, and adherence to the original song's style and mood. We extracted music from Yyoutube videos to get the audio tracks. For this evaluation we only used 30 seconds of audio. The following are examples of the lyrics before and after edits were made. Here we show three songs we evaluated, two reggae tracks and one popular rap song. We used the following prompt.

"I want you to make the lyrics for the following song safe for play on the radio. Just imagine kids will be listening to this. Ensure the overall mood of the song is still the same. There shouldn't be any strong sexual innuendos. Also note this is a Jamaican song so ensure to use patois. The lyrics are as follows: {lyrics} "

Song Name	Actual Lyrics	Clean Lyrics	Transcribed Lyrics (Unedited)
WAP-Cardi B feat Megan the Stallion	Bring a bucket and a mop for this wet ass pussy Give me everything you got for this wet ass pussy Beat it up, nigga, catch a charge Extra large, and extra hard Put this pussy right in yo' face Swipe your nose like a credit card Hop on top, I want a ride I do a kegel while it's inside Spit in my mouth, look at my eyes This pussy is wet, come take a dive Tie me up like I'm surprised Let's role-play, I wear a disguise I want you to park that big Mack truck right in this little garage	*Bring a bucket and a mop for this wild, wild party Give me everything you got for this wild, wild party Enjoy the fun, friend, dance till you're charged Extra energy, extra hard Put these moves right in your space Swipe your foot like you're in a dance race Hop on the beat, I wanna glide I do a twist, what is inside? Sing out loud, blink in my eyes This party is live, come take a dive Cheer me on like I'm surprised That's roleplay, I wear disguise I want you to park that big food truck Right in this little star-struck.*	*Bring a bucket and a mop for this wet-ass pussy Give me everything you got for this wet-ass pussy Beat it up, nigga, catch a charge Extra large and extra hard Put this pussy right in your face Swipe your nose like a credit card Hop on top, I wanna ride I do a kegel, what is inside? Spit in my mouth, lick in my eyes This pussy is wet, come take a dive Tie me up like I'm surprised That's roleplay, I wear disguise I want you to park that big Mack truck Right in this little garage*
A so we like it - Spice	*mi like it* *Bring yo buddy, come, yah meck mi ride it Ride it like a bike it Cock up and a sitdown and a wine it* *Yes a so mi like it* *Bring yo buddy, come, yah meck mi ride it Ride it like a bike it Cock up and a sitdown and a wine it* *Pon the bass line, mi a buss a wine Fold legs together two a wi a combine* *Tear out mi grind bend mi back bruk mi spine Caw no gyal cyaa do dem yah wine weh mi*	*Mi tek delight. Bring yuh energy come yah let's catch a vibe. Catch it like a kite, it. Stand yah in di rhythm an' a groove it. Yes, mi fi tek delight. Bring yuh energy come yah let's catch a vibe. Catch it like a kite, it. Stand yah in di rhythm an' a groove it. Pon di bassline, mi and boss a groove. Two hearts together we a combine. Steal di rhythm, mi move til mi back can't recline. Dem cyaa, dem cyaa match dis groove with mi.*	*me like it. Bring your body come here and make me ride it. Ride it like a bike it. Cock up on a sit down and a wine it. Yes a for me like it. Bring your body come here and make me ride it. Ride it like a bike it. Cock up on a sit down and a wine it. On the bassline, me and boss a wine. Four legs together two a we a combine. Steal road me grind bin me back broke me spine. Cannot y'all can't do dem y'all wine with me.*
Gvnam Shift - Skeng	*Diss mi? yuh mussi sick* *An' a 17 lodge up inna dis 45 fit, mi claat it naah miss* *Mi finga move swift, mi gun dem naah stick*	*Dis a me, yuh mus' be confused, an' 17 stars we a raise inna this, 45 beats mi make it can't miss Mi fingers move quick, mi tune*	*This me, you must be sick, and a 17 large up inna this, 45 feet me crack it nah miss Me finger move swift, me gun*

	Govi-govi dem sick Diss mi? An' mi nuh kill two a dem up inna dis? 45 fit, mi claat it naah miss Mi finga move swift, mi gun dem naah stick Listen up War dem waah tell dem get di tanker Every guns up, every killa mask up Travel pon dem endz 'bout dozen claat up Dozen man haffi dead a weh yuh feel mi lef' some	*dem nah slip, govvy govvy dem sick Dis a me, an' mi nah do dem ruff tings inna this, 45 beats mi make it can't miss Mi fingers move quick, mi tune dem nah slip Raise him up, what a man, tell him get the vibes up, vibes every tune up, every rhythm catch up Catch up pon dem ends, but it's just a warm up, just a man have to step up if he left some.*	*dem nah stick, govvy govvy dem sick This me, and me no kill two a dem up inna this, 45 feet me crack it nah miss Me finger move swift, me gun dem nah stick This him up, what a man, tell him get the tank up, tank every guns up, every kill a mass up Mass up upon dem ends, but does a crack up, does a man have to dead away if he me left some*

Table 1: Comparative Analysis of Original and Transformed Lyrics

The tool effectively eliminated all explicit language from the songs. While GPT-4 encountered challenges in transforming Jamaican lyrics initially, refining the prompt significantly enhanced the output quality. The model demonstrated an effort to preserve the overall flow and melody of the songs. However, in certain instances, the transformed lyrics fell short of expectations, often generating basic rhymes lacking in cleverness. We posit that refining prompts further and fine-tuning the model specifically for Jamaican language nuances could lead to a substantial improvement in output quality.

5. Discussion

The development and evaluation of our AI tool for transforming songs into clean radio-edit versions, with a primary focus on Jamaican music, have yielded noteworthy insights. The tool demonstrated proficiency in successfully eliminating explicit language. However, challenges arose when GPT-4 attempted to transform Jamaican lyrics, requiring careful prompt tuning to significantly enhance output quality. Despite these challenges, the model showcased an inherent capability to preserve the general flow and melody of the songs, underscoring its adaptability to musical nuances.

Yet, certain instances revealed shortcomings in the quality of the transformed lyrics, often resulting in basic rhymes that lacked cleverness. Looking forward, we envision substantial improvements through the incorporation of refined prompts and the fine-tuning of the model to better capture the intricacies of Jamaican language and music. This suggests a promising avenue for future research, aiming to create a tool that not only meets broadcasting standards but also respects and enhances the artistic and cultural nuances of diverse musical genres. We also plan to embark on a more extensive evaluation where we will use expert evaluators to score the output of our tool. We plan to use a dataset of a thousand songs. The scoring metric will be based on lyrical coherence, cultural appropriateness, and adherence to the original song's style and mood. In essence, our tool presents a stepping stone toward AI-driven

content modification, prompting a broader exploration of its potential and challenges in the dynamic realm of music.

6. Conclusion

In conclusion we have shown that there is potential in using generative AI technology to make music, more specifically Jamaican music fit for airplay. We imagine this tool can be useful to DJs as well as artists as it will make it easier to play their music. In the future we look to improve the quality of the outputs by finetuning some of the models we used such as Whisper and GPT-4.

7. Acknowledgements

We would like to acknowledge some other colleagues that helped to develop the tool namely Jordan Madden, Dimitri Johnson and Mickoy Banton.

References

Baevski, A., Zhou, Y., Mohamed, A., & Auli, M. (2020). wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations. Neural Information Processing Systems, 33, 12449–12460. https://proceedings.neurips.cc/paper/2020/file/92d1e1eb1cd6f9fba3227870bb6d7f07-Paper.pdf

Bolla, Bharath & Pattnaik, Soumya & Patra, Sambit. (2023). Detection of Objectionable Song Lyrics Using Weakly Supervised Learning and Natural Language Processing Techniques. 10.13140/RG.2.2.21414.29760.

Bonadio, E., Mimler, M., Mansingh, G., Corthesy, N., B., McDaniel, G., Parkes, A., & Harvey, S. C. (2022). Implications of Artificial Intelligence in Action – a Jamaican perspective. Social Science Research Network. https://doi.org/10.2139/ssrn.4232629

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., . . . Amodei, D. (2020). Language Models are Few-Shot Learners. arXiv (Cornell University). https://arxiv.org/pdf/2005.14165.pdf

Chen, X., Aljrees, T., Umer, M., Karamti, H., Tahir, S., Abuzinadah, N., Alnowaiser, K., Eshmawi, A. A., Mohamed, A., & Ashraf, I. (2023). A novel approach for explicit song lyrics detection using machine and deep ensemble learning models. *PeerJ*, *9*, e1469. https://doi.org/10.7717/peerj-cs.1469

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv (Cornell University). https://arxiv.org/pdf/1810.04805v2

Huang, R., Ren, Y., Liu, J., Cui, C., & Zhang, Z. (2022). GenerSpeech: towards style transfer for generalizable Out-Of-Domain Text-to-Speech. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2205.07211

OpenAI. (2023). GPT-4 Technical Report. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2303.08774

Pratap, V., Tjandra, A., Shi, B., Tomasello, P., Babu, A., Kundu, S., Elkahky, A., Ni, Z., Vyas, A., Fazel-Zarandi, M., Baevski, A., Adi, Y., Zhang, M., Hsu, W., Conneau, A., & Auli, M. (2023). Scaling speech technology to 1,000+ languages. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2305.13516

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust speech recognition via Large-Scale Weak Supervision. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2212.04356

Rospocher, M. Detecting explicit lyrics: a case study in Italian music. Lang Resources & Evaluation 57, 849–867 (2023). https://doi.org/10.1007/s10579-022-09595-3

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All you Need. arXiv (Cornell University), 30, 5998–6008. https://arxiv.org/pdf/1706.03762v5

4. An Examination of Industry Privacy Statements in Top New Zealand Websites

Sumedha Mukherjee
Auckland University of Technology, New Zealand
jjz7723@autuni.ac.nz

Jairo Gutierrez
Auckland University of Technology, New Zealand
jairo.gutierrez@aut.ac.nz

Abstract

This paper presents a comprehensive analysis of various website privacy statements within six distinguished industries, namely: Retail, Banking, Travel & Tourism, Government, Healthcare and Entertainment in line with the requirements of the New Zealand Privacy Act 2020. As data privacy concerns are highly prevalent in a time where data is a highly valuable asset and leveraged for organizational competitive advantage, this study seeks to evaluate the extent to which popular websites in these industries adhere to the legal requirements and industry best practices that are outlined in the latest Privacy Act using a content analysis questionnaire. This study also seeks to evaluate Government privacy statements against other industries and evaluate further the published information security practices of New Zealand websites.

Keywords: Privacy Statements, Privacy Policy, Website Analysis.

1. Introduction

Data privacy has become cardinal to the safety of online users' personal information. The term "privacy" in the online context has evolved to include an individual's inclination to be aware of,or have control over their data despite the nature of collection, and by whom (Tjhin et al, 2016). The limitless nature of the internet has exhibited a need for individuals to provide personal information to make use of services online, and a need for organizations to collect this data to improve their services (Tjhin et al, 2016). Business integrity in the data privacy context in New Zealand is often reflected in the "privacy policy" or "privacy statement" of websites (Hooper & Vos, 2009), through contractual obligation to comply with the New Zealand Privacy Act 2020. Privacy mechanisms such as "privacy policies" and "privacy statements" are perceived to be foundational for data privacy best practices, as they inform users on the privacy practices of the online organizations (Tjhin et al, 2016). Hooper & Vos (2009) stated that the public quality of a given website is cognizant of how the organization who hosts it is appraised; and it is argued that websites serve as a representation of the backing of the organization's values, particularly with respect to privacy. Privacy statements on these websites are an example of privacy assurance to users, which disclose the nature of collection, protection, and storage of personal data from visiting individuals online (Mutimukwe et al, 2020). In recent literature, the perceived effectiveness of privacy policy statements shows an increase over other mechanisms, such as third-party seals, when examining online individual's willingness to provide their information (Fabian et al, 2017).

In New Zealand, the information privacy principles and codes of practice detailed in the Privacy Act 2020 provide a definition for the privacy rights of individuals, and therefore online customers. Hooper et al. (2007) stated that the primary criteria for evaluating privacy statements of online organizations would be the details in the privacy legislation that is pertinent to the respective country. There are other privacy legislations that may apply to New Zealand organizations with services online, such as the EU General Data Protection Regulation (GDPR) (Mori et al, 2022). Mori et al. (2022) noted that privacy researchers observed more than 50% of organizational privacy policies in the EU did not communicate the variety of information they collected, despite this being a requirement of the GDPR. Another piece of prominent privacy legislature is the California Consumer Privacy Act (CCPA), which is debated to be the most extensive privacy law in the United States applicable to websites (Nortwik & Wilson, 2022). In a study by Chen et al. (2021), 95 privacy policies of popular websites in the United States were examined against CCPA, and the results showed that customer's privacy rights were vague when addressing their privacy rights. The motivation for this study is to assess to what level and extent, a selection of New Zealand industry privacy statements, are in line with the New Zealand Privacy Act 2020. Within the organization websites, the privacy notices are examined using a content analysis questionnaire, with the motivation to understand how certain industries in New Zealand are communicating privacy practice information to their visiting users. In numerous studies published previously, the analysis of privacy policy material was measured against the older version of the Privacy Act 2020, the New Zealand Privacy Act 1993 (Hooper & Vos, 2009).

2. Literature Review

2.1 Privacy Policies

Online privacy statements are used as a method to provide assurance to online consumers that their personal information would not be revealed, and privacy respected (Tjhin et al, 2016). It is a common belief that privacy policies are the main way that users must keep note of an organization's privacy practices (Tjhin et al, 2016). Privacy statements on websites in general are rarely read by consumers, likely due to well-known assumption detailing that policies are difficult to comprehend (Fabian et al, 2017). Results by Fabian et al. (2017) after an analysis of 50,000 privacy policies of the most popular English-speaking websites show that on average, privacy policies in general are complex to read. This complexity undermines the idea that consumers gave their explicit agreement for the collection of their data, and arguably that consent is not implied if users had a lack of understanding for the policy (Tjihn et al, 2016). Internet users usually favor third-party seals and other options that relay data privacy practices over privacy policies (Fabian et al, 2017). In addition to the level of compliance, the nature of dubious wording used within the statements can further impair the experience of informing internet users of data-handling practices (Fabian et al, 2017).

2.2 The New Zealand Privacy Act

In New Zealand, the most current data privacy legislation is the Privacy Act 2020. Its purpose is to provide directions on the collection, storage, utilization and disclose of personal information and protect the privacy of individuals. The Privacy Act 2020 came into effect on the 1st of December 2020, which replaced the older Privacy Act of 1993 (Office of the Privacy Commissioner, 2013) and the older Act is frequently referenced in literature on privacy policies. Within the Privacy Act, there are thirteen privacy principles that regulate the way

businesses and organizations should gather, manage, and utilize personal information (Office of the Privacy Commissioner, 2013). Principle 3 dictates the way organizations should be transparent about why they are collecting personal information and what they intend to do with it; and this includes websites (Office of the Privacy Commissioner, 2013). Office of the Privacy Commissioner (2013) states that by following this principle, the individual is informed about the purpose of collecting the personal information, and often this is through a privacy statement. The privacy statement should ensure that individuals are aware why data collection is necessary, what it will be used for, and who it will be given to, that they can access this information and correct it if required. Furthermore, Principle 5 of the Privacy Act reiterates the requirement for storage and security of personal information, where agencies must ensure data is protected and preclude unauthorized use or disclosure of this information (Vos et al, 2020). In relation to New Zealand website content related to these principles, older literature has stated that a significant number of websites were unclear when it came to specifying the methods for the safeguard, collection, and retention of personally identifiable information with the Privacy Act of 1993 (Hooper et al, 2007). This obscurity occurred while they "retained their right to gather and exchange non personally identifiable information". Previous literature also stated that retailer websites did not provide enough information to protect consumer rights, despite privacy law in New Zealand not posing any obstacles to the growth of e-commerce (Chung & Paynter, 2002).

2.3 Related Work

There are several previous studies that focused on the older version of the Privacy Act in New Zealand, the Privacy Act 1993, however there is little findings of website analysis with regards to the updated Act. Mori et al. (2022) stated that analysis of legal compliance assisted by categorising the contents of a privacy statement and comparing the findings against legal standards is evident in other studies. In a paper by Hooper and Vos (2009), they examined the scope at which New Zealand websites complied with the Privacy Act 1993 to better understand the value of individuals' information privacy and the practices of handling information in New Zealand's online environment. Tjhin et al. (2016) conducted a similar study, with an analysis of privacy governance in New Zealand websites through a content analysis questionnaire. In another similar study, the privacy policies of 200 organizations within New Zealand and Facebook pages were examined by Vos et al. (2020), in which they used a content analysis questionnaire to investigate website privacy practices. The various papers noted the importance of articulating the national principles on the rights of users to privacy of their information online.

Content analysis questionnaires were used by several studies within a similar research area to thoroughly investigate privacy practices in the New Zealand online environment, and therefore comprehensively assess the components of the websites (Tjhin et al, 2016). Vorster and da Veiga (2023) conducted a research study using a quantitative data analysis methodology to examine privacy policies of websites, with the greater part of their guidelines requiring a "yes" or "no" answer. If the guideline within the questionnaire was not addressed fully, it was interpreted as not met. Mori et al. (2022) used a convolutional neural network to classify privacy policy material to evaluate compliance with legal standards. They found that legal compliance was higher in the wholesale, telecommunication, and financial industries. Finally, more than one million privacy policies were analyzed over a twenty-year period using an

automated tool in a study mentioned by Vorster and da Veiga (2023). The results found that privacy statements are evolving in complexity and size, particularly lacking transparency regarding third-party data collection and tracking technologies.

3. Research Design and Questions

Various challenges faced by similar research studies include lack of user trust in websites and heterogeneity of industry target and range. A large amount of research on privacy statements predates social media, smartphones, and the Big Data era (Obar & Oeldorf-Hirsch, 2020). As service delivery has increased on the internet, there has been a paralleled rise in privacy concerns due to the higher potential for organization to retain, process and exploit personal information (Mutimukwe et al, 2020; Vorster & da Veiga, 2023). Concerns are rising, as a recent survey detailed over half of participants spanning 25 countries are more apprehensive regarding their online privacy than one year before (Mutimukwe et al, 2020).

Lack of user trust in an online organization perpetuates a reluctancy to share information (Busalim et, 2019). Detailed in the study by Vorster and da Veiga (2023), many users associated privacy statements with a lack of transparency about the processing and usage of personal information. An earlier paper by Hooper and Vos (2009) found that a primary reason people had not given up information to a website was due to a lack of trust, which led to fabricated personal information; trust has been an issue since long. Vagueness and a lack of straightforwardness within privacy statements are frequently found within privacy statements and policies. Many studies referred to by Chaudhury and Choe (2023) detailed that results show websites are influencing customer's decisions on data-sharing using deceptive user tactics, such as implied consent and nudging for data collection.

3.1 Research Questions

The following research questions address gaps in the recent literature and contribute insight into the nature of alignment with the updated Privacy Act:

1. Do privacy statements on the most visited New Zealand websites reflect the information privacy principles of the Privacy Act 2020?
2. Are government privacy statements closely aligned to the information privacy principles of the Privacy Act 2020 compared to other industries?
3. Is an effort made to address personal information security in New Zealand privacy statements?

The Government category is often precluded from privacy statement data collection due to assumed compliance, therefore this study addresses this gap directly. The study also provides insight into the specific addressing of information security on top New Zealand websites which can aid in assessing transparency and security efforts by various industries. This last inquiry is of interest because information security is not explicitly addressed in the New Zealand Privacy Act 2020 other than when discussing security of stored data in Principle 5.

3.2 Research Design and methods

3.1.1 Research Instrument

A content analysis questionnaire was used to conduct the research in this study. The content analysis technique, as stated by Hooper and Vos (2009), is used to evaluate, and analyze elements of communication and draw accurate conclusions based on the context of the data on the websites. Content analysis can be described as an objective and systematic method that is

used to infer conclusions by recognizing specific attributes within content such as privacy statements (Salva & Martino, 2012). Through a content analysis questionnaire, the analyzed websites' established principles and intentions can be inferred accordingly, justified by the publicly available information being a representation of the values of the organization behind the website (Hooper & Vos, 2009). Strengths of using the content analysis technique include the flexibility of the method that is effective in both large sets of data and specific, smaller datasets (Salva & Martino, 2012). Content analysis as a methodology has been demonstrated as a reliable and productive method of attaining results (Salva & Martino, 2012).

The content analysis questionnaire was designed to analyze the alignment of various websites with the Information Privacy Principles in the New Zealand Privacy Act 2020. The content analysis questionnaire was developed without involving questions from Information Privacy Principles 4, 8, 10 and 13.

Principle 4 states that the collection of personal information must be lawful and seen as fair and reasonable in the circumstances. The manner of collection may vary greatly between organizations and is often an internal process that is not available publicly, rather implemented through internal policies. "Lawful" and "fair" collection may differ between industries and individuals.

Principle 8 states that an organization must check if personal information is up-to-date, complete, relevant and not misleading when disclosing it. Internal checks of personal information are not inclined to be published in a privacy statement and would likely require actual communication with the website where information has been given. Given the limitations of the study, Principle 8 was omitted from the questions.

Principle 10 lists exceptions for the use of personal information not for the original obtained purpose. These exceptions are not often disclosed within privacy statements as they are informational restrictions for an organization's internal data collection uses.

Principle 13 sets various restrictions on assigning unique identifiers to individuals, and that it may only be done when it is necessary to its processes. Unique identifiers are seldom mentioned in privacy statements and are often only implemented in distinct scenarios, such as NHI numbers within New Zealand healthcare. Including Principle 13 may unfairly favor the developed question toward the Government and Healthcare fields, which often require a unique identifier.

The selection and omission of various IPPs from the questionnaire development was carefully considered. In addition to the reasons listed above, time, scope and resources prevented the questionnaire from exceeding 17 questions. However, the questions were designed to encapsulate the IPPs that would take little effort for organizations to disclose within a privacy statement. For example, Principle 6 states that individuals have the right to access their personal information held by the organization. This principle can be easily explicitly stated without revealing any internal processes that the organization may have and is fundamental and important information for an individual to know when accessing the website.

Various questions may address more than one IPP, and this is because there is often overlap between the content of the principles. For example, Question 2 on the content analysis questionnaire asks if the purpose of information collection is stated. This addresses both

Principle 1 (organizations must only collect information for a lawful purpose connected to their functions), and Principle 3 (what to tell the individual about collection). Due to the nature of language and the expression of business activities across industries, an overlap of acknowledgement of the Privacy Act is to be expected, and this was considered when developing the questions. There were seventeen questions formed for this purpose, that were based upon one or many principles. Answers were either a "Yes" or "No" for each question. If a keyword from the question was not mentioned, or the answer was not explicitly stated, it was regarded as a "No". The selected websites generally did not have to delve into detail to warrant a "Yes" on the questionnaire; primarily they had to acknowledge the principle in some way, with some exceptions as detailed in the Findings section.

3.1.2 Sample and Data Collection

120 websites were selected and measured against the above content analysis questionnaire. 96% of websites sampled contained a privacy statement and were able to be analyzed.

The content analysis was performed on a sample of the most visited New Zealand websites in July 2023, using the website Semrush "semrush.com". The websites were selected according to the six different categories on the Semrush website: "Retail", "Banking", "Travel and Tourism", "Government", "Healthcare" and "Entertainment".

The Top 20 websites visited in each category listed above were selected for the study. The scope for domain selection was limited to extensions of ".co.nz", ".org.nz" and ".govt.nz". As the objective was to examine the privacy practices of the most visited websites in the New Zealand online environment, websites void of privacy statements were still selected for the study. If a privacy statement could not be located on the website, the data collected was a "No" for all research questions.

3.1.3 Content Analysis

To be regarded as a privacy statement, the body of text or hyperlink needed to be titled similar to "privacy statement" or contain the word "privacy". Alternatively, if the privacy statement was found inside the "Terms of Conditions" or similar, it was still regarded as a statement on the privacy values, and therefore considered as the "privacy statement" of the organization and was included in the study. Organizations with websites void of a body of text containing "Privacy" in the title were not considered.

To locate the privacy statement on the selected websites, several steps were taken and recorded to create reasonable judgement. Where there was a lack of direct link to a privacy statement found on the homepage of the website, the following steps were taken: 1) The terms and conditions were located and checked (if present), and 2) A search was performed in the website search bar for the world "privacy" (if present), and 3) A search was performed in Google with the website name and the word "privacy" (a check was then performed to ensure the domain was the same). Following these steps, there were no further actions to locate the privacy statement of the website, and the content analysis questions were all answered with "No". All privacy statements that were sampled for the study were subject to seventeen content analysis questions.

4. Findings

The highest alignment percentage of 92.5% with the Privacy Act was regarding Principle 2; *Is the information collected directly from the individual?* It was evident that most websites stated explicitly that they collect information directly from the user themselves, and many also stated the circumstances in which it was not directly collected from the individual.

Alternatively, the lowest alignment percentage of 42.5% with the Privacy Act was regarding Principle 12; *Is disclosure of personal information to an organization outside of New Zealand stated?* Most organizations did not state whether information would be disclosed overseas, rather that it would be disclosed to "third parties".

4.1 Information Privacy Principles: Questionnaire Analysis

17 questions were developed and designed from various Information Privacy Principles in the Privacy Act 2020. The specific IPP(s) addressed, and the data collected per question gives insight into the alignment levels of the sampled websites.

Question 1: *Does the website contain a privacy statement?* This question addressed Principle 3 of the Privacy Act. Approximately 96.7% of websites contained a privacy statement on their website.

Question 2: *Is the purpose of information collection stated?* This question addressed Principles 1 and 3 of the Privacy Act. Most websites directly addressed the purpose of which they were collecting information at 85%.

Question 3: Is it stated whether the data collection is compulsory or voluntary? This question addressed Principle 3 of the Privacy Act. Approximately half of websites did not state whether data collection was compulsory or voluntary, and consent was assumed.

Question 4: *Is it stated the consequence of not providing the information?* This question addressed Principle 3 of the Privacy Act. Approximately half of websites did not state the consequences for not providing information (48.3%), closely aligning with Question 3.

Question 5: *Is the information collected directly from the individual?* This question addressed Principle 2 of the Privacy Act. Most websites did state that the information was collected directly from the individual, with a large amount also mentioning the case in which it would not be directly collected.

Question 6: *Are the names of the collecting agencies stated?* This question addressed Principle 3 of the Privacy Act. Approximately 85% of websites stated the names of the collecting agencies.

Question 7: *Are the addresses of the collecting agencies stated?* This question addressed Principle 3 of the Privacy Act. In contrast to Question 6, only 59.2% of websites accompanied the name of the collecting agency with an address.

Question 8: *Are the names of the holding agencies stated?* This question addressed Principle 3 of the Privacy Act. Holding agencies were mentioned occasionally, however 24.2% of websites failed to disclose this information.

Question 9: *Are the addresses of the holding agencies stated?* This question addressed Principle 3 of the Privacy Act. The addresses of the holding agencies were mentioned by just over half of all websites selected.

QUESTION NO.	PRINCIPLE NO. ADDRESSED	QUESTION	YES	NO	% "YES"
1	3	Does the website contain a privacy statement?	116	4	96.7%
2	1, 3	Is the purpose of information collection stated?	102	18	85%
3	3	Is it stated whether the data collection is compulsory or voluntary?	59	61	50.8%
4	3	Is it stated the consequence of not providing the information?	62	58	51.7%
5	2	Is the information collected directly from the individual?	111	9	92.5%
6	3	Are the names of the collecting agencies stated?	102	18	85%
7	3	Are the addresses of the collecting agencies stated?	71	49	59.2%
8	3	Are the names of the holding agencies stated?	91	29	75.8%
9	3	Are the addresses of the holding agencies stated?	73	47	60.8%
10	11	Has the website stated the purpose for disclosing information to other agencies?	82	38	68.3%
11	5	Are there safeguards in place to prevent loss or misuse of information collected?	97	23	80.9%
12	3,6	Can people request access to their personal information?	106	14	88.3%
13	6	If there are restrictions to accessing personal information, is the reason stated?	50	70	58.3%
14	3,7	Can people correct their personal information?	105	15	87.5%
15	9	Is the length of retention time of personal information for the purpose stated?	58	62	48.3%
16	12	Is disclosure of personal information to an organization outside of New Zealand stated?	51	69	42.5%
17	3	Is the Law under which personal information is required to be collected explicitly stated? (e.g. Tax Act)	67	53	55.8%

Table 1. Privacy Act Questions and Corresponding Answers from Most Visited New Zealand Websites

Question 10: *Has the website stated the purpose for disclosing information to other agencies?* This question addressed Principle 11 of the Privacy Act. A lower number of websites at 68.3% addressed the purpose for disclosing information to other agencies. Often there was only observed to be a statement on which third parties would be disclosed such information, however no reason accompanied.

Question 11: *Are there safeguards in place to prevent loss or misuse of information collected?* This question addressed Principle 5 of the Privacy Act. 80.9% of websites addressed that there were security measures in place, however most websites did not state any specific safeguards.

Question 12: *Can people request access to their personal information?* This question addressed Principle 3 and 6 of the Privacy Act. Approximately 88.3% of websites explicitly stated that personal information can be accessed, usually accompanied by an email or address.

Question 13: *If there are restrictions to accessing personal information, is the reason stated?* This question addressed Principle 6 of the Privacy Act. Almost half, 41.7% of websites did not state if there were restrictions to accessing personal information.

Question 14: *Can people correct their personal information?* This question addressed Principle 3 and 7 of the Privacy Act. Most websites detailed that personal information could be corrected, usually stated in line with Question 12.

Question 15: *Is the length of retention time of personal information for the purpose stated?* This question addressed Principle 9 of the Privacy Act. Over half of websites did not address retention time of personal information (51.7%).
Question 16: *Is disclosure of personal information to an organization outside of New Zealand stated?* This question addressed Principle 12 of the Privacy Act. Most websites, 57.5%, did not state whether personal information would be disclosed overseas for any reason.
Question 17: *Is the Law under which personal information is required to be collected explicitly stated? (e.g. Tax Act).* This question addressed Principle 3 of the Privacy Act. A low number of websites addressed the Privacy Act 2020 at all, and there were several websites that addressed it incorrectly due to a variety of reasons.

The number of websites that aligned with the principles is shown to be generally varied with each question asked, with particular questions, such as Question 16, having a high misalignment with the accompanying principle (IPP 9). The answers show clear priorities in being explicitly aligned with the Privacy Act, while lacking in addressing other areas.

4.2 Discussion

Further analysis was performed to compare the six different industries targeted in this research: Retail, Banking, Travel and Tourism, Government, Healthcare and Entertainment. Question 17 on the questionnaire is detailed: *Is the Law under which personal information is required to be collected explicitly stated? (e.g. Tax Act).* If the New Zealand Privacy Act was not mentioned, it was regarded as a "No", even if other Laws were mentioned, as the website is still subject to the Privacy Act as an organization operating in the New Zealand online environment. Similarly, if the correct year was not mentioned, or there was no year listed at all, it was also regarded as a "No".

Question 17 was not addressed or addressed incorrectly by 44.2% of all selected websites. Websites that earned a "No" answer on the questionnaire for Question 17 were split up by Reason in Table 2. This table shows the percentage of websites that did not comply with this subsection of Principle 3 of the Privacy Act and the reasons observed.

52.8% of these websites did not address the Privacy Act 2020 at all, and various other reasons for earning a "No" for Question 17 include addressing the old Act (Privacy Act 1993), no year stated and stating other laws such as GDPR but lacking a statement on New Zealand privacy legislation. An interesting observation about Question 17 is the lack of explicit statement of the New Zealand Privacy Act 2020. Many websites fail to mention the Act, but rather would state "relevant privacy legislation", "New Zealand legislation", or simply "Privacy Act". Many of the privacy statements had not been updated since the release of the new Act that replaced the old Privacy Act of 1993. The selected websites seem to portray the acknowledgement of the word "privacy" as an organization with a New Zealand domain as enough regard for the Act itself. The explicit statement of New Zealand privacy law is an area requiring improvement.

Question 11 addressed Principle 5: *Are there safeguards in place to prevent loss or misuse of information collected?* The websites (80.9%) that contained a statement on safeguards were analyzed further on the level of specificity their statement contained. Figure 1 shows the

percentage of privacy statements against two criteria:“Specific” and “Non-Specific”. Websites were classified as having “Specific” safeguards if they stated any detailed security processes or protocols, for example, encryption methods, internal procedures, or physical security measures. Otherwise, a generic statement earned a “Non-Specific” in the data collection.

REASON FOR “NO”	COUNT OF “NO”	% OF “NO”
NOT ADDRESSED	28	52.8%
ANOTHER LAW ADDRESSED	12	22.6%
OLD NZ ACT ADDRESSED	9	16.9%
NO YEAR STATED	4	7.5%

Note. The count of “NO” is derived from n=53 expressed in Question 17.

Table 2. Reasons for Failure to address Question 17

Information security is not specifically addressed in the New Zealand Privacy Act 2020; there is no legal requirement stated on how specific organizations need to be when disclosing the safeguards put in place to protect personal information. The 5th principle states that information should be protected “by such security safeguards as are reasonable in the circumstances to take”, which in its articulation, lacks a distinctness that is essential when addressing security. The Privacy Act states that reasonable security safeguards must be in place to prevent loss, access, use, modification, disclosure that is not authorized by the organization. Most websites directly quoting the Privacy Act were ambiguous, using statements such as “We will take reasonable steps to ensure personal information is protected from loss, misuse, disclosure or modification.”.

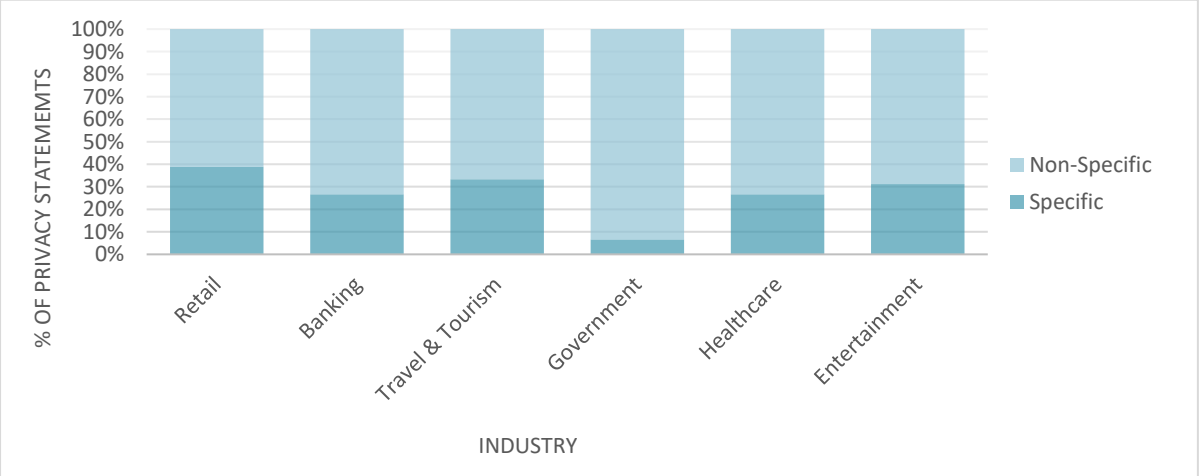


Figure 1. Specificity of Personal Information Safeguards

In Figure 1, we observe that 70 out of 97 of selected websites that had a safeguard statement, did not state what specific safeguards were in place to protect personal information. This indicates that around 71% of websites did not specify information security best practices such as encryption methods, employee training or security of servers storing information. Rather, user blame was high amongst all selected websites; emphasis was placed on the user's responsibility in protecting their information and displacing responsibility onto the user of the website if any security breach were to occur. The largest number of non-specific safeguards mentioned in privacy statements was from the Government sector, followed by the Healthcare and Banking industries. However, all industries were observed to have poor details of their actual security practices.

The assumption that government websites, particularly New Zealand government websites (that primarily have an extension of ".govt.nz") are more trustworthy and transparent in their respective privacy statements is challenged within the results of the study. Figure 2 displays the spread of "Yes" vs. "No" responses to all seventeen questions in the content analysis questionnaire in the Government industry.

Figure 2 shows that only Questions 1 and 5 were fully addressed by the selection of Government websites. Question 16, which is regarding the disclosure of personal information to overseas agencies, is a poorly addressed area in the Government sector. This is closely followed by Question 13, regarding restrictions to accessing personal information. This is arguably an important question to address, as government websites receive a wide variety of personal information, such as employment, immigration, criminal, and financial data. If there are restrictions to accessing such sensitive information about individuals, it should be explicitly stated. The Government sector was observed to have the greatest number of websites that did not state *specific safeguards* for protecting personal information. The statements were noted to be generic and ambiguous if any were stated at all.

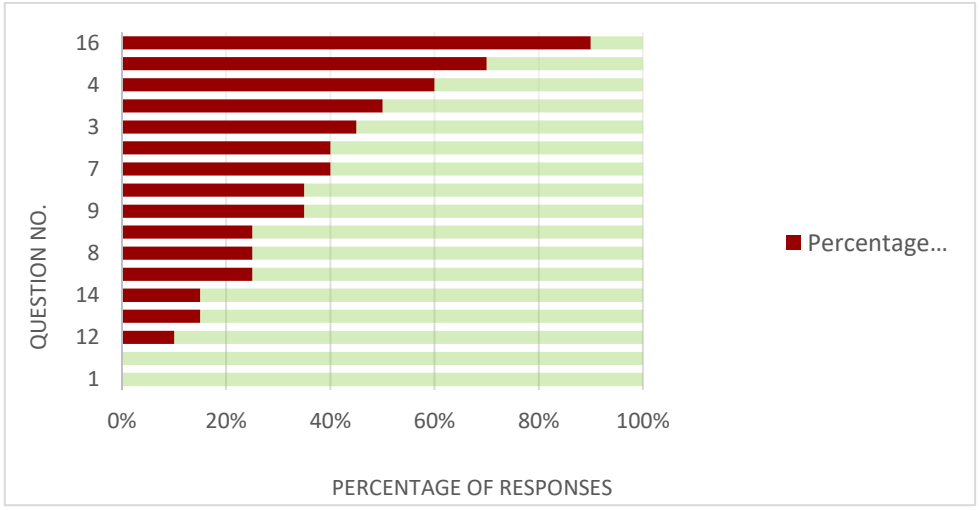


Figure 2. Government Privacy Statements: Bridging the Gap in Perception

5. Conclusions

The purpose of this study was to research the extent of alignment of a sample of New Zealand websites' privacy statements with the Privacy Act 2020. The Privacy Act defines the principles of privacy for users in the New Zealand online environment, and therefore provides an accurate basis for the content analysis questionnaire used in this study. The first research question addressed whether privacy statements on the most visited New Zealand websites can reflect the privacy principles in the latest New Zealand Privacy Act. This was not supported by the data collected, due to the fluctuating nature of explicit content. None of the questions were addressed by 100% of the examined sites, however in addition, the percentage of non-coverage was sometimes as low as 40%.

The second research question addressed whether the Government sector contained privacy statements that are closely aligned to the Act compared to other industries – and this was also not supported. The Government sector consistently assumed that the user could refer to the Privacy Act itself, and these sites also had a high level of non-coverage for multiple questions on the questionnaire. The third research question addressed whether efforts were made to specify personal information security. Little to no effort was made by most websites to address the security of personal information in detail, and the ambiguity of their statements point toward the generalized statement in the Privacy Act itself. Improvements are needed to ensure that users are aware of how their information will be protected.

5.1 Limitations and Future Work

A sample of the Top 120 New Zealand websites were selected for the study. Limitations include that the findings from these websites could not draw definite and conclusive results that represent all New Zealand websites with privacy statements. A bigger sample size would need to be obtained to represent the population of privacy statements more accurately in New Zealand. Another limitation is that the representation of industries was limited to only six: Retail, Banking, Travel & Tourism, Government, Healthcare and Entertainment.

Privacy statements that model the Privacy Act 2020 and any future Acts need to have a profile of elevated prominence for users to observe them more and raise user trust in organizations. From the findings of this study, it is evident that even the most visited New Zealand websites have not established a strong correlation between articulated content in their privacy statement and its potential to guide customer behavior through the promotion of their privacy values. A positively framed privacy statement that avoids user blame and displaced security responsibility may be a good starting area for improvement. Future work could include assessing governmental privacy statement practices against the non-profit industry in attempt to draw insights into the private and public sector relationships with New Zealand privacy legislation. Further research into published privacy statements against the organizations' real practices would be a significant insight into the importance of how privacy legislation needs to be refined.

References

Busalim, A., Hussin, A., & Iahad, N. (2019). Factors influencing Customer Engagement in Social Commerce Websites: A Systematic Literature Review. *Journal of Theoretical and Applied Electronic Commerce Research, 14*(2), 1-14. https://dx.doi.org/10.4067/S0718-18762019000200102

Chaudhury, R. D., & Choe, C. (2023). Digital Privacy: GDPR and Its Lessons for Australia. *Australian Economic Review*, *56*(2), 204-220. https://doi.org/10.1111/1467-8462.12506

Chen, R., Fang, F., Norton, T., McDonald, A. M., & Sadeh, N. (2021). Fighting the fog: Evaluating the clarity of privacy disclosures in the age of ccpa. *Proceedings of the 20th Workshop on Workshop on Privacy in the Electronic Society*, 73-102. 10.1145/3463676.3485601

Chung, W., & Paynter, J. (2002). Privacy issues on the Internet. *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, 1-9. https://doi.org/10.1109/HICSS.2002.994191

Dehling, T., Gao, F., & Sunyaev, A. (2014). Assessment Instrument for Privacy Policy Content: Design and Evaluation of PPC. *Proceedings of the Pre-ICS Workshop on Information Security and Privacy*, 1-16. https://papers.ssrn.com/abstract=2646460

Fabian, B., Ermakova, T., & Lentz, T. (2017). Large-scale readability analysis of privacy policies. *Proceedings of the International Conference on Web Intelligence*, 18-25. https://doi.org/10.1145/3106426.3106427

Hooper, A., Bunker, B., Rapson, A., Reynolds A., & Vos, M. (2007). Evaluating Banking Websites Privacy Statements – A New Zealand Perspective on Ensuring Business Confidence. *PACIS 2007 Proceedings.* https://aisel.aisnet.org/pacis2007/25/

Hooper, T., & Vos, M. (2009). Establishing business integrity in an online environment: An examination of New Zealand web site privacy notices. *Online Information Review*, *33*(2), 343-361. https://doi.org/10.1108/14684520910951258

Mori, K., Nagai, T., Takata, Y., & Kamizono, M. (2022). Analysis of Privacy Compliance by Classifying Multiple Policies on the Web. *2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC)*, 1734-1741. https://doi.org/10.1109/COMPSAC54236.2022.00276

Mutimukwe, C., Kolkowska, E., & Grönlund, Å. (2020). Information privacy in e-service: Effect of organizational privacy assurances on individual privacy concerns, perceptions, trust and self-disclosure behaviour. *Government Information Quarterly*, *37*(1). https://doi.org/10.1016/j.giq.2019.101413

Nortwick, M., & Wilson, C. (2022). Setting the Bar Low: Are Websites Complying with the Minimum Requirements of the CCPA? *Proceedings on Privacy Enhancing Technologies*, 608–628. https://doi.org/10.2478/popets-2022-0030

Obar, J. A., & Oeldorf-Hirsch, A. (2020). The biggest lie on the Internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society*, *23*(1), 128-147. https://doi.org/10.1080/1369118X.2018.1486870

Office of the Privacy Commissioner. (2013). Privacy Act 2020 and the Privacy Principles. https://www.privacy.org.nz/privacy-act-2020/privacy-principles/

Privacy Act 2020. https://www.legislation.govt.nz/act/public/2020/0031/latest/LMS23342.html

Savla, P., & Martino, L. D. (2012). Content Analysis of Privacy Policies for Health Social Networks. *2012 IEEE International Symposium on Policies for Distributed Systems and Networks*, 94-101.10.1109/POLICY.2012.20.

Tjhin, I., Vos, M., & Munaganuri, S. (2016). Privacy Governance Online: Privacy Policy Practices on New Zealand Websites. *Pacific Asia Conference on Information Systems.* https://www.semanticscholar.org/paper/Privacy-Governance-Online%3A-Privacy-Policy-Practices-Tjhin-Vos/4a2dc7fac576b8f9d7523c0f5de2036eec866c22

Vorster, A., & da Veiga, A. (2023). Proposed Guidelines for Website Data Privacy Policies and an Application Thereof. *Human Aspects of Information Security and Assurance*, *674*, 192-210. https://doi.org/10.1007/978-3-031-38530-8_16

Vos, M., Hu, M., & Du, B. (2020). Privacy on Facebook Brand Pages: A Content Analysis Study of New Zealand Organisations. *ACIS 2020 Proceedings.* https://aisel.aisnet.org/acis2020/53/

5. An Exploration of Online Instructor Performance Prediction Model Using LMS Data with Machine Learning Techniques

Gary Yu Zhao
Northwest Missouri State University
zhao@nwmissouri.edu

Cindy Zhiling Tu
Northwest Missouri State University
cindytu@nwmissouri.edu

Yufei Yuan
McMaster University
yuanyuf@mcmaster.ca

Abstract

Predicting online instructors' performance can help higher education institutions find issues and problems in the learning and teaching process as early as possible, giving them timely interventions to help ensure the quality of educational services. One of the significant challenges of evaluating and predicting online instructors' performance is to determine the key indicators from a large amount of data generated from the learning management system (LMS). The recent advancements in the continuous development of machine learning technology have led to a new momentum of teaching performance prediction in online education. In this paper, we follow the design science research methodology. Firstly, a dataset is collected from a Midwest university LMS platform, and we use a wrapper-based method to explore a simple and effective prediction model to select four sets of key influence variables. Then, we fit, tune, and compare various machine learning prediction models on different selected variable sets. Finally, we suggest that the Random Forest (RF) on eight selected variables with the Synthetic Minority Over-sampling Technique (SMOTE) is the best prediction model based on our work.

Keywords: Online instructor performance, Prediction model, LMS data, Machine learning, Design science.

1. Introduction

Online programs and students have increased dramatically in higher education. Higher educational institutions have large volumes of data from ERP (Enterprise Resource Planning) and LMS (Learning Management System) platforms which can be used to improve the quality of academic programs, services, and managerial decisions (Anwar et al., 2014). Researchers have done much research on predicting students' academic performance in the online education context based on different modeling methods (Ahajjam et al., 2022; Chandna et al., 2021; Karimi et al., 2020; Khan et al., 2021; Yang, 2021; Yu, 2021). Most existing studies use student demographic information, student socioeconomic data, students' perception of the academic learning processes, and students' key behavioral characteristic data to train and evaluate the prediction models (Costa et al., 2020). However, students' academic performance heavily relies on teaching quality and can reflect instructors' performance per se. In the concurrent learning analytics field, there is a significant lack of studies on constructing instructor performance prediction models (Abunasser et al., 2022; Li & Su, 2020; Wang et al., 2020).

There are many performance-related indicators that we can retrieve from LMS, e.g., instructor activity time, page views, class dropout rate, assignment grading days, student page views, student activity time, student assignment submission rate, on-time submission rate, student participant rate, student grades, and so on. Various machine learning methods can be employed for building the prediction model. Our study aims to explore mainstream prediction models of machine learning algorithms with various selected variable sets in the same data scenario to seek the improvement of performance and interpretability of the prediction model. Specifically, we answer the following research questions: (1) How do different selected variable sets affect the prediction performance of machine learning models? (2) How do we choose a better prediction model and what is the suggested one for online instructors' performance prediction?

By following design science research methodology (DSRM) (Hevner et al., 2004; Peffers et al., 2007), our research focuses on using LMS data to predict online instructors' performance with various widely used machine learning algorithms as well as a few ensemble methods, employing a wrapper-based variable selection method within the same data context. Additionally, we assess and compare these prediction models using metrics such as accuracy rate, precision and recall rate, F1 score, and time cost. The contribution of our work lies in the artifact of proposed online instructors' performance prediction model and the methodological contribution in modeling processes and methods.

2. Related Work

Fewer studies have been done on instructors' performance prediction in the online education context compared to student performance prediction. However, the methodologies of studies on students' performance prediction can be used for reference in our work because instructors' teaching performance can be reflected in the student's academic achievement. We review and synthesize the existing research from three critical components of constructing the performance prediction model: LMS data, machine learning algorithms, and evaluations.

2.1 LMS Data

As the LMS platform becomes widely adopted, scholars can use ample records of instructors' and learners' activities to construct performance prediction models (Jiao et al., 2022; Mi et al., 2022; Yang, 2021; Yu, 2021). Khan et al. (2021) used the data of students' gender, age, exam score, and assignment grades to predict students' academic performance. Karimi et al. (2020) utilized student demographic data, assessment results, and daily login data to the university's virtual learning platform. Pan et al. (2016) utilized quantile regression analysis to explore the impact of login frequency, time spent browsing teaching materials, and engagement in curriculum discussions on the final-term performance assessment of teachers. However, due to the high acquisition ability and availability, perception data still has been extensively used for building the teaching prediction model in concurrent research (Abunasser et al., 2022; Agaoglu, 2016; Ahmed et al., 2016; Ye et al., 2020). The perception data in online learning and teaching context is generally extracted from the response to the questionnaire primarily in four categories: student-student interaction, teacher-student interaction, content, system flexibility and convenience (Martínez-Caro et al., 2015).

2.2. Machine Learning Algorithms

Machine learning algorithms are the primary technical methods used in the learning analytics context. Traditional machine learning algorithms regard performance prediction as a classification problem. In contemporary literature, four machine learning algorithms are extensively used as the mainstream methods: Logistic Regression (LR) (Abunasser et al., 2022; Ahmadi & Ahmad, 2013; Mi et al., 2022), Support Vector Machines (SVM) (Abunasser et al., 2022; Agaoglu, 2016; Ye et al., 2020), Decision Tree (DT) (Abunasser et al., 2022; Agaoglu, 2016; Ahmadi & Ahmad, 2013; Ahmed et al., 2016; Kathiroli et al., 2020), and Multi-Layer Perceptron Neural Network (MLPNN) (Abunasser et al., 2022; Agaoglu, 2016; Ahmed et al., 2016; Pan et al., 2016; Yu et al., 2018).

LR is a supervised learning method based on generalized linear regression analysis. By integrating a logical model into the Linear Regression model, LR converts the continuous values predicted by the Linear Regression model to discrete values (Raschka, 2015). In logistic regression, the Sigmoid function is generally selected as the mapping function to solve the binary classification problem. SVM is considered one of the most popular methods in machine learning algorithms, which is an outstanding algorithm to fit small-scale datasets and is not sensitive to the dimensionality of data. It can transform a non-linear separable dataset into a linearly separable dataset, which can better solve the non-linear classification problem (Tian et al., 2012). The DT method uses a tree-shaped graph to model the possible decisions and the corresponding results. Some well-known impurity measures are used to find the homogeneity of instances in a node of the tree, such as information gain (ID3), gain ratio (C4.5, C5.0), and Gini index (CART) (Agaoglu, 2016). MLPNN is an artificial neural network consisting of multiple layers of nodes or artificial neurons. The network is organized into an input layer, one or more hidden layers, and an output layer. Each node in one layer is connected to every node in the next layer, and each connection has a weight associated with it (Mi et al., 2022).

Other than the abovementioned four machine learning algorithms, some ensemble algorithms show promising prediction performance, such as Random Forest (RF) (Kathiroli et al., 2020), Gradient Boosting Machine (GBM), and XGBoost (Abunasser et al., 2022). In addition, Naive Bayes' (NB) (Ahmed et al., 2016) and KNN's (Naser et al., 2015) algorithms are used to construct the performance prediction model in existing research.

2.3. Evaluation metrics

In existing studies, the primary evaluation metrics of prediction model performance include accuracy, precision, recall, and F1-score (Abunasser et al., 2022; Agaoglu, 2016; Ahmadi & Ahmad, 2013; Ahmed et al., 2016; Kathiroli et al., 2020). Accuracy assesses the proportion of correct predictions out of all predictions made. Precision quantifies the correctness of positive class predictions made by the classifier, while recall gauges the rate of correctly predicted positives by the classifier. In addition, the time cost is counted to get a better understanding of the efficiency of the models.

3. Methodology

In accordance with Hevner et al. (2004), design science research (DSR) is characterized by a build-and-evaluate approach aimed at creating artifacts. Our primary objective was to conceptualize and construct artifacts that facilitate the prediction of online instructors' teaching

performance. Given that suggested prediction model represents a novel perspective, its construction process can be likened to a search process, involving iterative assessment and enhancement of these artifacts. The research approach follows Peffers et al. (2007) DSR processes. As shown in Figure 1, we start with the problem identification and objectives of the solutions. The design and development process encompasses data collection, data preprocessing, variable selection, fitting and tuning machine learning models. The demonstration and evaluation processes include metric measurement, model optimization, and model comparison analysis. The DSR methodology ends with communication which means the suggested model, model interpretation, and academic publications. To obtain a good solution, we focus on the core processes: develop the solution (data collection and model construction), demonstrate and evaluate the solution. Notably, these processes can be iteratively conducted entirely or partially.

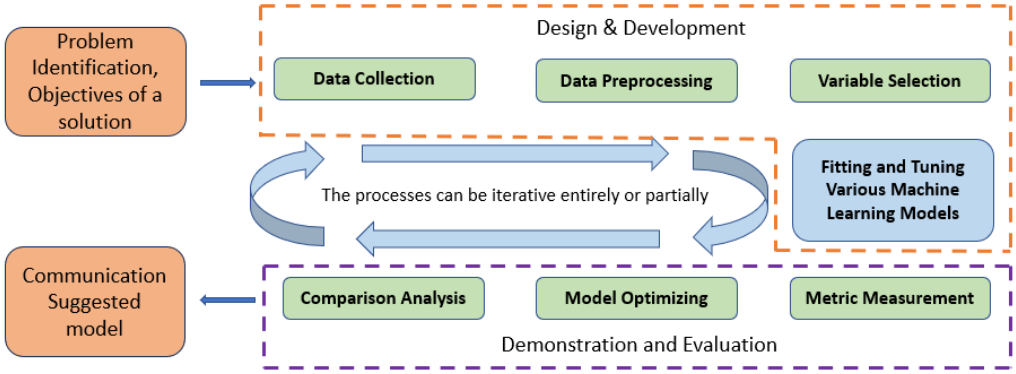


Figure 1: Research Methodology (Following Peffers et al. 2007)

3.1. Data Collection

Data was collected from a public Midwest university containing 120 online instructors' 2000 online classes taught from 2019 to 2022. The raw data was retrieved from a data warehouse of Canvas - an LMS platform. All data were aggregated to the course level if needed. We use instructor's evaluation scores to indicate instructor performance (dependent variable). The independent variables consist of behavioral features, including teaching and learning process data and summative data in the LMS platform. All variables are described below:

(1). Course_id_hash_encoded: Hashed Identity string of online class
(2). Instructor total_activity_time: Accumulated instructor action time in a specific class
(3). Instructor page_views: Instructor's viewed pages on the course site
(4). Published_on_time: Instructor publishes the course site on the term start date or late
(5). Mean_grading_days: Average days of all assignment grading in a course
(6). Total_student_enrolled: Number of students who enrolled in a specific class
(7). Dropout_rate: Number of dropped students out of total students in a course
(8). Retention_rate: 1- Dropout rate
(9). Evaluation total_responders: Total number of students who should evaluate the course instructor
(10). Evaluation total_responded: Total number of students who have completed instructor evaluation
(11). Total_response_rate: Total responses out of total responders

(12). Student average_activity_time: Average student action time on the course site
(13). Student average_page_views: Average student-viewed pages on the course site
(14). Student avg_participation_rate: Average of student participation in online discussions/quizzes out of total discussion assignments and quizzes
(15). Avg_assignment_submission_ rate: Assignment submissions out of total assignments
(16). Avg_on-time submission_rate: On-time submissions out of total assignments
(17). Avg_late_submission_rate: Late submissions out of total assignments
(18). Avg_missing_submission_rate: Missing submissions out of total assignments
(19). Student average score: Student current overall scores (grades)
(20). Evaluation_mean_value: Instructor's evaluation scores evaluated by students

3.2. Model Construction

The dependent variable, the Evaluation_mean_value, is the indicator of instructors' performance evaluated by students. The raw data of evaluation average score is numeric values between 1 - 4, which needs to be encoded to three classes based on the historical experience and university policies, i.e., "2" (Excellent, score >= 3.85), "1" (Good, score 3.25 – 3.85), "0" (Improve, score < 3.25). All student individual-level features need to be aggregated to the course level. For example, we calculate the average student's current score for each class, the mean value of student on-time assignment submission rate for all students and all assignments in a particular class, the mean value of instructor grading days for all assignments in a specific class, and so on. In addition, we encode Faculty ID and Course Section ID using the hash encode method to remove identity information. Further, data quality checks and preprocessing are conducted for all 20 variables using Python libraries, including Pandas, Numpy, Matplotlib, and Seaborn. The process includes univariate analysis to identify and remove records with outliers, missing values, and errors. This analysis provides summary statistics like mean, median, mode, variance, and standard deviation, offering a concise data overview. Univariate analysis aids in identifying independent variables with skewed distribution, suggesting potential transformations. The dataset is split into two subsets: 80% as training data and 20% as testing data.

Variable selection is a crucial step in building a predictive model. Focusing on the most influential and informative variables can reduce the complexity of the model, mitigate the risk of overfitting, and enhance model interpretability. We get four sets of variables using a wrapper-based variable selection method – Logistic Regression-based Sequential Forward Selection (SFS). Then, we compare and analyze the performance of different machine learning models on these four sets of selected variables.

We fit four sets of datasets to eight machine learning algorithms to construct the prediction model: logistic regression (LR), support vector machine (SVM), decision tree (DT), Naïve Bayesian (NB), Multi-Layer Perceptron Neural Network (MLPNN), Random Forest (RF), Gradient Boost Machine (GBM), and XG Boost (XGB). Moreover, we tune the parameters iteratively till we get an optimized result for each model.

3.3. Evaluation

We evaluate the prediction models using standard metrics of accuracy, precision, recall, F1-score, and time cost. Firstly, we compare and analyze the performance metrics. Then, we use cross-validation and SMOTE techniques to improve the model performance. Finally, we suggest a good solution.

4. Results

All exploration jobs, such as variable selection, model training, testing, evaluation, etc., are completed in Jupyter Notebook (an open-source web application that allows users to create and share documents containing live code, equations, visualizations, and narrative text) with Python 3 environment. Most popularly used machine learning-related Python libraries were imported for our work, e.g., pandas, numpy, matplotlib, seaborn, sklearn, and mlxtend. This section reports the results.

4.1. Variable Selection

As shown in Table 1, we initially extracted 19 independent variables from the LMS platform based on the educational literature and institutional experience. Then, we used a wrapper-based (wrapper methods evaluate feature subsets by directly training and testing a machine learning model) variable selection method to obtain four sets of variables: 14, 12, 10, and 8 variables were selected.

4.2. Machine Learning Model Performance Comparison

We compare the performance of eight machine learning models with four sets of independent variables. We also compare the performance of four models with the cross-validation method on eight independent variables.

As shown in Table 2, when we fit the dataset with 14 variables, the RF model has the best performance (0.71 Accuracy / 0.69 F1-Score), followed by GBM (0.71/0.68), XGB and MLPNN (0.70/0.68), and LR (0.67/0.63). For computing time, MLPNN consumes more time than others (20.5425 seconds), followed by GBM and RF.

By fitting the dataset with 12 variables, the RF model shows the best performance of 0.71 accuracy and 0.69 F1-Score, followed by GBM, XGB, and MLPNN. MLPNN takes 20.3255 seconds for the time consuming, followed by GBM and RF. Results are shown in Table 3.

Table 4 shows the models' performance on ten variables selected from the wrapper-based method. RF gets the highest accuracy (0.73 Accuracy / 0.70 F1-Scroe), followed by GBM, XGB, MLPNN, and DT. For the cost, MLPNN uses 20.1258 seconds, followed by GBM (1.2713 seconds) and RF (0.3863 seconds).
Table 5 shows the models' performance on eight variables. RF has the highest accuracy

(0.74) and highest F1-Score (0.73), followed by GBM, XGB, and MLPNN. For the cost,

MLPNN consumes 19.8566 seconds to output the result, followed by GBM and RF.

#	Independent Variables from LMS	14 Selected Variables	12 Selected Variables	10 Selected Variables	8 Selected Variables
1	Course_id_hash_encoded	√	√		
2	Instructor total_activity_time	√	√	√	√
3	Instructor page_views				
4	Published_on_time				
5	Mean_grading_days	√	√	√	√
6	Total_student_enrolled	√	√	√	√
7	Dropout_rate	√			√
8	Retention_rate		√	√	
9	Evaluation total_responders	√			
10	Evaluation total_responded	√			
11	Total_response_rate		√	√	√
12	Student average_activity_time	√	√	√	√
13	Student average_page_views	√	√	√	
14	Student avg_participation_rate	√	√		
15	Avg_assignment_submission_ rate	√			
16	Avg_on-time submission_rate	√	√	√	
17	Avg_late_submission_rate	√	√	√	√
18	Avg_missing_submission_rate				
19	Student_avg_score	√	√	√	√

Table 1: The Selected Independent Variables

ML Model	Accuracy	Precision	Recall	F1-Score	Computing time (sec.)
Logistic Regression	0.67	0.61	0.67	0.63	0.0165
Decision Tree	0.58	0.60	0.58	0.59	0.0163
SVM	0.66	0.60	0.66	0.60	0.0903
Naïve Bayes	0.63	0.62	0.63	0.62	0.0040
MLPNN	0.70	0.69	0.70	0.68	20.5425
Random Forest	0.71	0.70	0.71	0.69	0.3695
GBM	0.71	0.69	0.71	0.68	1.6986
XGB	0.70	0.68	0.70	0.68	0.2624

Table 2**:** Machine Learning Models' Performance on Fourteen Variables

We compare the performance of four machine learning models with eight independent variables using 5-fold and 10-fold cross-validation train processes. Results are shown in Table 6. LR performs best in 5-fold and 10-fold cross-validation train processes, followed by SVM, RF, and DT.

ML Model	Accuracy	Precision	Recall	F1-Score	Computing time (sec.)
Logistic Regression	0.67	0.61	0.67	0.63	0.0134
Decision Tree	0.56	0.60	0.56	0.58	0.0200
SVM	0.66	0.59	0.66	0.59	0.0815
Naïve Bayes	0.64	0.61	0.64	0.62	0.0044
MLPNN	0.68	0.68	0.68	0.67	20.3255
Random Forest	0.71	0.71	0.70	0.69	0.4084
GBM	0.71	0.67	0.71	0.68	1.5744
XGB	0.70	0.67	0.70	0.68	0.3122

Table 3: Machine Learning Models' Performance on Twelve Variables

ML Model	Accuracy	Precision	Recall	F1-Score	Computing time (sec.)
Logistic Regression	0.67	0.61	0.67	0.62	0.0246
Decision Tree	0.63	0.65	0.63	0.64	0.0099
SVM	0.65	0.42	0.65	0.51	0.0725
Naïve Bayes	0.64	0.61	0.64	0.62	0.0070
MLPNN	0.68	0.68	0.67	0.66	20.1258
Random Forest	0.73	0.74	0.73	0.70	0.3863
GBM	0.69	0.67	0.69	0.67	1.2713
XGB	0.69	0.67	0.68	0.66	0.2003

Table 4: Machine Learning Models' Performance on Ten Variables

ML Model	Accuracy	Precision	Recall	F1-Score	Computing time (sec.)
Logistic Regression	0.67	0.61	0.67	0.62	0.0081
Decision Tree	0.60	0.61	0.60	0.60	0.0121
SVM	0.65	0.42	0.65	0.51	0.0665
Naïve Bayes	0.66	0.63	0.66	0.64	0.0048
MLPNN	0.68	0.69	0.68	0.67	19.8566
Random Forest	0.74	0.74	0.74	0.73	0.3958
GBM	0.70	0.68	0.70	0.68	0.9755
XGB	0.70	0.67	0.70	0.68	0.0201

Table 5: Machine Learning Models' Performance on Eight Variables

cross-validation	5-fold			10-fold		
ML Model	**Mean Cross-Validation Score - Accuracy**	**Standard Deviation of Cross-Validation Scores**	**Computing Time (sec.)**	**Mean Cross-Validation Score - Accuracy**	**Standard Deviation of Cross-Validation Scores**	**Computing Time (sec.)**
Logistic Regression	0.63	0.02	0.0330	0.64	0.05	0.0699
Decision Tree	0.51	0.01	0.0495	0.52	0.04	0.0986
SVM	0.62	0.01	0.2955	0.63	0.03	0.6925
Random Forest	0.62	0.02	1.5208	0.63	0.06	3.2927

Table 6: Machine Learning Models' Performance Using Cross-Validation Method

5. Discussion

Since the target variable is unevenly distributed, we use the F1-score as the primary performance evaluation metric to address the research questions.

To answer our first research question, how variable selection affects prediction performance, we conclude the following from the abovementioned comparison:

- As shown in Figure 2, the light blue line represents the set of 14 variables. It shows the highest average F1-Score. Moreover, it has promising results on LR, MLPNN, RF, GBM, and XGB. The orange line (12 variables) performs similarly to the 14-variable line. The grey line (10 variables) shows the higher performance on DT and lower performance on SVM compared to 14 and 12 variables and similar performance on other models. Notably, the yellow line (8 variables) shows the worst performance on the SVM model. However, it gets the best results on NB and RF models.
- Variable selection can improve cost-efficiency: fewer variables consume less computing time for all machine learning models.
- Different sets of variables fit different machine learning algorithms; no one selected set is the best for all.
- Conduct variable selection iteratively to balance the trade-off between the number of features selected and model performance.

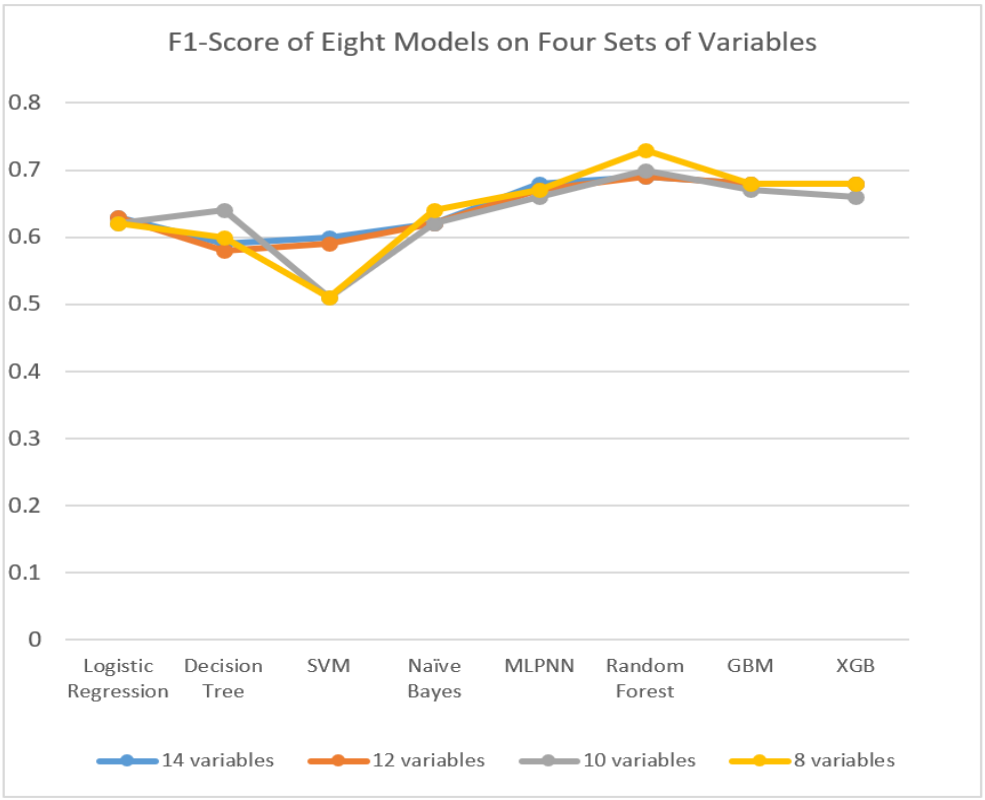


Figure 2: F1- Score of Eight Models on Four Sets of Variables

For the second research question, what machine learning method is the best suggestion for online teaching performance prediction, we summarize from the above comparison:

- As shown in Figure 2, RF achieves the highest performance with all different variable sets. Moreover, it reaches the highest F1-score (0.73) with eight selected variables.
- All three ensemble algorithms outperform other non-ensemble methods except MLPNN.
- MLPNN outperforms all other non-ensemble models and gets the same average F1-Score as GBM and XGB models. However, it consumes the most computing time of all machine learning models.
- SVM is the most popular used algorithm in solving classification problems. However, it performs worst in our work with the lowest average F1-Score.
- The Naïve Bayes model consumes the least computing time of all tested models.

In addition, when testing cross-validation on eight independent variables:

- All models need to use more training time than the non-cross-validation process.
- None of the models show a significant improvement in performance other than the non-cross-validation process.
- All models use about doubled time on 10-fold cross-validation rather than 5-fold.
- All machine learning methods show a tiny performance improvement but higher standard deviations from 5-fold to 10-fold.
- RF shows non-advantages of ensemble algorithms on cross-validation.

Due to the significant imbalance of the target variable, we try to use the Synthetic Minority Over-sampling Technique (SMOTE) to improve the performance of the RF model with eight variables, which is the best model from the above comparison. SMOTE can help mitigate the imbalance by oversampling the minority class and generating synthetic instances to balance the class distribution, which prevents the model from being biased toward the majority class and enhances its ability to recognize patterns in the minority class (Azhar et al., 2023). Further, SMOTE introduces additional instances of the minority class, providing more diverse examples for the model to learn from and reducing the risk of overfitting (Azhar et al., 2023). The dataset is about doubled using the SMOTE technique from 1864 to 3495 data items. Figure 3 shows the performance of the RF model using SMOTE. The F1-Score achieves 0.83, a 14% improvement compared to the previous 0.73 F1-Score.

```
Accuracy: 0.8283261802575107
Time:  1.0572922229766846
              precision    recall  f1-score   support

           0       0.84      0.98      0.90       217
           1       0.84      0.73      0.78       249
           2       0.81      0.79      0.80       233

    accuracy                           0.83       699
   macro avg       0.83      0.83      0.83       699
weighted avg       0.83      0.83      0.83       699
```

Figure 3: Performance of the RF Model Using SMOTE

Our suggested prediction model is the RF on eight variable sets. This model achieved 73% F1-Score initially and 83% after using SMOTE, which is reasonable but still slightly lower than expected. This may be caused by the small dataset which includes only 1,864 records with 19 features. According to the literature, we may need more than 5,000 data items. For example, Abunasser et al. (2022) achieved better performance using a dataset including 5820 with 33 features. Moreover, the low performance may be from the imbalanced dataset, inappropriate model tunning parameters, low informative variable selections, etc. These issues may lead to the ineffectiveness of the cross-validation training process, which will be turned into our future work for this study.

6. Conclusion

Compared to existing studies on predicting online instructors' performance, our study is outstanding in three aspects. First, this study focuses on using LMS activity and behavioral data instead of perception data. Second, we present how different variable sets (via variable selection methods) impact the performance prediction in our experiment. Third, we train, test, and compare eight machine learning models using the dataset with different variable sets. We suggest the best solution for satisfying the research objectives - the RF model on eight variables with the SMOTE technique. Theoretically, this work enriches the research of applying machine learning to the learning analytics field, especially in online instructors' performance prediction with LMS data. The methodological contribution can be generalized to the context with similar characteristics. Practically, the artifact of the model construction methods including data

preprocessing, feature selection, model tuning, and imbalance data handling techniques, as well as a proposed optimal model with a specific variable set that can be used for online instructors' performance prediction effectively and efficiently. It can help the early intervention during the learning and teaching process and improve overall educational performance.

References

Abunasser, B. S., AL-Hiealy, M. R. J., Barhoom, A. M., Almasri, A. R., & Abu-Naser, S. S. (2022). Prediction of Instructor Performance using Machine and Deep Learning Techniques. *International Journal of Advanced Computer Science and Applications*, *13*(7), 78-83.

Agaoglu, M. (2016). Predicting Instructor Performance Using Data Mining Techniques in Higher Education. *IEEE Access*, *4*, 2379-2387.

Ahajjam, T., Moutaib, M., Aissa, H., Azrour, M., Farhaoui, Y., & Fattah, M. (2022). Predicting Students' Final Performance Using Artificial Neural Networks. *Big Data Mining and Analytics*, *5*(4), 294-301.

Ahmadi, F., & Ahmad, M. E. S. (2013). Data Mining in Teacher Evaluation System using WEKA. *International Journal of Computer Applications*, *63*(10), 12-18.

Ahmed, A. M., Rizaner, A., & Ulusoy, A. H. (2016). Using Data Mining to Predict Instructor Performance. *Procedia Computer Science*, *102*, 137-142.

Azhar, N. A., Pozi, M. S. M., Din, A. M., & Jatowt, A. (2023). An Investigation of SMOTE Based Methods for Imbalanced Datasets With Data Complexity Analysis. *IEEE Transactions on Knowledge and Data Engineering*, *35*(07), 6651-6672.

Chandna, R., Saini, S., & Kumar, S. (2021). Fuzzy AHP based performance evaluation of massive online courses provider for online learners. *Materials Today: Proceedings*, *46*, 11103-11112.

Costa, L. A., Pereira Sanches, L. M., Rocha Amorim, R. J., Nascimento Salvador, L. d., & Santos Souza, M. V. d. (2020). Monitoring Academic Performance Based on Learning Analytics and Ontology: A Systematic Review. *Informatics in Education*, *19*(3), 361-397.

Jiao, P., Ouyang, F., Zhang, Q., & Alavi, A. H. (2022). Artificial intelligence-enabled prediction model of student academic performance in online engineering education. *Artificial Intelligence Review*, *55*(8), 6321-6344.

Karimi, H., Derr, T., Huang, J., & Tang, J. (2020, 2020/07/00). Online Academic Course Performance Prediction Using Relational Graph Convolutional Neural Network. *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, 444-450.

Kathiroli, P., Vijayalakshmi, V., Panimalar, K., & Janarthanan, S. (2020). Predicting the performance of instructors using Machine learning algorithms. *High Technology Letters*, *26*, 694-705.

Khan, I., Ahmad, A. R., Jabeur, N., & Mahdi, M. N. (2021). An artificial intelligence approach to monitor student performance and devise preventive measures. *Smart Learning Environments*, *8*(1), 1-18.

Li, M., & Su, Y. (2020). Evaluation of Online Teaching Quality of Basic Education Based on Artificial Intelligence. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(16), 147-161.

Martínez-Caro, E., Cegarra-Navarro, J. G., & Cepeda-Carrión, G. (2015). An application of the performance-evaluation model for e-learning quality in higher education. *Total Quality Management & Business Excellence*, *26*(5/6), 632-647.

Mi, H., Gao, Z., Zhang, Q., & Zheng, Y. (2022). Research on Constructing Online Learning Performance Prediction Model Combining Feature Selection and Neural Network. *International Journal of Emerging Technologies in Learning*, *17*(7), 94-111.

Naser, S. A., Zaqout, I., Ghosh, M. A., Atallah, R., & Alajrami, E. (2015). Predicting Student Performance Using Artificial Neural Network: In the Faculty of Engineering and Information Technology. *International journal of hybrid information technology*, *8*(2), 221-228.

Pan, W.-T., Huang, C.-E., & Chiu, C.-L. (2016). Study on the performance evaluation of online teaching using the quantile regression analysis and artificial neural network. *Journal of Supercomputing*, *72*(3), 789-803.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105.

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.

Raschka, S. (2015). *Python Machine Learning, 1st Edition*. Packt Publishing.

Tian, Y., Shi, Y., & Liu, X. (2012). Recent advances on support vector machines research. *Technological and Economic Development of Economy*, *18*(1), 5-33.

Wang, W., Zhuang, H., Zhou, M., Liu, H., & Li, B. (2020). What Makes a Star Teacher? A Hierarchical BERT Model for Evaluating Teacher's Performance in Online Education. *arXiv preprint arXiv:2012.01633.*

Yang, Y. (2021). The Evaluation of Online Education Course Performance Using Decision Tree Mining Algorithm. *Complexity*, *2021*, 1-13.

Ye, J., Isaeva, E., & Rocha, Á. (2020). Modeling of performance evaluation of educational information based on big data deep learning and cloud platform. *Journal of Intelligent & Fuzzy Systems*, *38*(6), 7155-7165.

Yu, C.-M., Chang, H.-T., & Chen, K.-S. (2018). Developing a performance evaluation matrix to enhance the learner satisfaction of an e-learning system. *Total Quality Management & Business Excellence*, *29*(7/8), 727-745.

Yu, J. (2021). Academic Performance Prediction Method of Online Education using Random Forest Algorithm and Artificial Intelligence Methods. *International Journal of Emerging Technologies in Learning*, *15*(5), 45-57.

6. Artificial Intelligence Meets Seafood Supply Chain Management

Behrang Ijadi
University of Canterbury
behrang.ijadi@pg.canterbury.ac.nz

Stephen C. Wingreen
University of Canterbury
stephen.wingreen@canterbury.ac.nz

Pedro M. Reyes
Baylor University
pedro_reyes@baylor.edu

Abstract

Seafood supply chains (SSCs) face intricate challenges throughout the food journey, from harvest locations to our tables. The complexity demands a holistic view of supply chain management (SCM) to ensure the interests of each respective entity within the chain are met. The perishable nature of seafood products complicates the proper distribution efforts by industry owners. They are grappling with mismanagement and inefficient operational logistics strategies, which lead to lower profits and hinder the goal of meeting market demands. Amidst these challenges, emerging technologies offer promising solutions for decision-makers to control the flow of physical goods in their businesses. Using the Design Science Research (DSR) method, this study developed an artifact to highlight the potential of artificial intelligence (AI) to be engaged in decision-making scenarios. In this regard, a mathematical model (MM) is developed to deepen the understanding of the SSC environment and make it possible to utilize novel machine learning (ML) techniques in order to help authorities make better decisions on their logistics operations, considering maximizing profits in their businesses.

Keywords: Seafood, Supply chain management (SCM), Mathematical modeling, Artificial intelligence (AI), Machine learning (ML).

1. Introduction

In today's complex world, the seafood supply chain (SSC) has a different story. Since the 1950s, global seafood consumption has risen significantly, reaching almost 186.6 million metric tons in 2023 (Statista, 2024). Fish consumption increased by 3.1% on average annually between 1961 and 2017, outpacing population growth (1.6 percent) (FAO, 2020). While increasing demand for fishery and aquaculture products has made seafood the most widely traded animal protein that can feed human beings (Rabobank, 2022), world hunger is on the rise (9.8% in 2021) (FAO et al., 2022), considering 20% of seafood spoils before reaching its end consumer (Ray, 2019).

Although such statistics can demonstrate the significance and crucial role of seafood products in providing people and the next generation with food and nutrition worldwide, seafood is a highly perishable product with a short shelf life. SSC actors are vulnerable to weather impacts and cold chain infrastructure conditions as long as they hold or move seafood product inventory

(Pramod et al., 2014), and operations mismanagement can make the situations more complicated. This is where the effective management of the SSC comes into play. To achieve a high level of food safety that meets the daily needs of involved parties, measures should be taken throughout the whole supply chain (SC), which will ultimately reflect in costs. Such financial barriers may occasionally compel suppliers to act inappropriately or make wrong decisions (Ray, 2019). If companies want to survive in a competitive industry ecosystem, they need to look at the existing opportunities for improving their SCs from a logistics operations perspective. This will lead to substantial and consistent improvements in the performance of the company, particularly in terms of operational efficiency and increasing revenue (Rai et al., 2006).

This research tries to highlight the power of AI in shaping an agent-based model that can interact with the physical side of the SSC and making autonomous decisions for the whole system and improve managerial decisions towards increasing seafood producers' income who are inclined to use such technologies. To achieve this goal, a MM needs to be made to depict the SSC environment. It is a must for a smart ML agent who wants to interact with the environment, explore it, collect data, take actions considering its optimal policy, and learn from all of them. This is a novel technique in the realm of ML that be called Deep Reinforcement Learning (DRL) which uses a combination of Deep Learning (DL) and Reinforcement Learning (RL) (Mnih et al., 2015).

1.1. Literature Review

New technological advances are paving the way for the effective management of SCs (Shekarian et al., 2022). Improving operations management involves creating more intelligent supply chains to control all procedures. While the market demand for seafood products is changing due to the improvement in peoples living standards (Ray, 2019; Wang et al., 2019) and people pay for the healthiest options (Jahanbin et al., 2023; Rahmaniya & Sekharan, 2018), seafood producers expect to increase their profits. Although the contrast between food producers' benefits and food safety has recently received much interest in the agricultural sector (Chen et al., 2021; Fan, 2019; Himmelstein et al., 2017; Lin et al., 2019; Toledo-Hernández et al., 2020), little has been done in seafood products. In addition, mismanagement and inefficient strategies among different parts of a SC such as production, processing, storage, distribution, and retail may lead to lower profits for involved parties (Lin et al., 2019). It will have a negative impact on their revenue when they cannot properly allocate their production level or storage capacity based on their market demand (Chen et al., 2021). In practice, continuously making decisions for profit optimization in a complex and dynamic environment within a SC can be an extremely challenging task, as various factors can impose constraints on its normal operations (Castro & Jaimes, 2017). In this scenario, decision-makers can concentrate on optimizing their SCs and increasing their revenues by controlling the quantity of production and storage of their products while taking consumers' priorities into account (Chen et al., 2021).

This section provides the related literature on the SC optimization context. There are lots of MMs to solve SC optimization problems most of them are based on heuristic (Dwivedi et al., 2020; Kocaoglu et al., 2020; Samadi et al., 2018; Shongwe & Bezuidenhout, 2019), game theory (Halat & Hafezalkotob, 2019; Raj et al., 2018; Vasnani et al., 2019), and control theory

(Wu & Chen, 2019; Zhao & Wang, 2018) that are commonly developed for specific applications in static environments without uncertainty and might be not fit in dynamic or complex environments like real-world SCs (Hubbs, Perez, et al., 2020). With advances in ML, especially in its branch known as reinforcement learning (RL) (Sutton & Barto, 2018), there is a capability to train a ML model to make better decisions (Hubbs, Li, et al., 2020). While the existing RL-based methods that employ traditional value-based RL algorithms like Q-learning (Habib et al., 2016; Kemmer et al., 2018; Sutton & Barto, 2018) can address SCM issues, they may face challenges in effectively handling large scale problems with highly dimensional state spaces. It is primarily due to the computational and storage requirements associated with computing and storing Q-values for each action across all states. In order to handle this issue, deep neural networks (DNNs) were combined to form deep Q-networks (DQN) (Goodfellow et al., 2018; Mnih et al., 2015; Oroojlooyjadid et al., 2017; Oroojlooyjadid et al., 2020). Considering the advantages offered by the DQN algorithm, Deep Reinforcement Learning (DRL) holds significant potential for dealing with SCM issues which has also made its way into this study.

1.2 Research Objectives and Research Questions

From a systems perspective, this study encompasses three interconnected elements: technology, people, and process, all of which require careful consideration. Within the realm of technology, research solutions will leverage emerging technologies like AI to manage the flow physical goods. From the process standpoint, MMs will define the real-world seafood industry environment, its existing variables, interactions, and constraints. On the people front, it includes all seafood industry people, and their effective managerial efforts that could be utilized to incentivize them. However, there are still several gaps on having a holistic view to SSCs while using emerging technologies or ML techniques together in SSC context. So, the primary objectives of this research are as follows; (i) understanding the seafood industry and its environment, (ii) developing an artifact based on emerging technologies in SSCs considering priority concerns of the industry people and their income issues, (iii) and providing recommendations to decision-makers. To move the DSR agenda forward, an integrated approach towards developing artifact for these objectives has been explored to solve the mismanagement issues in the SSC and contribute to the optimization of the model based on MM and DRL. To achieve research purposes, research questions are as follows:

Q1: How can a SSC be modeled based on mathematical programming, AI and ML?

Q2: How does the DRL would contribute to the effective operation management of a SSC?

2. Research Methodology

Design Science Research (DSR), which has been presented as the research methodology for this study, develops the problem and answers the research questions using novel mechanisms based on guidelines provided by some scholars (Hevner & Chatterjee, 2010; Hevner et al., 2004), as has been described in the following. They asserted that DSR's significance in Information Systems (IS) lies in its ability to identify novel IT capabilities that arise as a result of the extension of those systems into other fields. The design includes a process (activity) and a product (artifact), a verb, and a noun (Walls et al., 1992). It shows how the world is perceived (artifacts) as well as how it is acted upon (processes). A targeted IT artifact created to address a significant organizational issue is the result of DSR in IS (Hevner et al., 2004). Constructs, models, methods, and instantiations are just a few of the different shapes that the DSR-created

artifact can take (March & Smith, 1995). So, any designed object with a built-in answer to a known research topic is considered a DSR artifact (Peffers et al., 2007).

This study utilizes a structured process of DSR for developing and evaluating designs (Ågerfalk & Karlsson, 2020; Hevner et al., 2004; Jones & Gregor, 2007; Niederman & March, 2012), specially the Peffers et al. (2007) iterative approach. However, this is in-progress research. So, it has resulted in an artifact that handle management-related issues of seafood companies using emerging technologies in their SCs. Since MMs are particularly useful research tools when researchers wish to understand and describe an issue by variables, constraints, and models, the artifact of this study is a MM for illustrating interactions among real-world SSC members (March & Smith, 1995). By selecting a MM as the research artifact, there is a need for developing a simplified prototype for gathering feedback and adjustments as needed. In this regard, this study collects relevant concepts and variables from previous surveys, working papers, books, and research reports as well as interviews and field research to better represent decision variables and constraints of the MM for making ML model. Data standards related to SSC has been collected as real-world data settings which include production level, storage capacity, shipping costs, market demand, etc. So, the scope of this research includes the New Zealand seafood industry according to its environment and development.

3. Research Artifact

The first step before being engaged in MM to design the environment for the SSC is to become familiar with its basic foundations. RL is a mathematical framework for solving sequential decision-making issues in a variety of domains without needing an explicit environment model (Nazari, 2019). It defines an agent interacts with the environment while learning an optimal policy through trial and error (Bertsekas, 2012; Gijsbrechts et al., 2022; Li, 2017; Sutton & Barto, 2018). It has been a long time since neural networks and RL have been combined (Schmidhuber, 2015; Sutton & Barto, 2018). RL would be a promising research subject, especially with the combination of DNNs and RL, known as DRL or deep RL, which gains benefits from big data, advanced analytics, strong computational power, cutting-edge algorithms, developed software, powerful architectures, as well as significant financial support (Goodfellow et al., 2016; Krakovsky, 2016; LeCun et al., 2015; Li, 2017; Reyes et al., 2020). The availability of huge processing power and recent algorithmic advancements have spurred a new interest in DRL, despite the fact that both RL and DRL have been extensively explored in the Computer Science and Operations Research (OR) literature (Gijsbrechts et al., 2022; Mnih et al., 2015). Several of the existing RL libraries, such as OpenAI Gym, present numerous challenges that may not be directly applicable in an industrial context (Hubbs, Perez, et al., 2020). So, NumPy and OR-Gym would be the other options. The OR-Gym package offers simple customization through configuration dictionaries, allowing researchers in both OR and RL to easily utilize the library for their specific needs (Hubbs, Perez, et al., 2020). To model the problem, it is necessary to present the mathematical foundation of what is mentioned above based on Snyder and Shen (2019).

To employ the AI-based method for developing the research artifact, a DRL-based model has been developed for the SSC environment. Due to the similarity between the gameplay and SCM environment, DRL method can be benchmarked in SC scenarios (Nazari, 2019; Nazari et al., 2018; Oroojlooyjadid et al., 2022) as both environments involve complex decision-

making processes, resource allocation, and the pursuit of specific objectives. DRL has also made its way into SCM as a result of the recent accomplishments it has had in identifying winning strategies in (computer) games, which helped it acquire popularity (De Moor et al., 2022). So, there is a simplified real-world scenario for the seafood SCM network as the following, denoted by different variables, constraints, and an objective function. To begin with, Table 1 presents a summary of the variables involved in the mathematical model and their respective definitions;

Variable	Definition
$L = [l_i; \forall i = 0, 1, \ldots, n]$	Set of all nodes in the model includes seafood factory and n retailers
t	Time period
T	The number of time period
TR	Total revenue gains from selling seafood products to customers at retailers
TC	Total cost
NP	Net profit
p_0	Unit price of seafood products
D_i	Total seafood demand at retailer i
D_t	Retailers' demands at the time t
$d_i(t)$	Seafood demand for retailer i at time period t
d_0	Base line or average seafood demand at a retailer
d_m	Maximum potential seafood demand at a retailer
$\varepsilon(i,t)$	Noise term
$AD(i,t)$	Additional demand
$\varepsilon_0(i,t) \in (0,1)$	Random value
CP	Cost of seafood production at the seafood factory
c_p	Production cost per unit of seafood at the seafood factory
$a_0 \in [0,\alpha]$	The amount of production for the seafood factory
α	Maximum production capacity at the seafood factory
a_i	The amount of seafood being shipped to retailer i
$s_0 \in [0,\beta_0]$	Stock level at the factory
β_0	Maximum storage capacity at the factory
$s_i \in [0,\beta_i]$	Stock level at retailer i
β_i	Maximum stock capacity at retailer i
CS	Cost of storage for seafood at the factory and retailers
cs_0	Unit cost of storage at the factory
cs_i	Unit cost of storage at retailer i
CF	Cost as a fine for dissatisfying demands of seafood products
cf_0	Unit fine cost for dissatisfying demand at retailer i
CT	Transportation cost
ct_i	Shipping cost of a truck for transporting seafood products to retailer i
$v_i \in (0,v]$	Capacity of a truck for shipping seafood products to retailer i
v	Maximum capacity of a truck
S_t	State space
A_t	Action space
R_t	Reward function

Table 1. Variables and Corresponding Definitions Used in Mathematical Model

It consists of one processor as the seafood factory and multiple retailers which are represented by l_0 and (l_1, l_2, ..., l_n) respectively in a certain number of periods. So, $L = (l_0, l_1, l_2, ..., l_n)$ can represent the integration of all these nodes into a single set. The factory ships seafood products to retailers via distributors by truck. The amount of seafood to be produced, kept in storage at the factory, and transported to retailers would all be determined during each period. Now, there are some constraints for seafood demand and storage that should be formulated. There is a seasonal demand for seafood products for each retailer. A penalty will be applied until the

demand is met if the retailer is unable to satisfy it. Some limitations on the factory's capability to produce and store seafood products, as well as other restrictions on retailers' storage space, bring the issue closer to the actual world. Furthermore, we anticipate that the demand for seafood could surpass the factory's ability to produce. As a result, retailers should have an adequate supply of seafood in stock. To tackle this issue, both the factory and retailers must possess the capability to efficiently replenish their stock in accordance with the seafood demands.

The total revenue generated from the sale of seafood at retailers is determined by $TR = p_0 \times \sum_{i=1}^{n} D_i$. In this equation, p_0 represents the price per unit of seafood, $D_i = \sum_{t=1}^{T} d_i(t)$ represents the aggregate retailer's seafood demand, while T denotes the number of time periods. In particular, take into account the timeframe for managing the SSC as a one year. A 12-month timeframe provides a regular and consistent interval for analyzing demand patterns which simplifies the complexity of the modelling process and its implementation. This is a pilot timeframe for this model that can be extended for longer periods. In addition, the use of vacuum packing and cold storage through modern seafood packaging ensures its freshness for a duration exceeding 12 months. These methods are known for effectively preserving the freshness and quality of seafood products for extended periods, often surpassing 12 months (Kontominas et al., 2021). Consequently, the shelf life of the seafood becomes essentially negligible in this context which implies that the primary concern shifts from managing perishability to optimizing other aspects of the supply chain operations management, such as production, distribution, and demand fulfilment rather than solely focusing on mitigating spoilage and wastage due to product expiration. As a result, the shelf life of the seafood becomes insignificant in this scenario. As previously stated, seafood experiences seasonal demand across various retailers, each with distinct potential values. Therefore, the seafood demand for each retailer during a specific time period is represented as $d_i(t) = \frac{d_m}{2} \times (1 + sin(\frac{(2i+t) \times \pi}{6})) + \varepsilon(i,t)$ where d_m denotes the predetermined maximum potential demand at a retailer or the highest level of seafood products that can be demanded by a retailer.

The first part of the seasonal demand function denoted by $\frac{d_m}{2}$ represents the average or baseline demand for the seafood product which is a constant value that indicates the typical demand level. Multiplying it the sine term scales the amplitude of the seasonal pattern to half of the maximum demand. The second part introduces a seasonal pattern to the demand. The *sin* function calculates the sine of the given angle. By varying the angle $\frac{(2i+t) \times \pi}{6}$, the demand function creates a cyclical pattern over the course of 12 months. The $(2i+t)$ term introduces variations based on the retailer number i and the month t, influencing the amplitude and phase of the seasonal pattern. The last part of the seasonal demand function is the noise term that is denoted by $\varepsilon(i,t) = \varepsilon_0(i,t) \times AD(i,t)$. In this equation $\varepsilon_0(i,t)$ represents the stochastic factor, specific to retailer i in month t, which is allocated a random value ranging from 0 to 1. Also, $AD(i,t)$ represents the additional demand throughout various time periods for a retailer. To simulate the unexpected rise in demand, the noise factor is applied based on the stochastic nature or random variation in demand. It can introduce random fluctuations around the seasonal pattern and base demand. Besides, it can simulate various factors that affect demand, such as retailer preferences, external events, or market conditions. So, this demand function calculates the overall demand for a specific retailer i in a particular month t for seafood products. The

function output reflects the total demand value, considering both the seasonal variations and the random fluctuations introduced by the stochastic factor. Thus, mathematical representation of total revenue can be summarized in Equation 1:

$$TR = p_0 \times \sum_{i=1}^{n} \sum_{t=1}^{T} [\frac{d_m}{2} \times (1 + sin\ (\frac{(2i+t) \times \pi}{6})) + \ \varepsilon_0(i,t) \times AD(i,t))] \quad (1)$$

Moving forward, the production cost of seafood is determined by $CP = c_p \times a_0$. In this equation, c_p is the cost per unit for producing seafood, while a_0 is the factory's production level. Subsequently, the cost of transporting seafood is stated as $CT = \sum_{i=1}^{n} (ct_i \times \left(\frac{a_i}{v_i}\right))$, where ct_i represents the cost of transporting seafood to a retailer via truck and v_i signifies the truck's capacity for delivering seafood to a retailer which can be predefined based on the type of trucks. The value a_i represents the volume of seafood being transported to a retailer, and it is crucial to ensure that the total volume being transported to all retailers remains below the factory's stock level. Furthermore, the storage cost of seafood is denoted as $CS = \sum_{i=0}^{n} (cs_i \times s_i)$, where cs_i represents the seafood storing cost, and s_i signifies amount of stock at either the factory (i = 0) or retailers (i = 1, 2, …, n). In addition, there is a penalty cost incurred once a retailer fails to satisfy the demand. This penalty cost, denoted as $CF = cf_0 \times \sum_{i=1}^{n} (D_i - s_i)$, is determined by the unit punishment cost, cf_0. Now, the total cost can be considered as the sum of costs and denoted by $TC = CP + CT + CS + CF$, that can be formulated in Equation 2:

$$TC = \sum_{i=1}^{n} [ct_i \times \left(\frac{a_i}{v_i}\right) + \left(cs_i - cf_0\right) \times s_i + \sum_{t=1}^{T} cf_0 \times (\frac{d_m}{2} \times (1 + sin\ (\frac{(2i+t) \times \pi}{6})) + \varepsilon_0(i,t) \times AD(i,t))] + (c_p \times a_0 + cs_0 \times s_0) \quad (2)$$

Hence, the net profits resulting from the production and sale of seafood products within this SC can be computed as $NP = TR - TC$. This net profit is the objective function of the problem that should be maximized. Therefore, according to the above mathematical formulation, it can be represented in Equation 3:

$$NP = \sum_{i=1}^{n} \{[\ \sum_{t=1}^{T} (p_0 - \ cf_0) \times (\frac{d_m}{2} \times (1 + sin\ (\frac{(2i+t) \times \pi}{6})) + \ \varepsilon_0(i,t) \times AD(i,t))] - [ct_i \times \left(\frac{a_i}{v_i}\right) + \left(cs_i - cf_0\right) \times s_i]\} - (c_p \times a_0 + cs_0 \times s_0) \quad (3)$$

Once the SSC environment is formulated, the next step is to incorporate it into a DRL framework. In this research DRL agent can learn to allocate resources efficiently by determining the optimal amount of seafood to be produced by the factory and distributed to retailers. This ensures that resources are utilized effectively. It also can dynamically adjust inventory levels at the factory and retailers and improve the overall efficiency of inventory management by reducing relevant costs which ultimately results in increased profits through optimal operations management. To achieve this, DRL components must all be adjusted for the MM. So, first, the state space must be introduced. Then, action space needs to be defined, as well as the reward function. Now, this AI-based model can be employed to effectively optimize the production level and the amount of storage of seafood products, with the primary objective of maximizing profits. In Figure 1, the DRL agent interacts with the environment to take actions, while the proposed SSCM scenario is the environment. Hence, the state space

would be a snapshot of the whole system that shows the amount of fish stock produced by a seafood factory and the amount of existing stock at retailers' warehouses at a given moment. Also, some safety stock must exist at retailers' warehouses to meet their excessive demand. At period *t*, the state space is represented as $S_t = (s_0, s_1, s_2, \ldots, s_n)$, wherein s_0 corresponds to the stock level of the factory, and $(s_1, s_2, \ldots, s_n)$ signifies the retailers' stock levels. The stock levels are subject to the constraint of the maximum storage space, β_i. Retailers' demands at the time t represented by $D_t = d_i(t)$ subject to $i = (1, 2, \ldots, n)$ for n retailers, which has been described earlier in the equation relevant to the seasonal demand. The DRL agent's action involves deciding how much seafood must be produced and kept by the factory, as well as how much seafood can be distributed to retailers to fulfill their respective demands. Hence, at period *t*, the action space is denoted as $A_t = (a_0, a_1, a_2, \ldots, a_n)$ that a_0 represents the factory's level of production subject to $a_0 = [0, 1, \ldots, \alpha \mid \alpha \in \boldsymbol{N}]$. It should be noted that the production level is constrained by the predetermined maximum level of production, α. The quantities $(a_1, a_2, \ldots, a_n)$ represent how much seafood is being sent to each retailer. It is important to ensure that the sum of these quantities does not exceed the factory's stock level, indicated by $\sum_{i=1}^{n} a_i \leq s_0$. The reward function serves as a guide for the DRL agent, enabling it to learn an optimized policy for managing the SSC operations. It incentivizes the agent to maximize rewards, ultimately aiming to improve product profits. Total profits gained from selling seafood products will be affected by different types of costs including production, storage, and transportation costs as well as punishment costs of dissatisfying demands as mentioned. Thus, at period *t*, the reward function should reflect a positive increase in net profits. It means that reward function of DRL agent is defined as the net profits or objective function of the MM, denoted as $R_t = NP$ (see, Equation 3).

4. Research Contribution

"The contribution of design-science research is the artefact itself" (Hevner et al., 2004). The main contribution of this DSR method would be the development of a MM as prototype artifact. This idea joins the OR problems and the IS perspective together. This research will make some other contributions in the field as; i) According to the literature, for the complex and dynamic SSC environment, ML techniques such as the DRL method would be a good option. In fact, while the state space is highly dimensioned, the RL method needs to apply DNNs to better address the problem. So, DRL is a good option for this research due to the high level of complexity of the problem. It has been applied to optimize fish farmers' revenues by focusing on seafood production and storage capacity according to market demand in this research. ii) A technical AI-based solution for the SSC problem from the IS perspective would be the other contribution of this research. So, this research proposes MM as the research artifact in DSR combined with ML as a novel solution to the well-known problems in the realm of SSCM. So, the contribution has been made by this research would be in the Improvement section of the DSR Contribution Matrix, where, there is Research Opportunity and Knowledge Contribution while developing new solutions for known problems (Gregor & Hevner, 2013).

As mentioned, this research is in progress. Thus, to ensure the practical applicability and effectiveness of the proposed DRL method in addressing real-world problems in alignment with the principles of DSR, comprehensive experiments will be conducted via various scenarios in the following steps: i) Gathering data sets or establishing experimental settings that could accurately simulate and represent real-world conditions of the SSC in different

timeframes. ii)Designing and implementing the DQN algorithm relevant to this research via coding for gaining simulation results. iii) Analyzing experimental results to assess the effectiveness and advantages of the DRL model over other existing state-of-the-art methods such as heuristics or other ML techniques. These steps aim to provide a thorough evaluation of the proposed DRL model and demonstrate its superiority over existing methods. This additional analysis will significantly enhance the contribution and robustness of this research, which will complete the design of the DSR's artifact and evaluate methods feasibility and performance to finalize the DSR next steps ahead. Moreover, given that the research period is confined to 12 months, it would be interesting to investigate the application of the MM to seafood supply chain operations over durations exceeding 12 months. This exploration could enhance the contribution of the model and its practical utility in real-world scenarios.

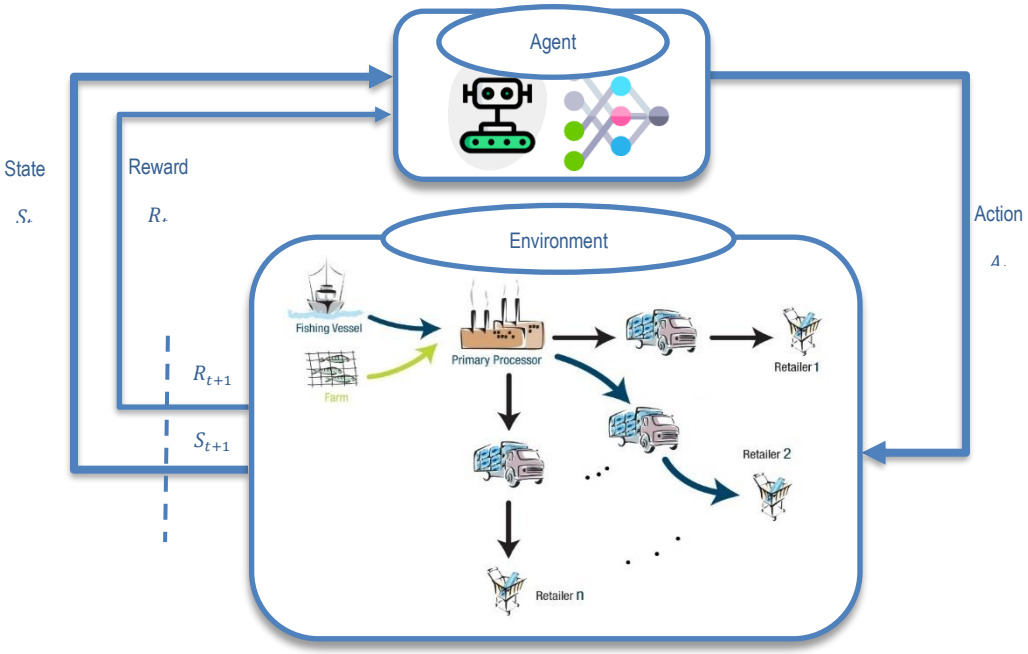


Figure 1: Developing DRL for Seafood SCM

(based on Sutton and Barto (2018)

5. Conclusion

This research aims to examine the applicability of AI in the realm of SSCM and evaluate the potential of using smart algorithms and their feasibility in real-world SC scenarios from the IS perspective. Bridging the gap between OR and IS, DSR has been conducted in this research to develop an AI-based MM as the research artifact to better understand the seafood industry in its environment. It also gives decision makers an overview of the complex structure of todays' SSC. The complexity in a seafood market increases once an inefficient system is at service along with an uncertainty in market demands. To address the problems, this study focused on the optimization and formulate the real-world SSC to a MM that would be solved via AI-based

methods. So, ML and DRL applied to cover the uncertain situations in the market demand which be resulted in an improvement in production and storage capacity through the entire SSC while answering the market demand and considering more profits for fish farmers that produce healthy food for consumers. As this study is going to be the pioneer in applying MM and DRL as the research artifact in seafood industry for effective management of its related SC, it would be an aggregate knowledgebase for additional research conducting. For example, the quick development of this IS-related paradigm would lead to the development and optimization of the artifact. Furthermore, if the proposed artifact be generally applied for better management in the SSC business environment in the future, experimental studies and tests can be done more comprehensively. Managerial authorities need to have a good perception of the application of AI-based feature services usefulness within SSCM. This perception should encompass recognition of how intelligent technologies can work together and an AI-based SC can be merged with the novel method of ML- in this research DRL, within the SSC context and be supported by applying such new technology, and how it impacts the overall SSC performance in future. It can plot an autonomous decision-making scheme for future applications of AI in real-world SCs. The results of this study not only will apply to the seafood industry but also can have a broader impact on other industries as well.

References

Ågerfalk, P. J., & Karlsson, F. (2020). Artefactual and empirical contributions in information systems research

Bertsekas, D. (2012). *Dynamic programming and optimal control: Volume I* (Vol. 1). Athena scientific.

Castro, J. A. O., & Jaimes, W. A. (2017). Dynamic impact of the structure of the supply chain of perishable foods on logistics performance and food security. *Journal of industrial engineering and management, 10*(4), 687-710.

Chen, H., Chen, Z., Lin, F., & Zhuang, P. (2021). Effective management for blockchain-based agri-food supply chains using deep reinforcement learning. *IEeE Access, 9*, 36008-36018.

De Moor, B. J., Gijsbrechts, J., & Boute, R. N. (2022). Reward shaping to improve the performance of deep reinforcement learning in perishable inventory management. *European Journal of Operational Research, 301*(2), 535-545.

Dwivedi, A., Jha, A., Prajapati, D., Sreenu, N., & Pratap, S. (2020). Meta-heuristic algorithms for solving the sustainable agro-food grain supply chain network design problem. *Modern Supply Chain Research and Applications*.

Fan, H. (2019). Theoretical basis and system establishment of China food safety intelligent supervision in the perspective of internet of things. *IEeE Access, 7*, 71686-71695.

FAO. (2020). *The State of World Fisheries and Aquaculture 2020. Sustainability in action.* (Licence: CC BY-NC-SA 3.0 IGO ed.). https://doi.org/https://doi.org/10.4060/ca9229en

FAO, IFAD, UNICEF, WFP, & WHO. (2022). *The State of Food Security and Nutrition in the World 2022. Repurposing food and agricultural policies to make healthy diets more affordable.* FAO. https://doi.org/https://doi.org/10.4060/cc0639en

Gijsbrechts, J., Boute, R. N., Van Mieghem, J. A., & Zhang, D. J. (2022). Can Deep Reinforcement Learning Improve Inventory Management? Performance on Lost Sales, Dual-Sourcing, and Multi-Echelon Problems. *Manufacturing & Service Operations Management*.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Goodfellow, I., McDaniel, P., & Papernot, N. (2018). Making machine learning robust against adversarial inputs. *Communications of the ACM, 61*(7), 56-66.

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly*, 337-355.

Habib, A., Khan, M. I., & Uddin, J. (2016). Optimal route selection in complex multi-stage supply chain networks using SARSA (λ). 2016 19th International Conference on Computer and Information Technology (ICCIT),

Halat, K., & Hafezalkotob, A. (2019). Modeling carbon regulation policies in inventory decisions of a multi-stage green supply chain: A game theory approach. *Computers & Industrial Engineering, 128*, 807-830.
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.
Himmelstein, J., Ares, A., Gallagher, D., & Myers, J. (2017). A meta-analysis of intercropping in Africa: impacts on crop yield, farmer income, and integrated pest management effects. *International Journal of Agricultural Sustainability, 15*(1), 1-10.
Hubbs, C. D., Li, C., Sahinidis, N. V., Grossmann, I. E., & Wassick, J. M. (2020). A deep reinforcement learning approach for chemical production scheduling. *Computers & Chemical Engineering, 141*, 106982.
Hubbs, C. D., Perez, H. D., Sarwar, O., Sahinidis, N. V., Grossmann, I. E., & Wassick, J. M. (2020). Or-gym: A reinforcement learning library for operations research problems. *arXiv preprint arXiv:2008.06319*.
Jahanbin, P., Wingreen, S. C., Sharma, R., Ijadi, B., & Reis, M. M. (2023). Enabling affordances of blockchain in agri-food supply chains: A value-driver framework using Q-methodology. *International Journal of Innovation Studies, 7*(4), 307-325.
Jones, D., & Gregor, S. (2007). The anatomy of a design theory. *Journal of the Association for Information Systems, 8*(5), 1.
Kemmer, L., von Kleist, H., de Rochebouët, D., Tziortziotis, N., & Read, J. (2018). Reinforcement learning for supply chain optimization. European Workshop on Reinforcement Learning,
Kocaoglu, Y., Cakmak, E., Kocaoglu, B., & Taskin Gumus, A. (2020). A novel approach for optimizing the supply chain: A heuristic-based hybrid algorithm. *Mathematical Problems in Engineering, 2020*.
Kontominas, M. G., Badeka, A. V., Kosma, I. S., & Nathanailides, C. I. (2021). Recent developments in seafood packaging technologies. *Foods, 10*(5), 940.
Krakovsky, M. (2016). Reinforcement renaissance. *Communications of the ACM, 59*(8), 12-14.
LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature, 521*(7553), 436-444.
Li, Y. (2017). Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274*.
Lin, Q., Wang, H., Pei, X., & Wang, J. (2019). Food safety traceability system based on blockchain and EPCIS. *IEeE Access, 7*, 20698-20707.
March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision support systems, 15*(4), 251-266.
Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., & Ostrovski, G. (2015). Human-level control through deep reinforcement learning. *nature, 518*(7540), 529-533.
Nazari, M. (2019). *Autonomous Decision-Making Schemes for Real-World Applications in Supply Chains and Online Systems* Lehigh University].
Nazari, M., Oroojlooy, A., Snyder, L., & Takác, M. (2018). Reinforcement learning for solving the vehicle routing problem. *Advances in neural information processing systems, 31*.
Niederman, F., & March, S. T. (2012). Design science and the accumulation of knowledge in the information systems discipline. *ACM Transactions on Management Information Systems (TMIS), 3*(1), 1-15.
Oroojlooyjadid, A., Nazari, M., Snyder, L., & Takáč, M. (2017). A deep q-network for the beer game: A reinforcement learning algorithm to solve inventory optimization problems. *arXiv preprint arXiv:1708.05924*.
Oroojlooyjadid, A., Nazari, M., Snyder, L. V., & Takáč, M. (2022). A deep q-network for the beer game: Deep reinforcement learning for inventory optimization. *Manufacturing & Service Operations Management, 24*(1), 285-304.
Oroojlooyjadid, A., Snyder, L. V., & Takáč, M. (2020). Applying deep learning to the newsvendor problem. *IISE Transactions, 52*(4), 444-463.
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.
Pramod, G., Nakamura, K., Pitcher, T. J., & Delagran, L. (2014). Estimates of illegal and unreported fish in seafood imports to the USA. *Marine Policy, 48*, 102-113.
Rabobank. (2022). *World Seafood Map* RoboBank Nederland https://research.rabobank.com/far/en/documents/729480_Rabobank_Seafood_Map_2021_DIGITAL.pdf

Rahmaniya, N., & Sekharan, M. (2018). Consumer behaviour towards seafood and seafood safety. *International Journal of current advanced research*, *7*(1), 8727-8736.
Rai, A., Patnayakuni, R., & Seth, N. (2006). Firm performance impacts of digitally enabled supply chain integration capabilities. *MIS quarterly*, 225-246.
Raj, A., Biswas, I., & Srivastava, S. K. (2018). Designing supply contracts for the sustainable supply chain using game theory. *Journal of cleaner production*, *185*, 275-284.
Ray, S. (2019). *Us consumer preferences for seafood traceability* Massachusetts Institute of Technology].
Reyes, P. M., Visich, J. K., & Jaska, P. (2020). Managing the dynamics of new technologies in the global supply chain. *IEEE Engineering Management Review*, *48*(1), 156-162.
Samadi, A., Mehranfar, N., Fathollahi Fard, A., & Hajiaghaei-Keshteli, M. (2018). Heuristic-based metaheuristics to address a sustainable supply chain network design problem. *Journal of Industrial and Production Engineering*, *35*(2), 102-117.
Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, *61*, 85-117.
Shekarian, E., Ijadi, B., Zare, A., & Majava, J. (2022). Sustainable Supply Chain Management: A Comprehensive Systematic Review of Industrial Practices. *Sustainability*, *14*(13), 7892.
Shongwe, M. I., & Bezuidenhout, C. N. (2019). A heuristic for the selection of appropriate diagnostic tools in largescale sugarcane supply systems. *AIMS Agriculture and Food*, *4*(1), 1-26.
Snyder, L. V., & Shen, Z.-J. M. (2019). *Fundamentals of supply chain theory*. John Wiley & Sons.
Statista. (2024). *Global fish production from 2002 to 2023*. Statista. https://www.statista.com/statistics/264577/total-world-fish-production-since-2002/
Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
Toledo-Hernández, M., Tscharntke, T., Tjoa, A., Anshary, A., Cyio, B., & Wanger, T. C. (2020). Hand pollination, not pesticides or fertilizers, increases cocoa yields and farmer income. *Agriculture, Ecosystems & Environment*, *304*, 107160.
Vasnani, N. N., Chua, F. L. S., Ocampo, L. A., & Pacio, L. B. M. (2019). Game theory in supply chain management: Current trends and applications. *International Journal of Applied Decision Sciences*, *12*(1), 56-97.
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.
Wang, S., Li, D., Zhang, Y., & Chen, J. (2019). Smart contract-based product traceability system in the supply chain scenario. *IEeE Access*, *7*, 115122-115133.
Wu, Z., & Chen, D. (2019). New optimal-control-based advertising strategies and coordination of a supply chain with differentiated products under consignment contract. *IEeE Access*, *7*, 170703-170714.
Zhao, W., & Wang, D. (2018). Simulation-based optimization on control strategies of three-echelon inventory in hybrid supply chain with order uncertainty. *IEEE Access*, *6*, 54215-54223.

7. Consumers' View and Discourse on ChatGPT through the Lens of YouTube Comments

Omar H. Fares
Toronto Metropolitan University
Omar.fares@torontomu.ca

Seung Hwan (Mark) Lee
Toronto Metropolitan University
Lee.mark@torontomu.ca

Joseph Aversa
Toronto Metropolitan University
javersa@torontomu.ca

Frances Gunn
Toronto Metropolitan University
fgunn@torontomu.ca

Abstract

This research explores consumers' views of AI conversational bots, explicitly focusing on ChatGPT by OpenAI and analyzing 45,844 YouTube comments. The study examines the societal implications of AI bots, emphasizing public perception as a key factor in their acceptance and integration. Three guiding questions shape the research: 1) What themes emerge from YouTube comments about ChatGPT? 2) How do these themes reflect the public's concerns and understanding of ChatGPT and AI? 3) What future research directions can be proposed based on these themes? The study extracts and interprets key themes and their relationships by employing manual and automated text analysis, including Leximancer, an AI-based software. Leximancer's AI capabilities are crucial in identifying patterns in large data sets, providing an in-depth understanding of public sentiment. The research categorizes attitudes towards AI bots into techno-skepticism, techno-realism, and techno-optimism, illustrating varied public opinions. Themes identified include concerns about AI's impact on employment, privacy, misinformation, and its potential for enhancing efficiency and creativity. These insights bridge AI's technical aspects with its social and ethical dimensions. The study underscores the importance of addressing public concerns and suggests informed approaches for AI development and societal integration, offering valuable insights for developers, researchers, and policymakers.

Keywords: Conversational Bots, ChatGPT, Consumer Perceptions, Thematic Analysis, Qualitative research

1. Introduction

In the current landscape of rapid technological progress, the emergence and widespread public accessibility of consumer-friendly Artificial Intelligence (AI) systems represent a transformative milestone (Goel et al., 2022; Gašević et al., 2023). Consumer-friendly AI may be referred to as AI tools designed with a focus on ease of use, accessibility, and positive consumer experience (Stige et al., 2023). Within the domain of AI, language models such as ChatGPT (Generative Pre-trained Transformer) by OpenAI have captured significant public attention (i.e., reaching a million consumers in five days) due to their sophisticated natural language processing abilities coupled with a consumer-friendly interface, which enable nuanced understanding and generation of human-like language (Exponential View, 2022; Ray, 2023). However, as AI systems grow increasingly sophisticated and embedded in societal functions, it is crucial to maintain a rigorous and critical perspective on both their potential and

limitations. The central objective of this study is to identify and analyze the key themes in public discourse related to ChatGPT, with a specific focus on YouTube as the primary data source. The selection of YouTube for this research is based on its unique position as a global platform that attracts many content creators and viewers (Snelson, 2011). This platform is a repository of information and a vibrant forum for public engagement and discussion. By focusing on YouTube as the key source for thematic analysis, this research aims to harness the platform's potential as a real-time barometer of public opinion.

2. Background and Context

Integrating conversational bots into daily life has been a subject of interest in recent years. However, the rapid advancement and increasing sophistication of these technologies, exemplified by AI systems like ChatGPT, have intensified their impact on society (Abdullah et al., 2022). The literature highlights the importance of understanding consumer perceptions of AI technologies for several reasons. Firstly, the acceptance and effective utilization of AI tools are contingent upon public trust and perception (Baek & Kim, 2023; Choudhury & Shamszare, 2023). Consumer feedback provides invaluable insights into how these technologies meet actual needs and expectations, highlighting areas for improvement and adaptation (Dignum, 2019). Secondly, exploring consumer views is crucial in addressing potential ethical and societal concerns associated with conversational bots. Issues related to privacy, data security, and the potential for misuse have been extensively discussed in the literature (Cath et al., 2018; Lupton, 2014). As conversational bots like ChatGPT gain enhanced capabilities, they raise complex questions about their role in society and the ethical boundaries of their application. Thirdly, analyzing consumer views aids in predicting future trends in AI development and application. Public opinion can significantly influence market demand and the direction of research and innovation in AI (Shin, 2020). By understanding how consumers view and interact with ChatGPT, developers and policymakers can better anticipate future needs and challenges, aligning AI advancements with societal values and expectations.

Our study builds upon this existing body of literature by focusing specifically on ChatGPT, a prominent example of a conversational AI bot, and analyzing public discourse on YouTube, a platform that has received relatively less attention in academic research on AI perceptions. By examining YouTube comments, this research aims to capture a wide range of consumer opinions and provide insights into the public's understanding and attitudes toward ChatGPT and AI technologies in general. This approach aligns with the work of Snelson (2011), who highlighted the importance of YouTube as a platform for public engagement and discussion. Furthermore, our study contributes to the ongoing discourse on digital technologies' societal and ethical implications (Cath et al., 2018; Lupton, 2014). Therefore, our study is structured around several sub-questions:

1) What themes and topics emerge from the analysis of YouTube comments regarding ChatGPT?

2) How do these themes reflect consumers' concerns, perceptions, and understanding of ChatGPT and AI technologies?

3) Based on the thematic analysis, what propositions can be made for future research to address key issues related to AI bots area?

3. Method and Findings

Our study analyzed the top five non-sponsored, English-language YouTube videos on ChatGPT, selected for their relevance, popularity, and content diversity. This selection balanced comprehensiveness with manageability, ensuring a broad representation of public opinions while minimizing biases from sponsored content and YouTube Shorts (Ahamad et al., 2017; Zian et al., 2021). The dataset, collected in 2023 using Octoparse, comprised 45,844 comments. Sentiment analysis revealed 22.1% negative and 77.9% positive comments, indicating a predominantly positive reception. The average sentence length was 10.54 words, and 26.7% of words were captured for Linguistic Inquiry and Word Count (LIWC) analysis. The channels varied in size, with subscribers ranging from 50,400 to 16.4 million, and video views from 4 million to 8.2 million, reflecting diverse public opinions. Ethical considerations were addressed by complying with YouTube's terms, maintaining transparency, and anonymizing data.

This study's thematic analysis of YouTube comments regarding ChatGPT follows a grounded theory approach (Strauss and Corbin, 1998). Initially, comments were examined line-by-line to identify first-order codes representing the data's most basic elements (Charmaz, 2006). For example, comments mentioning "privacy concerns" or "misinformation" were coded accordingly. Next, these first-order codes were grouped into second-order codes based on their similarities and relationships. For instance, codes related to "privacy concerns," "data security," and "surveillance" were grouped under the broader category of "Potential Risks and Pitfalls." Finally, the second-order codes were further synthesized to identify overarching themes that capture the essence of the data. The three distinct themes emerged from this process: "Potential Risks and Pitfalls," "Potential Benefits and Usage Areas," and "Attitudes Towards Technology." This methodology emphasizes the emergence of themes from the data itself rather than imposing preconceived categories or frameworks. This aligns with the principles of thematic analysis described by Braun and Clarke (2006), who emphasize the flexibility and iterative nature of identifying themes in qualitative data.

Following the qualitative coding, Leximancer analysis is employed to develop a conceptual map of concept occurrence and relationships. Leximancer identifies key concepts and their co-occurrences in the dataset, providing a quantitative perspective on the data (Sotiriadou et al., 2014). Manual coding aims to capture the depth and nuances of individual comments, while Leximancer quantifies the prominence and relationships of concepts in the entire dataset. The manual coding results provide insights into the specific themes and concerns expressed by YouTube users, while Leximancer analysis offers a broader view of the overall discourse landscape (Lemon and Hayes, 2020). By integrating these two sets of findings, we can appreciate both the detailed narratives and the overarching patterns in the public discourse on ChatGPT (Sotiriadou et al., 2014).

In the first part of the analysis, the initial categories were refined and grouped, consolidating three distinct themes: **a) Potential Risks and Pitfalls**: This theme emerged from comments expressing concerns and skepticism about ChatGPT. Such apprehensions are common in public discourse about emerging technologies, as discussed by Lupton (2014), who explores digital technologies' societal and ethical implications; **b) Potential Benefits and Usage Areas**: Reflecting the dual nature of public opinion towards technology, this theme captures the

positive perceptions and reported advantages of using ChatGPT. Rogers et al. (2014), in their diffusion of innovations theory, highlight how perceived benefits play a crucial role in the adoption of new technologies; and **c) Attitudes Towards Technology**: This broader theme encompasses the spectrum of opinions and philosophical stances towards AI and technology as a whole. This theme resonates with the concept of technological determinism and the social construction of technology theories, as explored by MacKenzie and Wajcman (1999). Table 1 shows the key themes with sample quotes from the data.

Themes	Sample Quotes
Potential Risks & Pitfalls	"How would you need to meet Basic Human needs through AI. Real needs would need real and physical work, not any artificial intelligence. How would you grow fruits and vegetables or produce milk and meet with it?" "higher authorities strictly told us not to use Chat GPT as it can provide false information." "there is a serious threat.. To use AI.. Almost privacy is finished in the modern era.." "Ai isn't gonna end anything, but will surely end some lives, gap between top & bottom line of population will widen... Privacy will be brought down, mass manipulation will be major part. Last one is AI will surely be used in lethal weapons development." "This may end up breaking the internet... online privacy will be impossible to implement... even with the best encryption available... in all this excitement we should not forget... we are playing with fire.." "ChatGPT is gone down the drain though, always telling wrong information." "I tried using ChatGPT to write product descriptions and it was terrible. Repetitive fluff and inaccurate information." "It gives inaccurate information when I study any particular topic deeply" "These models are trained on data, web content and media, it is unfortunate that most of this data is from the Western world, anglosphere, while data from developing countries who have hardly any online presence is lacking. This is what makes these AI models extremely biased and lopsided."
Potential Benefits and Usage Areas	"For the record, I had no knowledge of coding at all and have used Chat GPT to make several scripts that have helped me in my job and just general day to day computer efficiency. When I encountered an error in the code I would simply open a new chat with GPT and ask it what happened and how to solve it. Chat GPT has effectively taught me how to code on a basic level and I can now analyse code that it creates and figure out where problems are when they arise and then use Chat GPT's advice to solve the problems I find" "I am a student of IT. I use ChatGPT for my Assignments programs and daily study." "I asked it to create a Christmas menu with appetizers, including recipes and a shopping list. It was brilliant. I was able to make adjustments for something I didn't like." "Chat GPT successfully diagnosed a medical condition I had suffered from for many years. My doctors and specialists had come up with nothing after 12 years of tests and consultations. I was blown away." "I am myself a student of Class 12 and accept the fact that I used ChatGPT to complete all my projects and, really it saved a lot of time for me instead of scrolling through the web for hours to collect correct and factual information I got it within 2-3 mins from ChatGPT." "I used chatgpt to complete my program and it's very useful really very helpful I like chatgpt" "I just used chatgpt to write a description for my channel" "I have used chatgpt to write code for website I have to say it does better job than me" "Personally I used ChatGPT to make a rather simple mod for Space Engineers, to my surprise it did work. Not only did ChatGPT know what to do it constantly helped me finding the different files I needed, or which I needed to create myself."
Attitude towards Technology	"No AI can ever beat humans" "The biggest difference between a human and AI...is we don't want AI because AI made people very blind and brainless" "that is scary!" "AI will not replace human human will be replaced by people who have knowledge to use Ai" "Nope chatgpt must be erased from everywhere as must all Ai. You will all be controlled you will all become drones and do the bidding of Ai" "As an ai engineer, I want to say that it would take decades, or even centuries to be capable of what humans can do. It is only a tool to make our work much more efficient" "This technology is nothing but a mirror, it is the DEFINITION of confirmation bias." "This world is not for robots, it's for humans and animals, only living things...I hate robots, do Even like perfection, perfection don't have beauty of life, so only this generation of gen z at their age 30-40 after that human existence will be in robots hand, my god so we all have to understand that we don't need robots and ai, we need our old-culture we need peace, at this time only humans was in the race now artificial things also in the race, what if we don't need any artificial things anymore it can stop, but some humans are also have cruel mind like robots they are like that they want perfection they don't think about humans and generations..." "Artificial intelligence can never obtain the critical thinking and problem solving like human brains"

Table 1. Thematic analysis

Concerns about conversational bots' role in meeting basic human needs and their impact on employment are significant. Some consumers express skepticism, with one remarking, "How

would you need to meet Basic Human needs through AI. Real needs would need real and physical work, not any artificial intelligence." This sentiment highlights the skepticism that AI cannot fully replace the nuanced capabilities of human labour (De Cremer & Kasparov, 2021), particularly in vital areas like agriculture. The risk of misinformation also surfaces (De Angelis et al., 2023), with consumers warning about the reliability of AI-generated content, evidenced by the statement, "higher authorities strictly told us not to use Chat GPT as it can provide false information." Privacy concerns are also prevalent, with the sentiment that "almost privacy is finished in the modern era.." These comments underscore the challenges posed by AI's extensive data collection capabilities, leading to potential privacy issues (Wu et al., 2023).

Conversely, the practical benefits of ChatGPT in educational and personal contexts receive praise. Students utilize ChatGPT for coding assistance and assignment completion; for example, one consumer states, "I am a student of IT. I use ChatGPT for my Assignments programs and daily study." This highlights the role of conversational bots as a potential learning and efficiency tool, yet it raises critical questions about the future of education and learning, as highlighted by Lo (2023). Another consumer expresses satisfaction with ChatGPT's ability to aid in medical diagnosis, remarking, "Chat GPT successfully diagnosed a medical condition we had suffered from for many years." While conversational bots may provide accurate responses, such perceptions bring to light concerns regarding the medical application of conversational bots (Garg et al., 2023). Specifically, the risk is higher when providing inaccurate health-related information that seems coherent and could put consumers at significant risk (Kuroiwa et al., 2023). Creative tasks like menu planning and content creation are additional areas where ChatGPT may be helpful (Guo et al., 2023). One consumer shares, “I asked it to create a Christmas menu with appetizers, including recipes and a shopping list. It was brilliant. I was able to make adjustments for something I didn't like." These comments highlight ChatGPT's versatility and potential to enhance personal productivity and problem-solving while prompting significant questions about perceptions versus actual technical capabilities.

Attitudes toward technology exhibited a spectrum from caution to dismissal to optimism toward integrating AI bots into daily lives. Statements such as "No AI can ever beat humans" and "The biggest difference between a human and AI...is we don't want AI because AI made people very blind and brainless" reflect a deep-rooted skepticism and concern for AI's potential to diminish human skills and decision-making. There is a palpable fear of losing control, with one comment warning, "You will all be controlled you will all become drones and do the bidding of Ai." Yet, some expressed a more balanced view, acknowledging the limitations of AI but also its utility in augmenting human work; as one AI engineer mentioned, "It would take decades, or even centuries, to be capable of what humans can do." These findings demonstrate a complex web of attitudes where conversational bots are both a tool and a threat, a helper and a harbinger of potential societal disruption (Abdullah et al., 2022). The variance in viewpoints illustrates the multi-faceted nature of public sentiment towards conversational AI, highlighting an apparent dichotomy between the perceived risks and the acknowledged benefits and a dynamic range of attitudes toward technological advancement.

To further triangulate the analysis and provide further validity to the findings, we conduct a Leximancer analysis that yields a conceptual map based on concept occurrence—with 'concept'

signifying a collection of related words—and a relational map that highlights the co-occurrence of these concepts through a probabilistic percentage. This analytical strategy reveals four predominant concepts: *AI*, *People*, *Code*, and *Information*, each bearing significant thematic weight within the discourse. Figure 1 showcases the conceptual map.

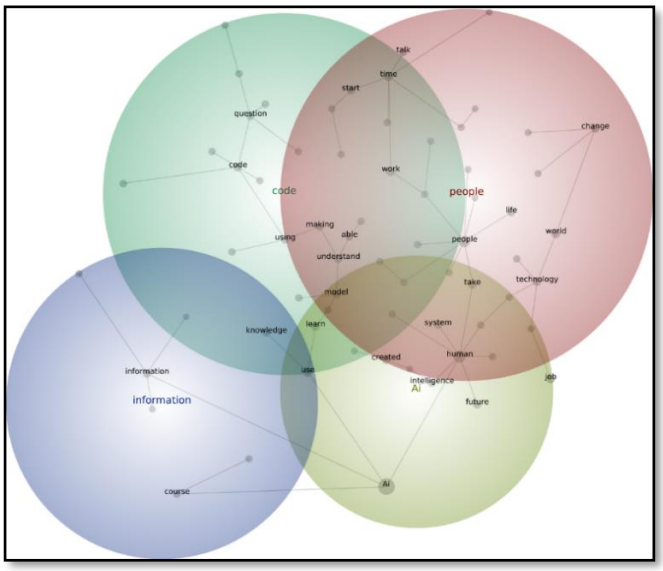


Figure 1. Conceptual Map

The most prominent theme, AI, with a hit count of 13,568, highlights the central focus of the discussions. This central theme of AI frequently intersects with pivotal societal elements (Vesnic-Alujevic et al., 2020), as evidenced by significant co-occurrences with terms such as 'human' (co-occurrence: 774, likelihood: 30%), 'future' (co-occurrence: 331, likelihood: 32%), and 'society' (co-occurrence: 125, likelihood: 32%). In the context of 'Humans and AI,' the discourse sways between apprehension and inevitability. Sentiments vary from wary skepticism, with one consumer stating, "AI can't Rule over human Intelligence," to the fatalistic acceptance in the comment, "AI will replace human." These dichotomous views reflect the complexity of integrating AI into the fabric of human activity and consciousness (Scharre, 2019). Moreover, the intersection of 'Future and AI' evokes a range of projections and speculations. While some express concern over AI's trajectory, noting "I think it is [AI] dangerous in future," others caution against over-enthusiasm: "don't get excited on AI influence in coming future. It is very harmful. People will suffer a lot." Conversely, there is a recognition of AI's inevitability and potential with the straightforward assertion, "No doubt Ai is our future," highlighting the anticipation of AI's integration into the evolution of human progress. Discussions around 'Society and AI' emphasize the collective responsibility and opportunity presented by AI development. One comment suggests that "there is a need for greater collaboration between industry, academia, and government to ensure that AI is developed in a way that benefits society as a whole," proposing a unified approach (Floridi et al., 2021).

The 'People' theme, with a notable 10,602 hits, further highlights the human-centric aspect of the discourse surrounding conversational bots. This theme is linked with financial, power, and societal dynamics. Notably, interactions between 'people and money' (co-occurrence: 85,

likelihood: 24%), 'people and power' (co-occurrence: 71, likelihood: 14%), and 'people and society' (co-occurrence: 54, likelihood: 14%) are prominent. In the 'People and Money' context, concerns about the financial implications of AI are evident. Comments such as “Don't worry soon bots are going to scam people's hard earned money” and “It's a tool for lazy people, made by people who are looking for quick money” reflect apprehensions about the monetization and potential misuse of AI technologies. These sentiments resonate with the research by Brown and Hurmuzlu (2023), who discuss the ethical challenges in AI's commercialization. Regarding 'People and Power,' the conversation goes into the dynamics of control and influence. One consumer points out, “People want to hold back AI simply because they realize that it indirectly forces their foot off the necks of all the people they currently have power over.” Another adds, “It further concentrates the wealth and power in the hands of a few people who own the technology.” These statements align with the observations by Gupta and Kumar (2022), highlighting the power shifts and potential inequality exacerbated by AI adoption. In the sphere of 'People and Society,' the focus shifts to the broader societal impact. Further, the 'People' theme reveals deep-seated concerns about economic, power, and societal dynamics. This analysis highlights the necessity of considering the human element in AI development and deployment, as echoed in the literature (e.g., Dova et al., 2022; Hurmuzlu, 2023; Mika et al., 2019, emphasizing the need for ethical considerations, equitable power distribution, and societal welfare in the age of AI.

The "Code" theme, with a significant 7,943 hits, emphasizes the technical aspects surrounding conversational bots in the discourse. The discussions notably center around the interactions between 'code and learn' (co-occurrence: 106, likelihood: 12%), 'code and knowledge' (co-occurrence: 57, likelihood: 8%), and 'code and work' (co-occurrence: 125, likelihood: 8%). Regarding 'Code and Learn,' the focus is on the educational aspects of coding with AI assistance. Consumers share their experiences, with one stating, “ChatGTP has been the most fun in the coding realm. I've had so much fun and learned so much recently.” Others reflect on the ease of learning coding now, as noted in remarks like “Cool, now I don't have to learn any more code” and “Anyone can easily learn how to write codes now. I wish I had this 10 years ago.” This mirrors the findings of Sundberg and Holmström (2023), who emphasize the transformative role of AI in democratizing coding education. Moreover, “Chat writes the code in a few seconds” highlight concerns about job displacement and the efficiency of AI in coding tasks. These observations resonate with Gruetzemacher et al. (2021), who examine the implications of AI automation on the labour market. The discourse indicates a growing interest in the role of AI in coding, from facilitating learning to transforming job requirements and impacting the nature of work. This trend highlights the need for ongoing discussions about the role of AI in technical education and the workforce, particularly considering the rapid advancements in publicly available AI capabilities and their impact on traditional coding roles. Lastly, the theme "Information" with 2,873 hits reflects diverse perspectives on the role and impact of information in the context of AI and ChatGPT. Discussions center around 'Information and Useful' (co-occurrence: 20, likelihood: 12%), 'Information and Data' (co-occurrence: 49, likelihood: 10%), and 'Information and Answer' (co-occurrence: 28, likelihood: 4%). Under 'Information and Useful,' the discourse focuses on the practical utility of ChatGPT in various fields. Consumers express mixed views, with some finding ChatGPT helpful, as one remarks, “ChatGPT is a programmer's rubber ducky that actually gives useful information.” Conversely, others caution against the potential misuse of assistance, highlighted in statements

like “That moment usefulness becomes harmful. And you’re helping them gather info to use against you.” This dichotomy highlights the fine line between beneficial and detrimental use of AI in information gathering, reflecting concerns similar to those raised by Gupta et al. (2023) on threats of ChatGPT. Regarding 'Information and Data,' the conversation highlights the intricacies of data processing and bias. A consumer notes, “I can see that deep neural networks are a great way to explore how to process data into information and explore the circuits of a brain.” This enthusiasm is tempered by critical views on bias, as another consumer comments, “Knowing the biases of the data scientists and sources of information is key to how accurate ChatGPT, or any other type of media is.” This highlights the importance of understanding data provenance and biases in AI systems, echoing research by Daneshjou et al. (2021) on transparency in AI.

4. Discussion

The findings from the thematic analysis and Leximancer analysis are complementary and together inform the conceptual framework of our study. Based on a grounded theory approach, the thematic analysis provided a deep understanding of the specific themes and concerns of YouTube users regarding ChatGPT. This qualitative analysis revealed three main themes: "Potential Risks and Pitfalls," "Potential Benefits and Usage Areas," and "Attitudes Towards Technology."

Leximancer analysis, on the other hand, offered a quantitative perspective by identifying key concepts and their relationships within the dataset. While the concepts identified by Leximancer did not directly match the themes from the manual coding, they provided a broader view of the discourse landscape, highlighting the prominence of concepts such as "AI," "People," "Code," and "Information."

Integrating these two sets of findings allowed us to develop a comprehensive conceptual framework that captures both the nuanced perspectives of individual commenters and the overarching patterns in the public discourse. The framework illustrates the dual nature of public sentiment towards ChatGPT, encompassing both optimism and skepticism and highlights the critical areas of concern and interest among YouTube users. By combining the depth of the thematic analysis with the breadth of the Leximancer analysis, we could construct a more holistic understanding of the public discourse on ChatGPT, which is reflected in our conceptual framework as highlighted in figure 2.

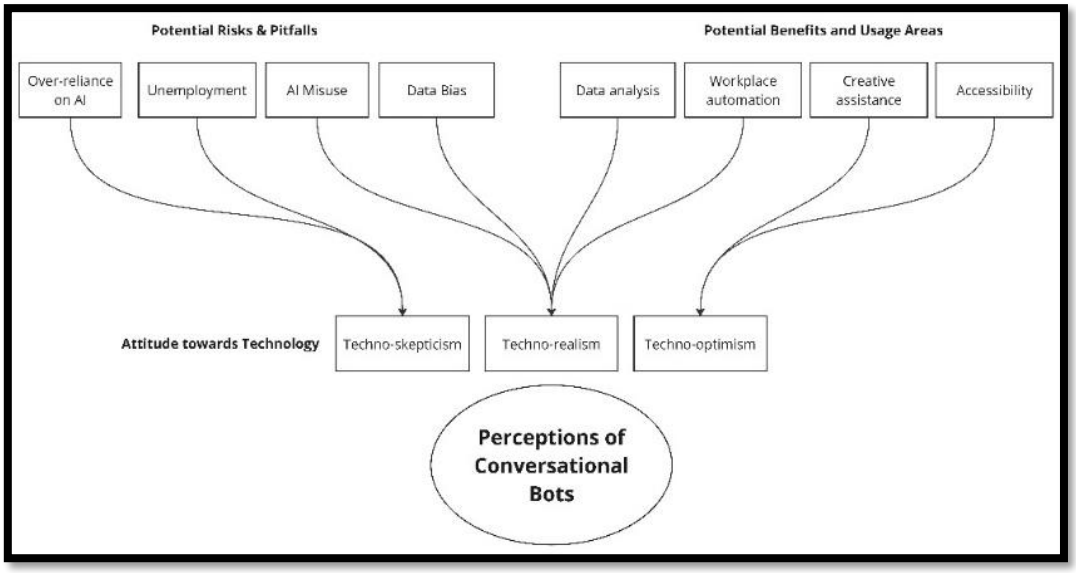


Figure 2. Framework for Public Discourse on ChatGPT

The framework serves as a systematic exploration of the dichotomous impact of AI on society. On one side, we observe apprehensions regarding over-reliance on AI, epitomized by the fear that machines may supplant human decision-making, leading to potential deskilling (Frey & Osborne, 2017). Moreover, the automation of jobs, while enhancing efficiency, presents a profound challenge to employment (Acemoglu & Restrepo, 2020), sparking fears of widespread unemployment and economic disruption. The misuse of AI, ranging from the perpetuation of fake news to unethical surveillance practices (O'Neil, 2017), along with the ever-present issue of data and algorithmic bias (Noble, 2018), compounds these anxieties, fostering a techno-skepticism that questions the unchecked proliferation of AI technologies.

Conversely, the framework also encapsulates the optimistic narrative surrounding AI, showcasing its potential to significantly enhance data analytics, enabling unprecedented insights and decision-making capabilities (Mayer-Schönberger & Cukier, 2013). The automation of mundane tasks liberates human creativity and allows for a shift towards more fulfilling work (Brynjolfsson & McAfee, 2014), while AI's ability to provide creative assistance and enhance accessibility signifies its role as an enabler and equalizer in society (Wu et al., 2021).

The framework's depiction of attitudes towards technology—techno-skepticism, techno-realism, and techno-optimism—reflects the spectrum of consumer sentiment identified in the literature. Techno-skepticism, wary of AI's threats to privacy, identity, and agency (Ngwenyama et al., 2023; Zuboff, 2023), advocates for rigorous ethical safeguards and a cautious approach to AI integration. This perspective resonates with consumers who prioritize data security and personal agency over the conveniences afforded by AI (Hagendorff, 2020). Techno-realism occupies a middle ground, advocating for a balanced, critical approach to AI adoption that acknowledges both its transformative potential and its fallibilities (Russell & Norvig, 2010). This stance aligns with pragmatic consumers who are willing to embrace AI, provided that its application is transparent, its governance is equitable, and its benefits are

tangibly realized (Crawford & Calo, 2016). Techno-optimism, on the other end of the spectrum, is fueled by a strong belief in the transformative power of AI. This outlook is shared by consumers who see AI as a progressive force that will drive innovation and societal growth and improve quality of life (Plastino & Purdy, 2018).

5. Future Directions and Propositions

We utilize the findings to formulate the following propositions that we argue are essential to address in future research.

Proposition 1: Systematic Integration of AI Literacy in Educational Curricula

The widespread presence of AI in society calls for an essential understanding of its mechanisms and implications. The first proposition advocates for integrating AI literacy into educational curricula to address the concerns of AI over-reliance and potential deskilling. This initiative aims to create an informed populace capable of responsibly engaging with AI (Mhlanga, 2023). Future research should evaluate the effectiveness of AI literacy programs and their impact on workforce adaptability, job role evolution, and reducing societal and economic disparities due to unequal AI access (Kulkarni et al., 2023). This proposition aligns with the "Potential Benefits and Usage Areas" theme from the thematic analysis, where ChatGPT's educational applications were highlighted. The Leximancer analysis's emphasis on "Code" and "Information" further underscores the importance of AI literacy in education to leverage AI technologies effectively.

Proposition 2: Establishment of Specialized Regulatory Bodies for AI Oversight

The evolving challenges of AI bots require dedicated regulatory bodies focused on AI governance (Crawford & Calo, 2016; Wu et al., 2023). These bodies would ensure a balance between innovation and societal welfare, shaping a future where AI benefits are equitably distributed. Future research should explore the optimal structure and function of such regulatory entities. This proposition is supported by the "Potential Risks and Pitfalls" theme, highlighting concerns like privacy and misinformation. The Leximancer analysis's emphasis on "People" and "Society" underscores AI's social implications, underscoring the need for regulatory oversight to address AI's ethical challenges.

Proposition 3: Regular Auditing and Bias Mitigation in AI Algorithms

The need for ongoing monitoring and evaluation of AI algorithms is underscored by their potential to reflect and amplify biases, as noted in consumer comments. Regular auditing is crucial, especially in critical areas like healthcare and law enforcement, to ensure ethical AI functionality (Challen et al., 2019). Future research should focus on developing methodologies to detect and correct biases in AI systems, including advanced algorithms for bias identification and new models with reduced bias susceptibility. This proposition is informed by the "Attitudes Towards Technology" theme, which revealed skepticism about AI fairness, and the Leximancer analysis that highlighted the significance of "Information" in AI systems. The emphasis on regular auditing and bias mitigation is a response to concerns about AI reliability and ethics.

6. Limitations

This research has multiple limitations. One primary limitation arises from the reliance on YouTube comments as the sole data source for public discourse analysis. While these comments offer a rich pool of consumer opinions, they do not represent the full spectrum of public perceptions, potentially exhibiting demographic or interest biases. However, this approach still provides valuable real-time insights into consumer perceptions, serving as a meaningful starting point for understanding public sentiment toward AI bots. Moreover, The findings are specific to ChatGPT and the time frame of data collection, may have limited applicability over time or in different contexts, potentially affecting their generalizability. However, these findings offer a snapshot of current public attitudes and can guide rapid strategies for AI bots' integration and policy development. Lastly, the thematic analysis relies on the subjective interpretation of comments, which could introduce bias. However, the use of methodological triangulation and Leximancer software adds a level of rigour and objectivity to the analysis, enhancing the reliability of the identified themes.

References

Abdullah, M., Madain, A., & Jararweh, Y. (2022). ChatGPT: Fundamentals, applications and social impacts. In 2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS) (pp. 1-8). IEEE.

Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. Journal of political economy, 128(6), 2188-2244.

Ahamad, D., Mahmoud, D. A. M. M., & Akhtar, M. M. (2017). Strategy and implementation of web mining tools. International Journal of Innovative Research in Advanced Engineering, 4(12), 1-7.

Baek, T. H., & Kim, M. (2023). Is ChatGPT scary good? How user motivations affect creepiness and trust in generative artificial intelligence. Telematics and Informatics, 83, 102030

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative research in psychology, 3(2), 77.

Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. WW Norton & Company.

Cath, C., Wachter, S., Mittelstadt, B., Taddeo, M., & Floridi, L. (2018). Artificial intelligence and the 'good society': the US, EU, and UK approach. Science and engineering ethics, 24, 505-528.

Challen, R., Denny, J., Pitt, M., Gompels, L., Edwards, T., & Tsaneva-Atanasova, K. (2019). Artificial intelligence, bias and clinical safety. BMJ Quality & Safety, 28(3), 231-237.

Charmaz, K. (2006). Constructing grounded theory: A practical guide through qualitative analysis. sage.

Choudhury, A., & Shamszare, H. (2023). Investigating the Impact of User Trust on the Adoption and Use of ChatGPT: Survey Analysis. Journal of Medical Internet Research, 25, e47184.

Clarke, V., & Braun, V. (2017). Thematic analysis. The journal of positive psychology, 12(3), 297-298.

Crawford, K., & Calo, R. (2016). There is a blind spot in AI research. Nature, 538(7625), 311-313.

Daneshjou, R., Smith, M. P., Sun, M. D., Rotemberg, V., & Zou, J. (2021). Lack of transparency and potential bias in artificial intelligence data sets and algorithms: a scoping review. JAMA dermatology, 157(11), 1362-1369.

De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023). ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. Frontiers in Public Health, 11, 1166120.

De Cremer, D., & Kasparov, G. (2021). AI should augment human intelligence, not replace it. Harvard Business Review, 18, 1.

Dignum, V. (2019). Responsible artificial intelligence: how to develop and use AI in a responsible way (Vol. 2156). Cham: Springer.

Exponential View. (2022). Adoption rate for major milestone internet-of-things services and technology in 2022, in days [Graph]. In Statista.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2021). An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Ethics, governance, and policies in artificial intelligence, 19-39.
Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation?. Technological forecasting and social change, 114, 254-280.
Garg, R. K., Urs, V. L., Agarwal, A. A., Chaudhary, S. K., Paliwal, V., & Kar, S. K. (2023). Exploring the role of ChatGPT in patient care (diagnosis and treatment) and medical research: A systematic review. Health Promotion Perspectives, 13(3), 183.
Gašević, D., Siemens, G., & Sadiq, S. (2023). Empowering learners for the age of artificial intelligence. Computers and Education: Artificial Intelligence, 4, 100130.
Glaser, B., & Strauss, A. (2017). Discovery of grounded theory: Strategies for qualitative research. Routledge.
Goel, P., Kaushik, N., Sivathanu, B., Pillai, R., & Vikas, J. (2022). Consumers' adoption of artificial intelligence and robotics in hospitality and tourism sector: literature review and future research agenda. Tourism Review, 77(4), 1081-1096.
Gruetzemacher, R., Paradice, D., & Lee, K. B. (2020). Forecasting extreme labor displacement: A survey of AI practitioners. Technological Forecasting and Social Change, 161, 120323.
Guo, C., Lu, Y., Dou, Y., & Wang, F. Y. (2023). Can ChatGPT boost artistic creation: The need of imaginative intelligence for parallel art. IEEE/CAA Journal of Automatica Sinica, 10(4), 835-838.
Gupta, M., Akiri, C., Aryal, K., Parker, E., & Praharaj, L. (2023). From chatgpt to threatgpt: Impact of generative ai in cybersecurity and privacy. IEEE Access.
Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. Minds and machines, 30(1), 99-120.
Hurmuzlu, O. (2023). Ethical Implications of the Commercialization of Artificial Intelligence (AI).
Kulkarni, M., Mantere, S., Vaara, E., van den Broek, E., Pachidi, S., L Glaser, V., ... & Greenwood, M. (2023). The future of research in an artificial intelligence-driven world. Journal of Management Inquiry.
Kuroiwa, T., Sarcon, A., Ibara, T., Yamada, E., Yamamoto, A., Tsukamoto, K., & Fujita, K. (2023). The potential of ChatGPT as a self-diagnostic tool in common orthopedic diseases: exploratory study. Journal of Medical Internet Research, 25, e47621.
Lemon, L. L., & Hayes, J. (2020). Enhancing trustworthiness of qualitative findings: Using Leximancer for qualitative data analysis triangulation. The Qualitative Report, 25(3), 604-614.
Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. Education Sciences, 13(4), 410.
Lupton, D. (2014). Digital sociology. Routledge.
MacKenzie, D., & Wajcman, J. (1999). The social shaping of technology. Open university press.
Mayer-Schönberger, V., & Cukier, K. (2013). Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt.
Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning.
Mika, N., Nadezhda, G., Jaana, L., & Raija, K. (2019). Ethical AI for the governance of the Society: Challenges and opportunities. In CEUR Workshop Proceedings (Vol. 2505, pp. 20-26).
Ngwenyama, O., Rowe, F., Klein, S., & Henriksen, H. Z. (2023). The Open Prison of the Big Data Revolution: False Consciousness, Faustian Bargains, and Digital Entrapment. Information Systems Research.
Noble, S. U. (2018). Algorithms of oppression. In Algorithms of oppression. New York university press.
O'neil, C. (2017). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown.
Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet of Things and Cyber-Physical Systems.
Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. In An integrated approach to communication theory and research (pp. 432-448). Routledge.
Russell, S. J., & Norvig, P. (2010). Artificial intelligence a modern approach. London.
Scharre, P. (2019). Killer apps: The real dangers of an AI arms race. Foreign Aff., 98, 135.
Shin, D. (2020). User perceptions of algorithmic decisions in the personalized AI system: Perceptual evaluation of fairness, accountability, transparency, and explainability. Journal of Broadcasting & Electronic Media, 64(4), 541-565.
Snelson, C. (2011). YouTube across the disciplines: A review of the literature. MERLOT Journal of Online learning and teaching.

Sotiriadou, P., Brouwers, J., & Le, T. A. (2014). Choosing a qualitative data analysis tool: A comparison of NVivo and Leximancer. Annals of leisure research, 17(2), 218-234.
Stige, Å., Zamani, E. D., Mikalef, P., & Zhu, Y. (2023). Artificial intelligence (AI) for user experience (UX) design: a systematic literature review and future research agenda. Information Technology & People.
Strauss, A., & Corbin, J. (1998). Basics of qualitative research techniques.
Sundberg, L., & Holmström, J. (2023). Democratizing artificial intelligence: How no-code AI can leverage machine learning operations. Business Horizons.
Vesnic-Alujevic, L., Nascimento, S., & Polvora, A. (2020). Societal and ethical impacts of artificial intelligence: Critical notes on European policy frameworks. Telecommunications Policy, 44(6), 101961.
Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q. L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. IEEE/CAA Journal of Automatica Sinica, 10(5), 1122-1136.
Wu, X., Duan, R., & Ni, J. (2023). Unveiling security, privacy, and ethical concerns of ChatGPT. Journal of Information and Intelligence.
Wu, Z., Ji, D., Yu, K., Zeng, X., Wu, D., & Shidujaman, M. (2021). AI creativity and the human-AI co-creation model. In Human-Computer Interaction. Theory, Methods and Tools: Thematic Area, HCI 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part I 23 (pp. 171-190). Springer International Publishing.
Zian, L. Q., Kumar, Y. J., & Zulkarnain, N. Z. (2021). Tools for Big Data and Analytics. Manuscript Editor, 2021, 76.
Zuboff, S. (2023). The age of surveillance capitalism. In Social Theory Re-Wired (pp. 203-213). Routledge.

8. Demystifying the Maze: An Investigation of Barriers to AI Marketing Adoption in SMEs

Mariya Rai
University of Bedfordshire
mariya.rai@study.beds.ac.uk

Yanqing Duan
University of Bedfordshire
yanqing.duan@beds.ac.uk

Annie Danbury
University of Bedfordshire
annie.danbury@beds.ac.uk

Abstract

AI is transforming business in an unprecedented way, but SMEs are still struggling to benefit from its huge potential through effective adoption and integration of AI tools. This is particularly true in terms of using AI in marketing by SMEs. This research-in-progress paper describes a study that aims to investigate the AI marketing adoption in UK SMEs by employing in-depth, semi-structured interviews with both SME owners/managers and AI providers. The investigation examines the challenges that SMEs have in adopting AI and explores the role of AI providers in this context. Interviews uncovered a widespread sense of ambiguity regarding AI marketing in the SMEs and offered unique insights into the adoption behavior of small and medium-sized enterprises (SMEs) towards AI marketing. Participants voiced strong apprehensions over the limited knowledge and understanding, resource limitations, and hesitation and skepticism regarding the suitability of the AI marketing tools for the type and size of their business, despite recognizing its potential advantages. The findings reveal a lack of trust and notable disparity in communication between AI providers and SMEs and calls for enhancing the knowledge and understanding of SME owners/managers by gaining a deeper grasp of the role of AI providers.

Keywords: Artificial Intelligence (AI), Marketing, Small and Medium Enterprises (SMEs), Adoption, AI Providers.

1. Introduction

The rise of artificial intelligence (AI) has sparked a surge in innovation. AI is transforming various sectors and presenting unparalleled prospects for enterprises of all scales. Revolutionizing various aspects of businesses. Artificial intelligence, in conjunction with the Internet of things (IoT) and Big Data, has the potential to significantly transform the economy (Agrawal et al., 2017). Within the field of marketing, AI offers powerful tools for data analysis, targeted campaigns, personalized customer experience, Chatbots, and data-driven decision making to name a few. Despite the enthusiastic embrace of AI marketing by larger organizations, its adoption within small and medium enterprises (SMEs) remains considerably low. SMEs are seriously lagging their large counterparts in harnessing significant business benefits of the latest digital technology applications, especially AI. This paper aims to address the need for an in-depth investigation of the current adoption behavior of SMEs towards AI in marketing; to understand the main barriers preventing them from actively exploring and adopting AI in marketing, and identifying the key role AI providers play in this phenomenon.

Hence, the adoption behavior of SMEs towards AI-based marketing tools, adoption barriers and the role of AI providers were investigated in a qualitative exploratory study through semi-structured interviews of UK SMEs and AI providers. This research-in-progress paper describes the methodology and discusses the preliminary findings and theorical considerations to address the barriers.

2. Literature Review

Extant literature was reviewed to get an in-depth understanding of the adoption behavior of SMEs towards AI in marketing and its influencing factors which revealed a lack of high-quality empirical investigations on this topic. Negnevitsky (2005) describes AI as systems/technology that provide automated and efficient solutions to problems that require intelligence when done by humans. Just as AI technologies have proliferated and optimized virtually every process in every industry at some level ranging from sales, manufacturing, travel, and health service (Davenport et al., 2020) researchers believe the technological advancement in AI, Internet of Things (IoT), Big Data and robotics are imparting profound effects on the marketing practices (Rust, 2020, Grewal et al., 2020), so much so that Wirth (2018) suggests it "would undoubtedly mean the end of market research as we know it." According to Dagenerate al (2020), the impact of AI depends on the type of industry where marketing is most affected due to industries such as consumer goods, retail, banking, and travel in which the businesses are involved in regular contact with customers that generates big amounts of customer transaction data and consumer attribute data.

AI based marketing tools can facilitate marketers to target the right customer at the right time with the right communications by analyzing the customer behavior based on attributes such as age, gender, location and many other data points and making smart recommendations (IAB, 2019). With great benefits to be gained from the adoption of AI marketing it is still imperative for the business decision makers to fully understand how and where to implement AI marketing tools for successful adoption, (Overgoor et al., 2019).

In SMEs the marketing activities and new technology implementation are heavily dependent on the attitudes, personal attributes, knowledge, and capabilities of the owners/managers (Bocconcelli et al., 2018). Scholars assert that decision-making authority is primarily held by owners and managers of SMEs, and their receptiveness and prejudice towards novel technologies are crucial determinants of SMEs' AI marketing adoption (Moeuf et al., 2020). Given the significant influence that owners/managers have as primary decision makers in the adoption and integration of AI technology (Duan et al., 2019), it is imperative to fulfill their requirement for adequate and precise AI knowledge and expertise (Jang et al., 2021). The best source of this knowledge can be the AI providers who are the expert professionals responsible for the designing, development, deployment, and marketing of AI applications. The existing literature does not provide empirical investigations on the adoption of AI marketing in SMEs and what role the AI providers play in the adoption process. In this respect this paper provides first of its kind insights into the adoption behavior of the SMEs towards AI marketing and the barriers, along with exploration of the key role the AI providers play in this paradigm.

3. Methodology

A preliminary study was conducted to gain an understanding around AI marketing, the suitability of existing AI marketing tools for SMEs, and to get initial idea of the perceptions and understanding of the SMEs towards AI marketing adoption. The preliminary study included individual in-depth semi-structured interviews with three AI providers, two SME owners/managers and two case studies of AI marketing use cases. Insights from the preliminary study informed the development of the semi-structured interview questions for the main study. The semi-structured interview questions for SMEs and AI providers were divided into four main themes that are presented in Table 1.

For the main research a three-stage qualitative study was carried out to explore the SMEs' AI adoption behavior, e.g. the current use of AI, their perceptions of AI in marketing, perceived challenges and barriers to the successful adoption, and the role of the AI providers in the adoption process. In stage one a focus group was conducted with SME owners/managers. Stage two comprised of 30 individual in-depth semi-structured interviews with SME owners/managers. The sample population of which was selected from UK based SMEs from retail, wholesale, charity, and service sectors. Each participant was confirmed to be involved in the decision-making process of new technology adoption specifically for marketing within their business. The final sample was a mix of participants that were SME owners, marketing managers, people in senior marketing positions (e.g. CMO, MD) and non-marketing senior decision makers. In stage three of the research in-depth semi-structed interviews were conducted with three individual AI providers involved in designing, developing, vending, and using AI marketing tools.

4. Data Analysis

The interview transcripts were uploaded to NVivo for data analysis. A reflective thematic analysis was carried out adapted from the process outlined by Braun and Clarke (2006). As it is an exploratory study a data driven bottom -up inductive approach was taken to allow themes to emerge from the data. Allowing the researcher to act as an 'excavator' in unveiling nuanced insights on the research phenomenon. In First Cycle coding 'Descriptive Coding' method was used to identify and label key phrases, concept, and segments of the data (Saldaña, 2021) . Next the codes were organized into a broader framework in line with the research questions using 'Structured Coding' method. In Second Cycle 'Pattern Coding' method was used to refine the initial codes to develop candidate themes from the data (Saldaña, 2021) which were later reviewed and renamed to establish final themes after detailed analysis.

5. Initial Empirical Results

The thematic analysis of the data included semantic as well as latent analysis of the interview transcripts. The themes related to the adoption barriers and role of AI providers are discussed in the sections below.

5.1 Barriers

Lack of knowledge and understanding*:* Participants expressed their lack of knowledge and understanding regarding AI marketing. P4 who had a positive attitude towards the AI marketing adoption revealed, *"Yes, I would go ahead with that. Yes, I would, but only providing*

that we are informed" stating that they *"need more training and understanding"* for adoption decision making. Participants A2 and A3 also suggest the lack of awareness amongst SMEs as a lead cause of low adoption rate, as they observed that *"people are talking about it because everyone's talking about it."* [A2] without really understanding what AI marketing is.

Research Instrument	Sample	Interview Themes
Stage 1- Focus Group	SME owners/managers	Knowledge and perception regarding AI marketing Influencing factors (facilitators and barriers) of AI adoption Mitigation of barriers Role of AI providers in the adoption of AI marketing by SMEs
Stage 2- Semi-structured interviews	SME owners/managers	Current marketing activities Knowledge and perception regarding AI marketing Influencing factors (facilitators and barriers) of AI adoption Mitigation of barriers Role of AI providers in the adoption of AI marketing by SMEs
Stage 3- Semi-structured interviews	AI providers	AI marketing Tools SMEs' AI marketing adoption Influencing factors (facilitators and barriers) of AI adoption Mitigation and role of AI providers

Table 1. Research Instruments

The participants also displayed a lack of knowledge regarding AI based marketing tools currently available in the market and their suitability for the size and type of their business. When asked about the current AI based marketing tools P4 replied *"I don't really know, to be honest, because I don't have that much technology here. No, I am lost with that one. To be honest. I don't know. I don't know what I could think of."* Echoing the thoughts of P6 who said, *"What do you mean by latest tools? What kind of thing…I don't even know what Jasper is."* However, amongst the participants that did have some awareness of the latest AI marketing tools, the lack of understanding on how these tools can help their marketing activities, and which tools to adopt, seemed to be affecting their decision-making process. As P6 stated, *"to understand it more would definitely help me make a more informed decision about it, but I don't really know enough about it to really comment too much on it.*" Similarly, P12 also expressed, *"my biggest one is just this whole knowledge around it, and then maybe also, again like the control would be a big one, and also how to go about setting it up."* Technology

adoption is a decision-making process that requires detailed and specific information especially when it comes to complex technology such as AI marketing, which SMEs are currently lacking.

Uncertainty and apprehensions*:* The lack of knowledge and understanding seems to have created ambiguity around the whole concept of AI marketing amongst the participants, giving rise to uncertainty, fear, and apprehensions towards the technology. As P5 explained *"I think a lot of it is fear if you know what I mean, giving control away. It's because the technology isn't fully proven. You are in the early adoption phase"* raising questions such as, "*how would it take that data? What would it do? How would it interact?"* Lack of clear answers to such questions has led to skepticism as to whether the AI marketing tools are suitable for the unique and niche industries the SMEs operate in, and the impact it will have on their customers. P10 voiced their concern as *"I'd be skeptical. And then I'd also be- just in terms of my customers' data going into something. So, I'd have- I'd still be skeptical, especially because we work in such a niche area."* Moving on to say, *"I don't think anyone in the smaller medium business is swayed by something that like, if it works for Waitrose, it doesn't mean it's gonna work for me".*

Lack of resources*:* Participants expressed a cascade of resource limitations, where lack of finances leads to insufficient human resources, leaving them strapped for both manpower and the luxury of time to invest in AI marketing education. Regarding this P9 explained *"Budget is a massive thing as well and resource, taking the time, my time- out of doing my every day [tasks], there's a lot, taking my time out to actually learn about it and how to better use it, and- so I would say I'm probably about 85% on board with using it. Just, time and resource and budget."* P2 echoed similar views, *"Well, I suppose time is one of the most important and knowledge. Access is usually, if you know what you're doing much, you got it. You're there, aren't you? Resources, yeah, which is part of time, isn't it? Part of the same thing. Time would probably be the biggest, and then the knowledge, getting the knowledge to do it."* Tying the lack of resources (finance, human and time) and lack of knowledge to the same string of limitations in the case of SMEs.

Lack of trust in AI*:* Lack of trust in the AI marketing tools was present in almost all participants' responses in one form or the other, such as distrust in its output, the tone of voice of the generated content and allowing full control to the AI without human gate keeping. Participants P7, P8, P13 and P15 all shared the concern of AI generated marketing content not being able to capture the essence and values of their business and lack of personalization leading to a negative impact on their customer relations. Just as P14 voiced their concern *"I don't think we are at a stage yet where we trust machines 100%. So, if it's a recommendation that I can review and just with a click of a button, say yeah, go on then yes, but letting a machine do, making it taking the decisions. Without being able to review that they are actually sending out the contents and the message I like, then no."*

Hesitation and confidence (wait and see): The analysis of the data revealed that although the participants perceived AI as the future, and adoption of AI in marketing to have positive gains, however most participants talked about AI marketing adoption in the future tense showing that they are either at the very initial stages of adoption and have not yet seen conclusive results or they may or may not adopt in it the future. Some participants like P9 seemed positive saying

"we are exploring different avenues of it because I do think it will be beneficial for a small team like us" whereas participant P7 believed *"the AI is the next new thing now. So, it's the next step."* However, *"We are not yet ready to embrace the technology as it is."* P3 also stated that they are *"always pro automation"* and *"absolutely for it"* but they will start *"to adopt [it] slowly, slowly."*

5.2 Role of AI providers

When the participants were inquired about the role of AI providers in potentially mitigating the adoption barriers, most participants expressed a lack of trust in the AI providers. P11 found that the *"the number of people that genuinely know and really have a deep knowledge is probably very, very small. But there's a large number of people that are tasked with getting this technology into businesses and getting new clients working with it."* Likewise, P3 voiced their concern that "*there could always be certain amount of doubt how genuine they are and what they're offering, how transparent they are at what they're offering as well. Because there is always something that you might not be aware of."* P6 was frustrated that *"9 times out of 10, not just for marketing, but for anything the people that you're talking to don't actually understand what they're talking about. That's the impression that you get a lot of the time. They're just reading from a script. They're just trying to get a signature…it's a lot of people selling stuff don't really know the detail about what they're selling."* P1 however spoke positively, *"I'm talking about agencies that I work with and I have worked with and I've got two or three very good agencies who I know. I know exactly when they talk to me about something I know and understand it."* They went on to say that they "*would only use an agency house [they were] very confident and very trusting of."* The AI providers were inquired regarding the influencing factors of AI marketing adoption in SMEs, to investigate whether they were aware of the current challenges SMEs faced and their perceptions towards the AI providers themselves. It was observed that the AI providers were to some extent aware of the lack of knowledge and understanding regarding AI marketing amongst the SMEs, however SMEs' lack of trust in AI provider was not mentioned. Bringing to conclusion that there exists a communication gap between the two parties.

6. Work in progress

The preliminary findings reveal that managers show a serious lack of knowledge and understanding on AI in general and AI in marketing in particular. Consequently, the lack of knowledge and understanding leads to the lack of trust and confidence, uncertainty, apprehension, and hesitation. These factors form major barriers to AI adoption and significantly affect managers AI adoption attitudes and behavior. The findings also uncover a significant communication mismatch between AI providers and SME managers due to lack of trust and the knowledge gap between the two parties. To improve SMEs' AI marketing adoption, it is imperative to improve SME managers' level of knowledge and understanding effectively by better understanding the role of AI providers. Although the investigation provides rich empirical data, the next stage of the study aims to make sense of the findings by mapping out the key themes using a suitable theoretical lens. Two relevant theoretical lenses from the psychology/sociology theories are currently considered. One is the Fear of the unknown (FOTU) (Carleton, 2016) which is defined as, "an individual's propensity to experience fear caused by the perceived absence of information at any level of consciousness or point of processing". Fear of the unknown may be a, or possibly the fundamental fear

impacting the behavior and decision-making of an individual. Another is the Uncertainty Reduction Theory (Berger and Calabrese, 1975). This theory proposes that there is a need to gain information through communication to reduce uncertainty to be able to predict and explain behavior of individuals. Underpinned by these relevant theoretical lenses, it is hoped that a theoretical framework will be developed to help address the AI marketing adoption knowledge gaps focusing on the role of AI providers.

7. Conclusion

The preliminary results of the research-in-progress determine that lack of knowledge and understanding has many implications on the adoption and integration of AI marketing in SMEs. However, SMEs' desire to adopt AI marketing without the help of professional experts, due to lack of trust in AI providers, may lead to failed attempts at adoption, which can reinforce their uncertainty and distrust towards AI marketing. Although barriers may seem distinct, they often intervene and influence each other. Employing a skilled professional or an external agency to help educate on the right knowledge and understanding can help alleviate the barriers such as uncertainty and lack of trust in AI marketing. The key findings of the empirical investigation suggest that both the SMEs and the AI provider need to recognize their communication gap and make active efforts in developing a trustworthy relationship to gain reciprocal benefits.

References

AGRAWAL, A., GANS, J. S. & GOLDFARB, A. 2017. What to expect from artificial intelligence. *MIT Sloan Management Review*, 58**,** 23-26.

BERGER, C. R. & CALABRESE, R. J. 1975. SOME EXPLORATIONS IN INITIAL INTERACTION AND BEYOND: TOWARD A DEVELOPMENTAL THEORY OF INTERPERSONAL COMMUNICATION. *Human Communication Research,* 1**,** 99-112.

BOCCONCELLI, R., CIOPPI, M., FORTEZZA, F., FRANCIONI, B., PAGANO, A., SAVELLI, E. & SPLENDIANI, S. 2018. SMEs and Marketing: A Systematic Literature Review. *International Journal of Management Reviews***,** 227.

BRAUN, V. & CLARKE, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology,* 3**,** 77-101.

CARLETON, R. N. 2016. Fear of the unknown: One fear to rule them all? *Journal of Anxiety Disorders,* 41**,** 5-21.

DAVENPORT, T., GUHA, A., GREWAL, D. & BRESSGOTT, T. 2020. How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science,* 48**,** 24-42.

DUAN, Y., EDWARDS, J. S. & DWIVEDI, Y. K. 2019. Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management,* 48**,** 63-71.

GREWAL, D., HULLAND, J., KOPALLE, P. K. & KARAHANNA, E. 2020. The future of technology and marketing: a multidisciplinary perspective. *Journal of the Academy of Marketing Science,* 48**,** 1-8.

IAB. 2019. *Artificial Intelligence in Marketing Report* [Online]. Interactive Advertising Bureau (IAB). Available: https://www.iab.com/insights/iab-artificial-intelligence-in-marketing/ [Accessed April 11, 2020].

JANG, M., JUNG, Y. & KIM, S. 2021. Investigating managers' understanding of chatbots in the Korean financial industry. *Computers in Human Behavior,* 120**,** 106747.

MOEUF, A., LAMOURI, S., PELLERIN, R., TAMAYO-GIRALDO, S., TOBON-VALENCIA, E. & EBURDY, R. 2020. Identification of critical success factors, risks and opportunities of Industry 4.0 in SMEs. *International Journal of Production Research,* 58**,** 1384-1400.

NEGNEVITSKY, M. 2005. *Artificial intelligence: a guide to intelligent systems*, Pearson education.

OVERGOOR, G., CHICA, M., RAND, W. & WEISHAMPEL, A. 2019. Letting the Computers Take Over: Using AI to Solve Marketing Problems. *California Management Review,* 61**,** 156-185.

RUST, R. T. 2020. The future of marketing. *International Journal of Research in Marketing,* 37**,** 15-26.

SALDAÑA, J. 2021. *The coding manual for qualitative researchers*, sage.
WIRTH, N. 2018. Hello marketing, what can artificial intelligence help you with? *International Journal of Market Research*, 60**,** 435-438.

9. Effective Churn Prediction in the Online Payment Sector: A Survival Analysis

Gözde Özalpay
Boğaziçi University, BIS
gozde.ozalpay@bogaziçi.edu.tr

Nazım Taşkın
Boğaziçi University, BIS
nazim.taskin@bogazici.edu.tr

Abstract

In the dynamic and competitive domain of financial technology, particularly within the online payment sector, the phenomenon of customer churn presents a significant challenge, necessitating advanced predictive strategies for sustainable customer relationship management. This study, centered on a leading payment provider in Turkiye, employs survival analysis to discern key churn determinants. Three distinct models were developed, each integrating variables such as demographics, payment history, and usage patterns, but with varying definitions of churn. Empirical findings from the study highlight the paramount importance of variables, such as changes in commission rates, refund rates, payment counts, merchant types, and sector affiliations in predicting customer churn. A significant insight of the research is the identification of a heightened churn risk post a 25-month customer-company engagement duration. Among the models, the one characterizing churn as a lapse in activity over a one-month interval exhibited superior efficacy. This study underscores the utility of survival analysis in the domain of churn prediction in online payment platforms, offering pivotal insights for strategizing enhanced customer retention approaches.

Keywords: Churn Prediction, Online Payment Systems, Survival Analysis, Fintech, Customer Retention.

1. Introduction

In the evolving digital landscape, online payment systems are vital, continually adapting to meet consumer demands for safe, convenient transactions. As competition intensifies, customer satisfaction becomes crucial for business success (Garg, Rahman, Qureshi, & Kumar, 2020). Central to maintaining business viability is customer retention, where churn prediction is a key strategic tool. Churn prediction enables businesses to understand and mitigate customer loss. This concept, although varying across business models, generally represents the loss of revenue-generating customers. Unlike subscription-based models, where customer departure is more explicit, the online payment sector faces ambiguity in defining churn due to the non-subscription nature of its services. Churn in the online payment sector is a fairly new topic and can benefit from new approaches to make companies more compatible.

This study addresses this challenge by proposing three distinct definitions of churn events and developing corresponding churn prediction models.First definition takes one-month absence as the sign of churn, second one takes two months and the third definition uses the upper confidence interval of time differences between transactions as a threshold for identifying churn. While choosing a fixed time interval is a well known approach, using the confidence intervals of time differences is a relatively new approach. The models are evaluated to

determine the most effective churn prediction period and influential factors. The study follows a systematic approach, encompassing objective definition, research gap identification, literature review, and methodology explanation, culminating in a comprehensive conclusion.

1.1 Objectives of the Study

This project's primary aim is to develop a churn prediction model for the online payment industry using survival analysis. The main reason for selecting survival analysis as the main algorithm is that traditional machine learning algorithms do not have a time dependent variable and can not reflect the relationship between a time dependent variable and target variable. The first objective is to identify key factors influencing customer churn at the selected online payment provider, achieved through an extensive review of churn literature and analysis of customer data from the company. The second objective involves comparing different churn periods aligned with the company's business strategy and customer behavior. Ultimately, the research seeks to offer insights to assist the online payment business in reducing churn and enhancing customer retention.

1.2 Research Gap

While churn has been extensively studied across various industries, its implications for the online payment sector remain relatively unexplored (Dheekollu et al., 2021) A financial technology firm, one of Türkiye's leading fintech firms, provided the data for this study. The primary business of the company is online payment systems. Online payment is a relatively new business, and our research indicates that churn prediction has not been thoroughly investigated in this sector (Aslan & Asan, 2021). The primary contributions of this study are the identification of characteristics associated with churn and the proposition of various approaches to characterize the churn event in the online payment industry.

1.3 Research Questions

Research questions in the study are as follows:

- What elements affect customer attrition at an online payment provider?
- What is the relationship between customer tenure and the timing of churn in online payment industry?

2. Literature Review

For a long time, businesses have prioritized attracting new customers over satisfying existing customers (Reinartz & Kumar, 2019). Customer retention is becoming a crucial component of research and planning in today's low-growth, fiercely competitive market (Lemon & Verhoef, 2016). Since acquiring new clients can necessitate a sizable expenditure in marketing and sales operations, it is frequently more expensive than keeping existing ones (Kumar, Aksoy, Donkers, & Venkatesan, 2019). According to Moghaddasi and Mojtaba (2019), keeping a customer for ten years is three times more profitable than keeping one for five. Customer retention and satisfaction are crucial for the stability of the firm considering the increasing competition in the market (Larivie're & Van den Poel, 2004). Advantages of customer retention include increased profitability, reduced marketing, costs, and improved customer loyalty (Fornell, Mithas, Morgeson, & Krishnan, 2006).

Churn prediction is a general problem for different sectors and can be held from different perspectives using different methods. Related sectors include telecom, banking, insurance, gaming, retail and e-commerce (Mahajan, Misra, & Mahajan, 2015). Kassem, Hussein, Abdelrahman, & Alsheref (2020) and Khan, Jamwal, & Sepehri (2010) have laid the foundation with their forays into predicting customer churn using machine learning techniques. The study by Lee, Wei, Cheng, & Yang (2012) employed a k-Nearest Neighbor Classifier to identify churners in cellular service providers, leveraging detailed user behavior and service usage data, suggesting that time series of call volume could be a potent indicator of churn. Huang & Kechadi (2013) took a more innovative route with a hybrid model-based learning system that fused supervised and unsupervised methodologies. The comprehensive variables used, including age, gender, and call details, contributed to a model that outperformed its predecessors in churn prediction. Khan, Jamwal, & Sepehri's (2010) study highlighted the lesser significance of demographic variables compared to behavioral data such as tenure, customer complaints, and usage patterns in predicting churn. This cross-sector perspective can offer a broader understanding of customer behavior patterns that might be applicable to the online payment sector. The studies by Kassem et al. (2020) and Khan, Jamwal, & Sepehri (2010) emphasize behavioral data in churn prediction, while Lee et al. (2012) and Huang & Kechadi (2013) incorporate demographic variables along with behavioral data, suggesting a different approach. This contrast in the types of variables prioritized for churn prediction across various sectors indicates further research is needed to reconcile these contrasting findings. It underscores the necessity for additional investigation to understand the comparative impact of demographic and behavioral variables in churn prediction, offering a distinct area for our study to make a meaningful contribution.

Machine learning techniques, such as artificial neural networks, Random Forest, and Extreme Gradient Boosting, are increasingly being applied due to their ability to handle diverse and complex data sets. The importance of network characteristics was underscored by Zhang, Zhu, Xu, & Wan (2012), and Kim, Jun, & Lee (2014), who demonstrated that understanding the social and communication networks could significantly enhance prediction accuracy. In a similar vein, the artificial neural network approach by Keramati et al. (2014) excelled, particularly when analyzing variables related to service usage and customer feedback. This highlights the growing consensus that customer interaction metrics are more telling than static demographic data. The work of Phua et al. (2012), and Mishra & Reddy (2017) further solidified the significance of service usage patterns and customer engagement metrics over traditional billing and demographic data. These studies found advanced ensemble methods like Random Forest to be particularly effective in handling the complex nature of churn. Gordini & Veglio (2017), and Patil et al. (2017) provided insights into B2B E-Commerce and retail churn prediction, with a strong focus on transactional behavior, service interaction, and customer engagement parameters, emphasizing the adaptability of machine learning algorithms like Extreme Gradient Boosting in various commercial contexts.

Survival analysis techniques applied to longitudinal data have shown that the timing of churn is as crucial as the event itself, highlighting the importance of customer interactions and changing environmental factors. Lu (2002), and Van den Poel & Lariviere (2004) contributed to the field with survival analysis techniques, focusing on longitudinal data to predict churn timing. Their studies highlight that while demographic information provides context, it is the

customers' interaction with the product and changing environmental factors that play a crucial role in predicting churn. Lastly, the research by Yiwen, Xiaoyang, Qiwen, Weining, & Fang (2021) uses the data of a company serving in the online payment sector which is the same sector that our study was built on. The study utilized sequential data with LSTM-based techniques to capture the timing characteristics in the data. The churn period used in the study was a one-month period of absence, and variables related to merchant information, transaction flow features, and risk control features were included. Proportion of days without transactions, number of chain stores, average number of transactions, and the sector of the merchant were identified as the most important features in the study.

3. Methodology

3.1 Model Selection

This study adopts a multifaceted approach to churn prediction within the online payment industry, addressing two pivotal questions: the factors influencing customer churn and the interplay between customer tenure and churn timing. Machine learning algorithms, such as decision trees, neural networks, logistic regression, XGBoost, and Random Forests were considered for their strengths in predictive modeling with survival analysis (Mahesh, 2020). To align with these objectives, survival analysis was identified as the optimal methodological framework. Survival analysis, a statistical method traditionally used in medical research, has been adapted to analyze the time-to-event data, in this case, the duration until a customer disengages from a service. Survival analysis was selected due to its effectiveness in modeling non-uniform customer engagement durations and its proficiency in event-time prediction. This approach is particularly suited to the online payment context, where customer interactions are non-uniform and varied in duration (Yiwen et al., 2021). In this study, we will develop three models using survival analysis all using the same variables but with different targets to define churn events. This approach will allow us to examine how varying definitions of churn impact our analysis and predictions. Model 1 will target the upper confidence interval of time differences between transactions, Model 2 will focus on a one-month absence as the churn event, and Model 3 will define churn as a two-month absence.

3.2 Survival Analysis and Survival Models

Survival analysis is versatile and applies to various domains, not just medical research. It is particularly suitable for examining phenomena such as churn, where the interest lies in predicting the time until an event (e.g., customer churn) occurs based on historical data (Leung, Elashoff, & Afifi, 1997). The study uses survival models, which are classified into non-parametric, parametric, and semi-parametric models, each with different assumptions about the distribution of survival times and the effect of covariates on the hazard function. For example; Kaplan and Meier proposed a non-parametric model in 1958 to check if the survival probability differs between sub-populations (Kaplan & Meier, 1958). This approach is frequently employed when estimating the survival curve from observed data without assuming the underlying distribution. Parametric survival models have the assumption that survival time follows a known distribution such as Weibull, exponential, log-logistic, log-normal, or generalized gamma (Kleinbaum & Klein, 2012).

The Cox Proportional Hazards model is a semi-parametric approach, is employed to relate the covariates' influence on the hazard function, which represents the instantaneous risk of the

event. Cox Proportional Hazards model implies that the input variables' effects are additive on a single scale and constant throughout time (Grambsch & Therneau, 1994). Variables used in the model are classified into three types:

- Event: Gets the value of one if the event of interest has occurred and value of zero if it has not.
- Time: Time to event.
- Predictors: Features related to the event.

In cox regression, the probability of an event occurring is referred to as the hazard. Hazard is represented as follows:

$$h(t) = h_0(t)\ exp\ (b_1X_1 + b_2X_2 + \cdots + bpXp) \quad (1)$$

where X_1, X_2 … Xp are the predictor variables used in the model and b_1, b_2 … b_p are the coefficients of these predictors and $h_0(t)$ shows the baseline hazard. When the model does not have any predictor variables, hazard is equal to the baseline hazard value. Survival probability can be calculated according to the formula below.

$$S(t) = exp\ (-h_0(t)\ .\ (b_1X_1 + b_2X_2 + \cdots + bpXp)) \quad (2)$$

Where S(t) represents the survival probability $-h_0(t)$ represents the baseline hazard X_1, X_2 … Xp are the predictor variables used in the model and b_1, b_2 … b_p are the coefficients of these predictors. The success metric used in cox regression is called the concordance index (Harrell, Lee, & Mark, 1996). When concordance value gets close to one this means your model is producing more accurate results and can be used to predict the survival probabilities.

3.3 Data Collection

Data for the study is gathered from the private database of an online payment provider in Turkey. Online payment businesses have become increasingly popular in recent years, particularly as e-commerce and online shopping have grown in popularity (Yiwen et al., 2021). The dataset used in this study contains 23,605 observations and 82 variables. The data spans from the beginning of 2020 to 2022 and includes information about transactions made during this period. The time frame selected for this study was based on the superior quality of data available during that period, which is believed to enhance the study's reliability. Data includes the information about all the merchants who have used the platform since the beginning of 2020, including details about their transaction history, payment methods, and other relevant data points. Variables range from categorical inputs such as the type of merchant (standard or marketplace), the source of merchant acquisition (field, online, telesales), settlement periods (daily, weekly, monthly), to the sector in which the merchant operates, and whether the merchant is designated as a key account. We also explore financial dynamics using variables that track monthly commission rates and their recent fluctuations, highlighting the notion that recent rises may be the cause of merchant churn.Transactional aspects are explored through refunded payment counts and volumes, refund rates by transaction and volume, alongside the usage of virtual POS terminals over recent months, offering insights into merchant satisfaction and operational trends. Additionally, sector-specific variables like average and total transaction volumes, merchant counts, and average basket values provide context on market behavior and performance.

3.4 Data Preprocessing Activities

Study adopts a systematic approach to data preprocessing, an essential step for ensuring data integrity and analytical accuracy (Maharana, Mondal, & Nemade, 2022). The initial phase

involved data cleaning, a meticulous process of rectifying inconsistencies and errors within the raw dataset. This included the handling of missing values, a critical issue in statistical analyses. The method of listwise deletion was employed for this purpose, where any records exhibiting missing data were excluded from the analysis. This method was chosen due to the relatively small amount of missing data, which accounted for only 0.25% of the total observations. Outlier detection and elimination were integral to the preprocessing strategy. Outliers, being anomalous data points, can significantly skew the results of statistical models. The study utilized Interquartile Range (IQR) technique to identify and mitigate the influence of these outliers, thereby refining the dataset for more accurate modeling. Furthermore, the treatment of categorical variables was approached through one-hot encoding. This technique effectively transforms categorical data into a binary format, enabling its seamless integration into the machine learning models. The encoding process was crucial for models reliant on numerical inputs and contributed significantly to the enhancement of model performance.

The final preprocessing step involved feature selection. Notably, variables with zero variance, where data points lack variability, are identified and excluded from the study to prevent skewed correlation results and inaccurate conclusions. Following this, correlation coefficients are calculated among the remaining features, adopting a threshold 0.5 or higher to signify strong correlations, based on Cohen's (1988) standards. Variables with strong correlations (coefficients over 0.5) are removed to eliminate redundant information. Standardization techniques are used in data processing to transform raw data into a common scale or format, in order to facilitate meaningful analysis and comparisons. Z-score normalization was used to standardize numerical variables in this study, given its simplicity and usefulness for the cox regression method.

4. Results

4.1 Baseline Hazard

A Cox proportional hazards regression model's baseline hazard function is a crucial element and helpful computation for model understanding. In the absence of any covariates, the baseline hazard function depicts the present-day risk of an occurrence at a specific period.

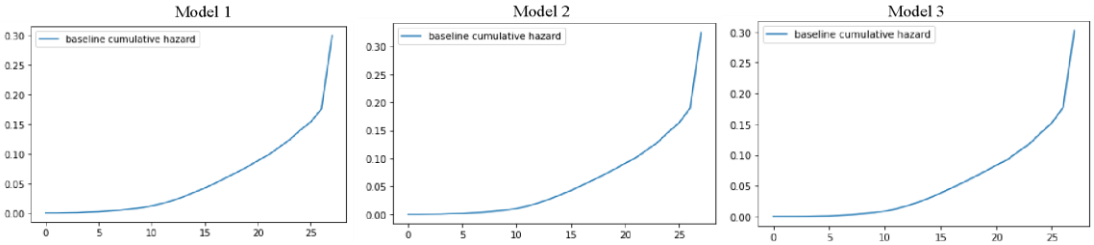


Figure 1: Baseline hazard graphs for models

As seen in Figure 1, the curves are very similar; they are relatively flat at first and then steeply increase. This implies that the risk of the event occurring is low in the early months, but rapidly increases over time, especially when the time value approaches 25. This means that after 25 months of working with the company, the risk of churn increases significantly in comparison to the early months.

4.2 Variable Importance

The analysis revealed that certain variables had a more pronounced influence on churn. These included factors related to transaction volume, frequency, and the nature of customer interactions with the service. Variables, such as the average transaction value, frequency of transactions, and customer engagement metrics emerged as significant predictors. The study also noted the importance of merchant-specific characteristics, such as the type and sector of the merchant, in determining churn likelihood. These findings provide critical insights into the factors that online payment companies should monitor and address to reduce customer churn. By understanding which variables are most impactful, companies can tailor their strategies to enhance customer retention and improve service offerings.

4.3 Model Performance Comparison

The concordance index, usually referred to as the C-index or Harrell's C, is a measurement of a Cox regression model's ability to discriminate across groups of survivors. The percentage of all subject pairs in the dataset for which the projected survival times are appropriately rated in relation to the observed survival times is known as the C-index. It evaluates the model's accuracy in predicting the relative risks of survival outcomes for pairs of individuals (Brentnall & Cuzick, 2018). The concordance index ranges from zero to one, where zero indicates complete disagreement between the model's predictions and the observed events, and one indicates perfect agreement. A C-index of 0.5 indicates a model that performs no better than random chance, while a C-index above 0.5 indicates better-than-random predictive accuracy. Model 1 shows a C-index of 0.81 for training and 0.80 for testing. Models 2 and 3 demonstrate slightly higher accuracy, with both scoring 0.83 in training and 0.82 in testing, suggesting these models have marginally better predictive capability for the defined churn events. Akaike Information Criterion, or AIC, is a statistical metric frequently used in Cox regression models to evaluate the model's goodness-of-fit and simplify model selection. It was developed by Hirotugu Akaike in 1974 as a way to balance a model's complexity and goodness-of-fit while also accounting for the number of estimated parameters. The model with the lowest AIC value is considered to have the best balance between goodness-of-fit and model complexity, and is typically preferred over models with higher AIC values. Model 1 has the lowest AIC value at 50,313, suggesting it fits the data more efficiently than the others. Model 2, with an AIC of 58,409, shows the least efficient fit among the three. Model 3 presents a moderate fit with an AIC value of 53,107, placing it between Models 1 and 2 in terms of data fit efficiency.

Accuracy is defined as the proportion of correctly predicted events (i.e., true positives) and non-events (i.e., true negatives) out of the total number of instances in the dataset. It is often used as a measure of overall model performance (Hastie et al., 2009). The F1 score is the harmonic mean of precision and recall, and it provides a balance between the two. Precision is the proportion of true positives out of the total predicted positives, and recall (also known as sensitivity or true positive rate) is the proportion of true positives out of the total actual positives (Powers, 2020). The F1 score is frequently employed as a gauge of the precision-to-recall trade-off.

The models were evaluated using different thresholds (0.4, 0.5, and 0.6) for accuracy, F1 score, precision, and recall. Model 1 had the highest accuracy for all thresholds in both training and test datasets. However, accuracy might be misleading with imbalanced data, therefore, the F1

score, which combines precision and recall, is also considered. Model 2 excelled in F1 score across all thresholds, indicating a better balance of precision and recall. Precision measures the model's ability to correctly identify positive cases, while recall assesses its ability to recognize all positive cases. Model 2 surpassed Model 1 and Model 3 in both precision and recall, particularly at higher thresholds. In conclusion, based on the provided tables, Model 2 appears to perform better than Model 1 and Model 3 in terms of predictive accuracy, as reflected by higher CI values, and better balance between precision and recall.

Variable	Model 1 - Hazard Rate	Model 2 - Hazard Rate	Model 3 - Hazard Rate	Variable	Model 1 - Hazard Rate	Model 2 - Hazard Rate	Model 3 - Hazard Rate
working_area:Limited Şirket	0,95	0.97	0.96	refund_volume_tr_l4m	0.93	0.93	0.89
mcc_category:Evcil Hayvan	1.5	1.58	1.64	commission_chg_mar_jan	0.92	0.91	0.92
mcc_category:Ayakkabı / Çanta	1.12	1.1	1.08	commission_rate_mar_log	0.92	0.94	0.94
merchant_source_name:Online	1.1	1.12	1.11	gross_payment_cnt_l1m	0.91	0.89	0.9
working_area:Şahıs Şirketi	1.05	1.07	1.08	refund_payment_cnt_l3m	0.87	0.86	0.83
merchant_source_name:Telesales	1.01	1.03	1.03	refundratebytx_l3m	0.84	0.81	0.81
avg_gross_volume_mcc_l1m	1	1	1	mcc_category:Bilgi Sistemleri Hizmetleri	0.74	1	1
total_gross_volume_mcc_l1m	1	1	1	merchant_type:STANDARD MERCHANT	0.73	0.76	0.8
settlement_period:DAILY	1	1.2	1.17	merchant_source_name:BOS	0.71	0.6	0.64
settlement_period:WEEKLY	1	1.09	1.08	mcc_category:Diğer Hizmetler	0.69	1	1
total_transaction_mcc_l1m	1	1	1	merchant_source_name:BTS	0.68	0.6	0.53
commission_chg_mar_feb	0.95	0.95	0.95	gross_volume_l1m	0.62	0.55	0.57
refund_volume_tr_l3m	0.94	0.94	0.92	mcc_category:Sosyal Hizmet	0.5	0.53	0.53

Table 1: Feature Importances

5. Conclusion and Discussion

The study aimed to develop a churn prediction model for an online payment company using survival analysis. By reviewing literature and analyzing customer data, significant factors contributing to customer churn were identified. The study also compared various churn periods to align with the company's business strategy and customer behavior, ultimately aiming to enhance customer retention. Model 2 emerged as the most accurate predictor of churn, outperforming other models in terms of predictive accuracy, precision, and recall. The study's unique contribution lies in its focus on the online payment sector, an area less explored in churn prediction research. It provided comparative insights with Yiwen et al.'s (2021) study in the same sector, particularly in defining churn and examining multiple churn periods.

		Threshold: 0.4				Threshold: 0.5				Threshold: 0.6			
Model	Data	Accuracy	F1	Precision	Recall	Accuracy	F1	Precision	Recall	Accuracy	F1	Precision	Recall
1	Train	0.85	0.06	0.55	0.03	0.83	0.41	0.44	0.38	0.69	0.47	0.32	0.89
1	Test	0.85	0.05	0.57	0.02	0.83	0.41	0.43	0.4	0.69	0.46	0.31	0.9
2	Train	0.82	0.36	0.53	0.27	0.78	0.6	0.45	0.92	0.64	0.49	0.33	0.97
2	Test	0.82	0.33	0.47	0.25	0.78	0.59	0.43	0.93	0.64	0.48	0.32	0.97
3	Train	0.84	0.26	0.53	0.17	0.8	0.56	0.44	0.77	0.66	0.47	0.31	0.97
3	Test	0.84	0.22	0.43	0.14	0.8	0.54	0.42	0.76	0.66	0.47	0.31	0.97

Table2: Performance Metrics comparison

Significant predictors of churn identified include commission rate change, refund rate, payment count, merchant type, and sector. Interestingly, the merchant's working area and settlement period showed no relation to churn, suggesting other factors like sector dynamics might be more influential. The study revealed that merchants from sectors like pet products are more likely to churn, while those in social services tend to maintain longer relationships. Human influence was evident in churn patterns. Merchants acquired through human interaction (BOS and BTS teams) showed higher loyalty compared to those onboarded online. This highlights the importance of personal touch in reducing churn. Merchant type feature shows if the merchant is a standard merchant or marketplace merchant. According to our model, marketplace merchants were found to be more churn-prone than standard merchants, suggesting a need for tailored strategies to address their specific churn risks.This also creates a contrast to Kassem et al.'s (2020) study where type of the customer was not selected as an important feature. Upon comparison, Yiwen et al.'s study and this study demonstrate a utilization of similar feature types. Both research efforts include merchant information, with Yiwen et al. (2021) also focusing on this aspect. Key variables identified in both studies, like the average number of transactions and merchant sector, highlight their significance in churn prediction. Despite the initial selection of different variables, the findings of both studies converge, revealing similarities in the importance of certain features. When it comes to comparing performance metrics. Overall, the study provides valuable insights into churn prediction in the online payment sector, emphasizing the role of industry-specific dynamics and human interaction in customer retention strategies.

5.1 Implications of the Study

This study's results offer significant insights for online payment companies, particularly in strategic policy-making and leadership. Key variables impacting merchant churn, such as commission rate changes, refund rates, payment volumes, and merchant sectors, were identified. Companies can leverage these findings to tailor strategies aimed at reducing churn rates and enhancing customer retention. Actions might include adjusting commission rates, closely monitoring refund rates, and directing marketing efforts toward sectors or merchant types more prone to churn. Additionally, these insights can guide improvements in risk control and customer retention systems. For instance, drawing upon the insights from our survival

analysis companies can optimize their customer onboarding and support strategies. For example, a payment platform could implement a simplified onboarding experience for merchants, supplemented with personalized tutorials and human interaction. Additionally, leveraging the insights into the importance of human interaction in reducing churn, businesses can elevate their customer support by ensuring accessibility and responsiveness, as emphasized by Gartner's research, which found 84% of B2B software buyers prioritize excellent customer support in their renewal decisions. This strategic alignment of onboarding and customer support, grounded in the analytical findings of our study, demonstrates a clear path for companies to mitigate churn through targeted, data-informed customer engagement initiatives.

5.2 Limitations of the Study

This study's limitations include using an imbalanced dataset with a lower proportion of churn cases, potentially affecting model reliability. Techniques like oversampling could correct this imbalance. Conducted with data from a single online payment company, the results might not apply universally. Additionally, the study considered a limited range of variables, possibly omitting other significant factors, and did not account for external influences like market or economic changes. Despite these constraints, the research provides useful insights into merchant churn in the online payment sector and demonstrates the value of machine learning in predicting churn and shaping retention strategies.

5.3 Future Studies

The study raises several important avenues for future research. Exploring additional variables not covered in this study, like risk control features or geographic factors, could offer deeper insights. Employing techniques like Synthetic Minority Over-sampling Technique (SMOTE) or Adaptive Synthetic (ADASYN) can help to balance the imbalanced dataset. Replicating the study with data from other online payment companies would help establish the generalizability of the findings. Studies findings highlight the potential importance of pricing strategies and customer engagement in maintaining loyalty, pointing to industry-wide trends that merit further investigation. On this matter an appropriate research question would be ‘What is the right price for each customer to prevent churn events’. Furthermore, investigating the effectiveness of various machine learning algorithms, or a combination of them, for churn prediction could further enhance strategies for customer retention in the online payment sector. For future research advanced ML techniques like LSTM (Long Short Term Memory) or less complicated approaches like lagging or windowing functions to build time dependent models.

References

Abou el Kassem, E., Hussein, S. A., Abdelrahman, A. M., & Alsheref, F. K. (2020). Customer churn prediction model and identifying features to increase customer retention based on user generated content. International Journal of Advanced Computer Science and Applications, 11(5).

Alberts, L. J. S. M., & Westra, D. R. (2006). Churn Prediction in the Mobile Telecommunications Industry: An application of Survival Analysis in Data Mining. Engineering & computer science, Maastricht University, Department of General Sciences.

Ali, Ö. G., & Arıtürk, U. (2014). Dynamic churn prediction framework with more effective use of rare event data: The case of private banking. Expert Systems with Applications, 41(17), 7889-7903.

Aslan, D., & Asan, U. (2022). Churn prediction in the payment services industry: an application at token financial technologies for IoT devices. In Industrial Engineering in the Internet-of-Things World: Selected Papers from the Virtual Global Joint Conference on Industrial Engineering and Its Application Areas GJCIE 2020, August 14–15, 2020 (pp. 317-333). Springer International Publishing.

Brentnall, A. R., & Cuzick, J. (2018). Use of the concordance index for predictors of censored survival data. Statistical methods in medical research, 27(8), 2359-2373.

Dheekollu, L., Wadhwa, H., Vimal, S., Gupta, A., Asthana, S., Arora, A., & Gupta, S. (2021, December). Modeling approaches for Silent Attrition prediction in Payment networks. In 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 409-414). IEEE.

Fornell, C., Mithas, S., Morgeson III, F. V., & Krishnan, M. S. (2006). Customer satisfaction and stock prices: High returns low risk. Journal of marketing, 70(1), 3-14.

Garg, R., Rahman, Z., Qureshi, M. N., & Kumar, I. (2012). Identifying and ranking critical success factors of customer experience in banks: An analytic hierarchy process (AHP) approach. Journal of Modelling in management, 7(2), 201-220.

Gordini, N. & Veglio, V. (2017). Customers churn prediction and marketing retention strategies. An application of support vector machines based on the AUC parameter-selection technique in B2B e-commerce industry. Industrial Marketing Management, 62, 100-107.

Harrell Jr, F. E., Lee, K. L., & Mark, D. B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in medicine, 15(4), 361-387.

Huang, Y., & Kechadi, T. (2013). An effective hybrid learning system for telecommunication churn prediction. Expert Systems with Applications, 40(14), 5635-5647.

Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecom churn management. Expert Systems with Applications, 31(3), 515-524.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, *53*(282), 457-481.

Keramati, A., Jafari-Marandi, R., Aliannejadi, M., Ahmadian, I., Mozaffari, M., & Abbasi, U. (2014). Improved churn prediction in telecommunication industry using data mining techniques. Applied Soft Computing, 24, 994-1012.

Khan, A. A., Jamwal, S., & Sepehri, M. M. (2010). Applying data mining to customer churn prediction in an internet service provider. International Journal of Computer Applications, 9(7), 8-14.

Kleinbaum, D. G., Klein, M., Kleinbaum, D. G., & Klein, M. (2012). Evaluating the proportional hazards assumption. Survival analysis: a self-learning text, 161-200.

Kim, K., Jun, C. H., & Lee, J. (2014). Improved churn prediction in telecommunication industry by analyzing a large network. Expert Systems with Applications, 41(15), 6575-6584.

Kukush, A., Baran, S., Fazekas, I., & Usoltseva, E. (2011). Simultaneous estimation of baseline hazard rate and regression parameters in Cox proportional hazards model with measurement error. Journal of Statistical Research, 45(2), 77.

Larivière, B. & Van den Poel, D. (2004). Investigating the role of product features in preventing customer churn by using survival analysis and choice modeling: The case of financial services. Expert Systems with Applications, 27(2), 277-285.

Lee, Y. H., Wei, C. P., Cheng, T. H., & Yang, C. T. (2012). Nearest-neighbor-based approach to time-series classification. Decision Support Systems, 53(1), 207-217.

Lemon, K. N. & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. Journal of marketing, 80(6), 69-96.

Leung, K. M., Elashoff, R. M., & Afifi, A. A. (1997). Censoring issues in survival analysis. *Annual review of public health*, *18*(1), 83-104.

Li, X., & Li, Z. (2019). A Hybrid Prediction Model for E-Commerce Customer Churn Based on Logistic Regression and Extreme Gradient Boosting Algorithm. Ingénierie des Systèmes d'Information, 24(5).

Lu, J. (2002). Predicting customer churn in the telecommunications industry—An application of survival analysis modeling using SAS. SAS User Group International (SUGI27) Online Proceedings, 114.

Mahajan, V., Misra, R., & Mahajan, R. (2015). Review of data mining techniques for churn prediction in telecom. Journal of Information and Organizational Sciences, 39(2), 183-197.

Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. Global Transitions Proceedings, 3(1), 91-99.

Mahesh, B. (2020). Machine learning algorithms-a review. International Journal of Science and Research (IJSR).[Internet], 9, 381-386.

Milošević, M., Živić, N., & Andjelković, I. (2017). Early churn prediction with personalized targeting in mobile social games. Expert Systems with Applications, 83, 326-332.

Mishra, A. & Reddy, U. S. (2017, November). A comparative study of customer churn prediction in telecom industry using ensemble based classifiers. In 2017 International conference on inventive computing and informatics (ICICI) (pp. 721-725). IEEE.

Moghaddasi, N. & Heidari, M. (2019). The impact of modern information technology on customer satisfaction considering the mediating role of organizational health. International Journal of Services and Operations Management, 34(3), 279-305.

Patil, A. P., Deepshika, M. P., Mittal, S., Shetty, S., Hiremath, S. S., & Patil, Y. E. (2017, August). Customer churn prediction for retail business. In 2017 International Conference on Energy Communication Data Analytics and Soft Computing (ICECDS) (pp. 845-851).

Phua, C., Cao, H., Gomes, J. B., & Nguyen, M. N. (2012). Predicting near-future churners and win-backs in the telecommunications industry. arXiv preprint arXiv:1210.6891.

Pondel, M., Wuczyński, M., Gryncewicz, W., Łysik, Ł., Hernes, M., Rot, A., & Kozina, A. (2021, July). Deep learning for customer churn prediction in e-commerce decision support. In Business Information Systems (pp. 3-12).

Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv preprint arXiv:2010.16061.

Reinartz, W. & Kumar, V. I. S. W. A. N. A. T. H. A. N. (2002). The mismanagement of customer loyalty. Harvard business review, 80(7), 86-94.

Therneau, T. M., Grambsch, P. M., Therneau, T. M., & Grambsch, P. M. (2000). Testing proportional hazards. Modeling survival data: extending the Cox model, 127-152.

Tsai, C. F., & Chen, M. Y. (2010). Variable selection by association rules for customer churn prediction of multimedia on demand. Expert Systems with Applications, 37(3), 2006-2015.

Tsai, C. F., & Lu, Y. H. (2009). Customer churn prediction by hybrid neural networks. Expert Systems with Applications, 36(10), 12547-12553.

Van den Poel, D., & Lariviere, B. (2004). Customer attrition analysis for financial services using proportional hazard models. European journal of operational research, 157(1), 196-217.

Vafeiadis, T., Diamantaras, K. I., Sarigiannidis, G., & Chatzisavvas, K. C. (2015). A comparison of machine learning techniques for customer churn prediction. Simulation Modelling Practice and Theory, 55, 1-9.

Zhang, X., Zhu, J., Xu, S., & Wan, Y. (2012). Predicting customer churn through interpersonal influence. Knowledge-Based Systems, 28, 97-104.

Yiwen, X. U., Xiaoyang, L. I., Qiwen, D. O. N. G., Weining, Q. I. A. N., & Fang, Z. H. O. U. Merchant churn prediction based on transaction data of aggregate payment platform. Journal of East China Normal University (Natural Science), 2020(5), 167.

10. Exploring the Intersection of Large Language Models (LLMs) and Explainable AI (XAI): A Systematic Literature Review
(Research-in-Progress)

Busra Sebin
Bogazici University
busra.sebin@std.boun.edu.tr

Nazim Taskin
Bogazici University
nazim.taskin@boun.edu.tr

Nijat Mehdiyev
Saarland University
nijat.mehdiyev@dfki.de

Abstract
In the evolving domain of natural language processing (NLP), the emergence of Large Language Models (LLMs) like Generative Pre-trained Transformer (GPT), Bidirectional Encoder Representations from Transformers (BERT), and their advanced successors has markedly transformed text generation and comprehension. These models, leveraging deep learning and transformer architectures, have significantly enhanced NLP capabilities. Concurrently, the importance of Explainable AI (XAI) in elucidating the complexities of AI models, particularly LLMs, has gained prominence, especially for decision-making processes. This research-in-progress paper, structured under the PRISMA guidelines, provides initial insights into the confluence of LLMs and XAI. It explores two critical inquiries: (1) Which XAI methods are utilized to understand LLMs? and (2) How are LLMs leveraged for their explanatory capabilities? Publications from the four essential databases—Web of Science, Scopus, Arxiv, and IEEE Explore—are processed for this systematic review. Initially comprising 96 articles, the selection was refined to 21, focusing on studies that integrate LLMs and XAI with empirical findings. This review highlights current XAI research primarily relies on agnostic, local, quantitative methods like SHAP and LIME, often combined, to comprehend AI decision-making. LLMs like ChatGPT excel in providing explanation of XAI outputs*, revealing a gap in direct LLM utilization as explainer methods.*

Keywords: Large Language Models, Explainable AI, ChatGPT, SHAP, Systematic Literature Review

1. Introduction
The introduction of the Transformer architecture has significantly advanced Large Language Models (LLMs) such as the Generative Pre-trained Transformer (GPT) and Bidirectional Encoder Representations from Transformers (BERT), revolutionizing machine language processing (Harrer et al., 2023). These LLMs, grounded in deep learning, often function as ‘black boxes’ obscuring the processes behind their decision-making. This opacity is a critical concern as LLMs increasingly influence important decisions, where transparency and reliability are paramount (Islam et al., 2022). In contrast to LLMs, simpler models like regressions and decision trees offer greater interpretability but lack the capability to process complex, unstructured data effectively (Malhotra & Jindal, 2023). This dichotomy between

performance and interpretability has underscored the importance of Explainable Artificial Intelligence (XAI). Since 2017, XAI has aimed to demystify the workings of complex ML models, including LLMs (Islam et al., 2022). The urgency for XAI is heightened by LLMs' potential to produce erroneous or 'hallucinatory' outputs, which can inadvertently spread misinformation (Ji et al., 2023). Thus, XAI emerges as an essential tool in ensuring that the deployment of LLMs is both responsible and trustworthy.

However, transparency, a key aspect of responsible AI, remains underexplored in discussions on LLMs (Liao & Vaughan, 2023). To the best of our knowledge, this is the first systematic literature review addresses this pivotal aspect of AI by exploring LLMs in the XAI context, examining their roles as both subjects of other XAI methods and model explainers.

2. Methodology

This systematic review adheres to the 2020 PRISMA guidelines (Page et al., 2021). The main focuses of the research-in-progress study are as follows: the development of XAI methodologies aimed at elucidating the complex internal mechanisms of LLMs; and the utilization of LLMs as instrumental explanation tools.

We employ a thorough literature search strategy across four key academic databases: Web of Science (WoS), Scopus, Arxiv, and IEEE Xplore Digital Library. The search, focused on XAI within the context of LLMs, utilized carefully selected keywords. WoS (Birkle et al., 2020) and Scopus (Boyle & Sherman, 2006) were chosen for their extensive bibliographic collections and academic influence, while IEEE Xplore's vast repository offering over 6 million documents in technology and engineering and Arxiv's role as a leading open-access platform (Ginsparg, 2021) were pivotal for capturing the latest and most innovative researches enabled via the arXiv python API. This approach ensured a comprehensive and up-to-date review, making it indispensable for an in-depth exploration of LLMs in these fields.

Each keyword within a group is combined using the OR operator, and different groups are linked with the AND operator to create an extensive search string. The final line of Table 1 shows how keywords from various groups are amalgamated to construct a search query deployed across all four bibliographic databases, a search of titles, abstracts, and indexing terms for all available years since all relevant papers in this field are of recent publication.

The keywords for query formulation is displayed below (see Figure 1):

Group 1: Explainable AI Related Keywords	Group 2: Large Language Model Related Keywords
XAI OR explainable AI OR explainable artificial intelligence OR black-box OR post-hoc explainability OR instrinsic explainability OR counterfactual approximation OR counterfactual explanation OR feature importance	LLM OR chatgpt OR GPT-* OR generative AI OR large language model* OR pretrained language model*
Search Query: (GROUP 1) AND (GROUP 2)	

Table 1: Selected keywords in different groups.

We established criteria for both inclusion and exclusion to ensure a focused and relevant review. This rigorous approach allowed us to refine the selection process, securing a collection of articles that precisely aligned with our research aims.

Details of inclusion and exclusion criteria are listed in Table 2:

Inclusion criteria
1) The paper claims that an LLM and XAI method used.
2) The paper is published in either a journal or conference proceedings in selected databases except arxiv.
3) The paper is in English.
4) The paper with accessible full text.
Exclusion criteria
1) The paper mentioning the used of LLMs without describing an XAI method.
2) The paper applying LLMs for purposes non-related with XAI.
3) The paper released in a workshop.
4) The paper without a clear demonstration of results for LLMs.

Table 2: Inclusion criteria and Exclusion criteria.

The PRISMA search methodology is displayed in Figure 1:

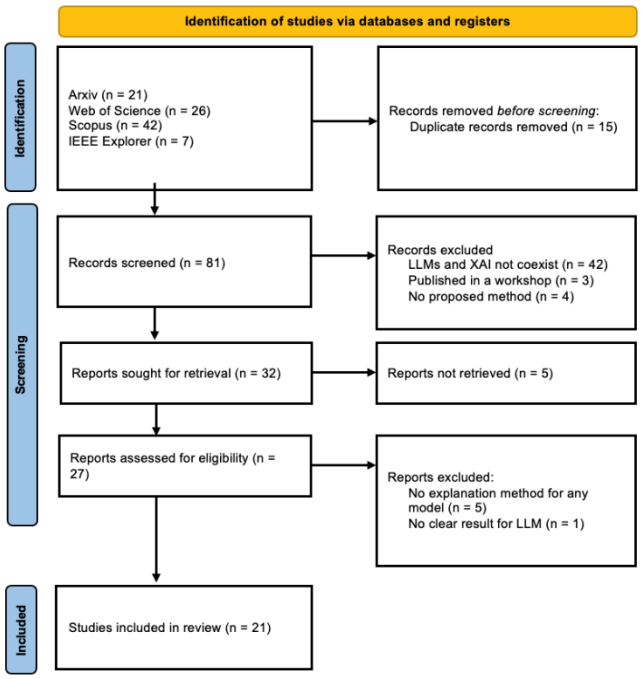


Figure 1: PRISMA search methodology.
(Adapted from Page et al., 2021)

The search across these databases yielded a total of 96 studies based on the selected keywords. Duplicate entries from different databases were meticulously examined, retaining only unique records for each primary sample. During this duplicate removal process, 15 studies were excluded. Through a manual selection process focusing on titles and abstracts, the initial set of 81 studies was narrowed down to 32 unique and pertinent articles. Then, the count was further

reduced to 27 papers, since access to 5 articles could not be retrived. Ultimately, after excluding 6 articles based on the defined exclusion criteria through full-text review, 21 articles remained for consideration.

3. Results

We initially would like to introduce a structured summary of recent scholarly contributions at the intersection of LLMs and XAI to identify which XAI methods utilized to explain the inner working of LLMs and what roles LLMs taken in this studies.

The summary of recent studies for the SLR is displayed in Table 3:

Author & Year	XAI Method	Utilized LLM	Role of LLM
Ali & Kostakos, 2023	Local Interpretable Model-agnostic Explanations (LIME), Shapley Additive Explanations (SHAP)	ChatGPT (GPT-3.5)	XAI Output Translator
Bhattacharjee et al., 2023	LLM, Counterfactual Explanation (CFE)	text-davinci-003, ChatGPT, GPT-4, LLaMA2 7B, LLaMA2 13B - DistilBERT	Explanation Generation, LLM as Subject Model
Cha & Lee, 2023	SHAP	BART	LLM as Subject Model
Chang et al., 2023	LIME	GPT-2, MASS, BART	LLM as Subject Model
Chen et al., 2023	Knowlegde Graph (KG), Graph Attention Network (GAT) for why-choose, why-not choose explanations, and reason elements	ROBERTA - GPT 3.5, GPT-4	LLM as Subject Model, XAI Output Translator
Demertzis et al., 2023	SHAP, CFE	ChatGPT	XAI Output Translator
Feldhus et al., 2023	Saliency Map	GPT-3.5	XAI Output Translator
Feleki et al., 2023	Heatmap	GPT-3.5	XAI Output Translator
Gat et al., 2023	LLM, CFE	ChatGPT, T5 - DistilBERT, BERT, ROBERTA, Llama2-7b, Llama2-13b	Explanation Generation, LLM as Subject Model
Kroeger et al., 2023	LLM, LIME, SHAP	ChatGPT (GPT-3.5, GPT-4)	Explanation Generation
Lucchese et al., 2023	Embedding Analysis for distance between query and output token embeddings	BERT, ROBERTA, ELECTRA	LLM as Subject Model
Malhotra & Jindal, 2023	LIME, SHAP	BERT, DistilBERT, RoBERTa, MentalBERT, PsychBERT, PHSBERT	LLM as Subject Model
Martens et al., 2023	SHAP, CFE	ChatGPT (GPT-4)	XAI Output Translator
Metzger et al., 2023	Explainable Reinforcement Learning for Decomposed Interestingness Elements (XRL-DINE)	ChatGPT (GPT-3.5)	XAI Output Translator
Mitrović et al., 2023	SHAP	DistilBERT, GPT-2	LLM as Subject Model
Nazary et al., 2023	Feature Importance	ChatGPT (GPT-3.5)	LLM as Subject Model
Sarti et al., 2023	Feature Importance (Perturbation-based, Internal-based, Gradient-based)	GPT-2	LLM as Subject Model
Susnjak, 2023	SHAP, CFE	ChatGPT	XAI Output Translator
Tursun et al., 2023	Heatmap	ChatGPT	XAI Output Translator
Wellawatte & Schwaller, 2023	LIME, SHAP	ChatGPT (GPT-4)	XAI Output Translator
Yang et al., 2023	Human Knowledge	GPT-3	LLM as Subject Model

Table 3: Exploring the Intersection of LLMs and XAI: A Summary of Recent Research.

3.1 Understanding the Inner Working of the LLMs

In the landscape of AI research, two primary explanatory approaches are recognized: local, focusing on individual model decisions, and global, which elucidates a model's overall behavior. SHAP is particularly noteworthy for its ability to offer both local and global insights (Wellawatte & Schwaller, 2023). Its game theory foundation ensures an equitable attribution of contribution across features, making SHAP a highly effective tool for delving into the nuanced dynamics of AI models (Mitrović et al., 2023). In contrast, LIME and CFE are tailored towards local interpretability, offering clarity on specific predictions and hypothetical decision

alterations, respectively. Despite SHAP's efficacy as a standalone tool—employed independently in 45% of studies—the trend towards integrating it with methods like LIME and CFE is emerging in about 40% of research. This integration amplifies the interpretability scope by combining SHAP's broad analytical reach with the detailed, scenario-specific insights provided by LIME and CFE. Such a composite approach not only enhances the granularity of local explanations but also enriches the understanding of global model dynamics.

Most of the employed methods are local explainers which focus on providing explanations for individual predictions. This trend towards local explanations may stem from the growing need to understand specific decision-making instances in AI, where individual decisions can have significant consequences (Martens et al., 2023). Moreover, methods like SHAP, LIME, CFE, and KG, which tend to offer more quantitative insights, are prevalent. This preference could be due to the increasing demand for measurable and data-driven explanations in AI, especially in research and development environments where quantifiable insights are crucial for model evaluation and improvement. Additionally, model-agnostic methods like LIME and SHAP are broadly applicable across various models, favored in over 40% of studies for their flexibility and universal insights. Conversely, model-specific techniques, such as KG, GAT, heatmaps, offer precise, tailored insights into the specific workings of particular model architectures (Tursun et al., 2023). This highlights AI research's dual needs: for versatile tools that broadly interpret AI behaviors and for specialized methods that deeply analyze specific models.

Moreover, the analysis of 21 articles on XAI methods shows that while around approximately 48% did not specify their target group, likely indicating a more general audience, the specified audiences predominantly cater to non-technical groups such as judicial officers, students, chemist, and healthcare practitioners, with a smaller portion targeting technical professionals like programmers and data scientists.

3.2 LLMs as Explainers

In the XAI realm, LLMs such as ChatGPT play a pivotal role in translating complex XAI method outputs into digestible narratives, bridging the gap between intricate analytical results and their practical implications. This translation is key for stakeholders without technical expertise, fostering transparency and trust in AI systems (Demertzis et al., 2023; Martens et al., 2023). These models proficiently convert intricate data from approaches like heatmaps or Shapley value distributions into lucid explanations embedded with contextual understanding, thereby making the AI decision-making process more intelligible.

Furthermore, ChatGPT extends beyond translation to actively create explanations. It generates counterfactual narratives and isolates the most influential features affecting the model's decisions, offering insight into the model's rationale (Bhattacharjee et al., 2023). Moreover, LLMs themselves are now under XAI investigation, with BERT and GPT variants undergoing scrutiny. Applying XAI techniques to these models is essential for maintaining transparency and trust, particularly in sensitive sectors like healthcare (Yang et al., 2023; Malhotra & Jindal, 2023). While LLMs are predominantly engaged in translating XAI outputs and are increasingly the focus of XAI studies, the direct application of XAI tools to analyze LLMs signals a shift towards a deeper comprehension of AI internals, as crucial as the external applications. This trend underscores the dual utility of LLMs: as pragmatic tools for explanation generation and

as subjects for in-depth analytical examination, promoting the development of transparent and interpretable AI systems. Despite their efficacy in interpreting outputs from various XAI techniques, the exploration of LLMs as independent explanators has not been fully realized, indicating a prospective area for growth within XAI research (Kroeger et al., 2023; Bhattacharjee et al., 2023).

4. Conclusion and Future Work

Based on this research-in-progress, it is evident that current research in XAI predominantly relies on agnostic, local, quantitative methods like SHAP and LIME, often in combination, to delve into the intricacies of AI decision-making processes. Moreover, LLMs, such as ChatGPT play a pivotal role in providing XAI output translator, making complex AI outputs accessible. However, a noticeable gap remains in the direct utilization of LLMs as explainer methods in explaining other predictive models. To further enhance the development of transparent AI systems, future work will expand its scope using the forward-backward method and additional dimensions such as application areas, XAI evaluation metrics, and prompting techniques.

References

Ali, T., & Kostakos, P. (2023). HuntGPT: Integrating Machine Learning-Based Anomaly Detection and Explainable AI with Large Language Models (LLMs). http://arxiv.org/abs/2309.16021

Bhattacharjee, A., Moraffah, R., Garland, J., & Liu, H. (2023). LLMs as Counterfactual Explanation Modules: Can ChatGPT Explain Black-box Text Classifiers? http://arxiv.org/abs/2309.13340

Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). "Web of Science as a data source for research on scientific and scholarly activity." *Quantitative Science Studies, 1*(1), 363-376.

Boyle, F., & Sherman, D. (2006). "ScopusTM: The Product and Its Development." *The Serials Librarian*, *49*(3), 147–153.

Cha, Y., & Lee, Y. (2024). "Advanced sentence-embedding method considering token importance based on explainable artificial intelligence and text summarization model." *Neurocomputing*, *564*, 126987.

Chang, B. R., Tsai, H. F., & Chou, H. L. (2023). "Implementation of Rapid Code Transformation Process Using Deep Learning Approaches." *CMES - Computer Modeling in Engineering and Sciences*, *136*(1), 107–134.

Chen, Z., Chen, J., Gaidhani, M., Singh, A., & Sra, M. (2023). XplainLLM: A QA Explanation Dataset for Understanding LLM Decision-Making. http://arxiv.org/abs/2311.08614

Demertzis, K., Rantos, K., Magafas, L., Skianis, C., & Iliadis, L. (2023). "A Secure and Privacy-Preserving Blockchain-Based XAI-Justice System." *Information (Switzerland), 14*(9), 477.

Feldhus, N., Hennig, L., Nasert, M. D., Ebert, C., Schwarzenberg, R., & Möller, S. (2022). Saliency Map Verbalization: Comparing Feature Importance Representations from Model-free and Instruction-based Methods. http://arxiv.org/abs/2210.07222

Feleki, A., Apostolopoulos, I. D., Moustakidis, S., Papageorgiou, E. I., Papathanasiou, N., Apostolopoulos, D., & Papandrianos, N. (2023). "Explainable Deep Fuzzy Cognitive Map Diagnosis of Coronary Artery Disease: Integrating Myocardial Perfusion Imaging, Clinical Data, and Natural Language Insights." *Applied Sciences, 13*(21), 11953.

Gat, Y., Calderon, N., Feder, A., Chapanin, A., Sharma, A., & Reichart, R. (2023). Faithful Explanations of Black-box NLP Models Using LLM-generated Counterfactuals. http://arxiv.org/abs/2310.00603

Ginsparg, P. (2021). "Lessons from arxiv's 30 years of information sharing." *Nature Reviews Physics*, *3*(9), 602–603.

Harrer, S. (2023). "Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine." *EBioMedicine, 90*.

Islam, M. R., Ahmed, M. U., Barua, S., & Begum, S. (2022). "A systematic review of explainable artificial intelligence in terms of different application domains and tasks." *Applied Sciences, 12*(3), 1353.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). "Survey of hallucination in natural language generation." *ACM Computing Surveys, 55*(12), 1-38.

Kroeger, N., Ley, D., Krishna, S., Agarwal, C., & Lakkaraju, H. (2023). Are Large Language Models Post Hoc Explainers? http://arxiv.org/abs/2310.05797

Liao, Q. V., & Vaughan, J. W. (2023). AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap. http://arxiv.org/abs/2306.01941

Lucchese, C., Orlando, S., Minello, G., Perego, R., Nardini, F. M., & Veneri, A. (2023). Can Embeddings Analysis Explain Large Language Model Ranking? *International Conference on Information and Knowledge Management, Proceedings,* 4150–4154.

Malhotra, A., & Jindal, R. (2024). "XAI Transformer based Approach for Interpreting Depressed and Suicidal User Behavior on Online Social Networks." *Cognitive Systems Research, 84,* 101186.

Martens, D., Dams, C., Hinns, J., & Vergouwen, M. (2023). Tell Me a Story! Narrative-Driven XAI with Large Language Models. http://arxiv.org/abs/2309.17057

Metzger, A., Bartel, J., & Laufer, J. (2023). An AI Chatbot for Explaining Deep Reinforcement Learning Decisions of Service-oriented Systems. http://arxiv.org/abs/2309.14391

Mitrović, S., Andreoletti, D., & Ayoub, O. (2023). ChatGPT or Human? Detect and Explain. Explaining Decisions of Machine Learning Model for Detecting Short ChatGPT-generated Text. http://arxiv.org/abs/2301.13852

Nazary, F., Deldjoo, Y., & di Noia, T. (2023). ChatGPT-HealthPrompt. Harnessing the Power of XAI in Prompt-Based Healthcare Decision Support using ChatGPT. http://arxiv.org/abs/2308.09731

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews." *International Journal of Surgery*, *88*.

Ramon, Y., Martens, D., Provost, F., & Evgeniou, T. (2020). "A comparison of instance-level counterfactual explanation algorithms for behavioral and textual data: SEDC, lime-C and shap-C." *Advances in Data Analysis and Classification*, *14*(4), 801–819.

Sarti, G., Feldhus, N., Sickert, L., & van der Wal, O. (2023). Inseq: An interpretability toolkit for sequence generation models. arXiv preprint arXiv:2302.13942.

Susnjak, T. (2023). "Beyond Predictive Learning Analytics Modelling and onto Explainable Artificial Intelligence with Prescriptive Analytics and ChatGPT." *International Journal of Artificial Intelligence in Education.*

Tursun, O., Denman, S., Sridharan, S., & Fookes, C. (2023). Towards Self-Explainability of Deep Neural Networks with Heatmap Captioning and Large-Language Models. http://arxiv.org/abs/2304.02202

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.

Wellawatte, G. P., & Schwaller, P. (2023). Extracting human interpretable structure-property relationships in chemistry using XAI and large language models. http://arxiv.org/abs/2311.04047

Yang, Q., Hao, Y., Quan, K., Yang, S., Zhao, Y., Kuleshov, V., & Wang, F. (2023, April 19). Harnessing Biomedical Literature to Calibrate Clinicians' Trust in AI Decision Support Systems. *Conference on Human Factors in Computing Systems - Proceedings*. 1-14.

11. Investigating Higher-Education Students' ICT Skills Learning Experiences: Extending the Technology Acceptance Model

Charlette Donalds
The University of the West Indies, Mona
charlette.donalds02@uwimona.edu.jm

Kweku-Muata Osei-Bryson
Virginia Commonwealth University
kmosei@vcu.edu

Abstract

While much of the literature applying the Technology Acceptance Model (TAM) focuses on the actual adoption of the information communication and technology (ICT) itself, this study differs significantly by focusing on the learning experiences of ICT adoption. The study's aim is to investigate what factors influence university students' learning experiences of ICT-based skills. Given the context-dependent nature of the ICT-based course, the TAM was extended and supplemented with components largely relevant to harnessing the students' learning experiences. The new research model is validated using data collected from 135 higher-education students, pursuing varying majors. The data is analyzed using structural equation modeling. The results suggest that students' expectations to use ICT-based course skills are determined by perceived usefulness, perceived enjoyment, and attitude towards the course. Perceived enjoyment, in turn, is influenced by attitude towards the course and perceived ease of learning. This study offers several contributions: (i) this is the first study that investigates the learning experiences of ICT adoption and not the actual adoption of the ICT itself; (ii) it presents and evaluates a new conceptual model; and (iii) it extends the TAM framework by integrating new constructs that may influence not only ICT adoption but ICT adoption learning experiences.

Keywords: Perceived Ease of Learning, TAM, ICT Skills, Perceived Enjoyment, Expectation to Use.

1. Motivation for the Study

The importance of information and communication technology (ICT) skills in modern economies is well accepted. A strong case can be made that students' performance in learning ICT skills represents a strong indicator of tomorrow's success and status, from the individual professional point of view, as well as from a country's economic point of view. For instance, ICT skills are identified as being increasingly important for individuals to benefit from technological innovations, specifically, better employment chances and higher wages; and in fact, in recent years, of the science, technology, engineering, and mathematics (STEM) fields' employment rates are highest among tertiary-educated individuals (on average 90%) who studied ICT (OECD, 2016, 2022). While the benefits of ICT skills are undeniable, less is known about how to create effective learning experiences that motivate students to acquire and retain these crucial skills. The question is how to best teach ICT-based courses so that students acquire and later use these skills. The last several decades have seen a plethora of research on ICT acceptance and use in higher education (e.g. Alismaiel, Cifuentes-Faura, & Al-Rahmi, 2022; Barclay, Donalds, & Osei-Bryson, 2018; Fathema, Shannon, & Ross, 2015; Lin & Yu,

2023; Mohammadi, 2015), yet an aspect that has been less researched is the learning experiences of ICT adoption by students.

For the ICT and other STEM-like courses, it has been taken as given by many that such courses should be taught in a manner that the average student finds participating in the given course to be an enjoyable experience (e.g. Wu, Deshler, & Fuller, 2018). However, it is important to do appropriate formal evaluation of such assumptions. Thus, this study's aim is to investigate what factors influence university students' learning experiences of ICT-based skills. By investigating the factors influencing students' learning experiences, we can develop more engaging and effective teaching methods that promote long-term ICT skill acquisition and utilization. As Li, Duan, Fu, and Alford (2012) note, it is essential to examine the relationship between e-learners' experiences, perceptions, and their behavioral intentions to use, because system use is an important indicator of the system's success. Similarly, we argue that it is essential to examine the relationship between students' experiences with ICT adoption skills and their expectation to use those skills in the future, because expectation to use may be an essential driver of actual learning the ICT-based skills. This reasoning motivates us to consider that adapting a Technology Acceptance Model (TAM) inspired framework would be appropriate for this study. Specifically, the TAM allows us to consider the perspective that it is the student's attitude (i.e. tendency to respond negatively/positively toward learning) that has a significant impact towards learning and frames the impact of enjoyment and other antecedents of expectation to use the taught ICT-based skills. Prior works support this viewpoint as researchers reveal that attitude is a critical requirement for STEM persistence (e.g. Graham, Frederick, Byars-Winston, Hunter, & Handelsman, 2013). This research has the potential to improve how ICT skills are taught in universities, ultimately preparing students for successful careers in the digital age.

2. Theoretical Background

2.1 Technology Acceptance Model in the Educational Context

The TAM, introduced by Davis (1989), is believed to be one of the most well-researched, robust, and parsimonious model in explaining and predicting ICT acceptance behaviour (Bagozzi, 2007; Cornell, Eining, & Hu, 2011; Venkatesh & Davis, 2000). The TAM model asserts that perceived usefulness and perceived ease of use form an end user's belief about a technology and therefore predict the end user's attitude (ATT) toward the technology, which in turn predicts the behavioural intention (BI) to use and actual ICT usage.

A review of the literature reveals that TAM is applied in many different settings, including the higher education context. In the higher education context, based on a systematic literature review, TAM is identified as a leading scientific paradigm and credible model for facilitating assessment of diverse educational technological deployments (Granić & Marangunić, 2019). For instance, TAM is applied to evaluate students' use of digital reading tools on computers for academic and research purposes (Lin & Yu, 2023), investigate students' behavioral intention to adopt social media in higher education (Al-Qaysi, Mohamad-Nordin, & Al-Emran, 2023), determine what factors influence pre-service teachers' intentions to utilize artificial intelligence (AI)-based educational applications (Zhang, Schießl, Plößl, Hofmann, & Gläser-Zikuda, 2023), determine factors that impact university students' behaviour and intentions to use social media to boost their academic performance during the COVID-19

Pandemic (Alismaiel et al., 2022), determine the factors that influence university students' adoption and use of an e-learning system in the English-speaking Caribbean (Barclay et al., 2018), examine factors that affect faculty members' learning management systems behaviour (Fathema et al., 2015), and investigate students' perceptions about e-learning intention and actual e-learning system usage (Mohammadi, 2015). While the TAM literature is replete with investigations of factors that influence ICT adoption/rejection in the educational context, to our best knowledge, studies have yet to investigate the learning experiences of ICT adoption; this study addresses this gap.

2.2 An ICT Adoption Learning Experience Model

To evaluate technology acceptance in diverse contexts, researchers commonly add to the TAM by introducing additional predictors. Generally, predictors are added to enhance the explanatory power of ICT adoption/use in the various contexts since the original TAM model does not consider context-specific factors that may influence such ICT adoption/use. For example, Lin and Yu (2023) added several predictors to explain and predict higher-education students' acceptance and actual use of digital academic reading tools on computers, to include self-efficacy, lecturers' positive response, ease of access to academic reading materials, perceived ease for collaborative learning, academic experience, and expectation of academic achievement. Zhang et al. (2023) included AI anxiety, job relevance, subjective norm, AI self-efficacy, and perceived enjoyment in their TAM-based model investigating pre-service teachers' acceptance of AI in education. Likewise, Alismaiel et al. (2022) added collaborative learning, student engagement, and perceived enjoyment to explain students' academic performance, facilitated by social media. Similarly, Barclay et al. (2018) added predictors computer self-efficacy, social norm, university climate, instructor engagement, access to computers, and system availability to evaluate the impact of adoption and use of e-learning systems in higher education. This approach is also observed in the other versions of TAM such as TAM 2 (Venkatesh & Davis, 2000) which included factors of social influence and cognitive instrumental processes, the unified theory of acceptance and use of technology (UTAUT) (Venkatesh, Morris, Davis, & Davis, 2003) which integrated factors across eight IT acceptance models, TAM 3 (Venkatesh & Bala, 2008) which integrated several determinants of perceived usefulness and perceived ease of use, and UTAUT2 (Venkatesh, Thong, & Xu, 2012), which incorporated hedonic motivation, price value, and habit. In spite of its various extensions, the original TAM remains popular, due in part to its parsimony (Bagozzi, 2007).

Consistent with prior investigations, this study extends TAM by incorporating new and adapted predictors to propose a new model of ICT adoption learning experiences. Since students' ICT adoption learning experiences have gained little to no attention, to properly interrogate the phenomenon, context-specific predictors that are relevant to harnessing students' ICT learning experiences supplement original TAM predictors. Specifically, this model includes perceived ease of learning (PEL) and expectation to use course skills in future (EXP), new predictors, perceived enjoyment (PNJ), perceived usefulness (PUF), and attitude toward course (ATT). Figure 1 shows the proposed relationships among these predictors. The reader may note that while prior works have presented some hypotheses similar to ours, but with the opposite direction, other studies have presented hypotheses with the same direction as ours. Much more importantly is that we present logical justifications (including support from the literature); and

our empirical evaluation strongly supports our research model. In the remainder of this section, we define each construct and provide the theoretical justifications for the hypotheses.

2.2.1 Attitude Toward Course (ATT)

According to (Triandis, 1971) attitude (*ATT*) is a representation of sentiments and understanding about a topic or situation, as well as an individual's positive or negative behaviour. *Attitude* can also be considered as the tendency of a person to respond negatively/positively towards something (i.e. idea, object, person, situation) (Vargas-Sánchez, Plaza-Mejía, & Porras-Bueno, 2016). Therefore, in this study *ATT* refers to a student's tendency to respond negatively/positively toward their effort in learning the ICT-based course materials/skills.

According to prior research, attitude plays an important role in influencing learning and successful learning. For instance, Wang (2010) finds that a negative attitude towards language learning results in language learning failure while a positive attitude yields positive results. Furthermore, Brown (2007) finds that a positive attitude towards language learning will shape the way an individual learn the language. Likewise for mathematics, research finds that a student's attitude towards the subject affects his/her learning process and eventual performance in the course (see Moussa & Saali, 2022). We extend this line of argument to suggest that a student's *ATT* toward the ICT-based course materials/skills will influence the student's learning process. Therefore, it is reasonable to expect that a student's *ATT* towards learning the ICT-based course materials/skills will likely influence his/her perception of how easy the course materials/skills are to learn, thus:

> **H1**: A student's *ATT* towards learning the ICT-based course materials/skills will influence his/her perception about how easy the course materials/ICT skills are to learn (*PEL*).

We argue that when an individual approaches something with a positive *ATT*, they are more likely to find it enjoyable. This is so because their expectations are set for a good experience, and they are likely more open to appreciating the positive aspects. The link between *ATT* and enjoyment is established. For instance, Spencer, Alison, and McKeough (2013) find that people who have more positive *ATT* towards exercise report enjoying it more and are more likely to stick with an exercise program over time. Similarly, Subramaniam and Silverman (2002) find that students with relatively positive *ATT* towards physical education have a higher degree of participation and express greater enjoyment. Russo and Minas (2020) also report that students with positive *ATT* towards a challenging task are more likely to find it enjoyable and be more persistent in the face of difficulty. Consistent with this line of argument, we propose that a student with a positive *ATT* towards learning the ICT-based course materials/skills, which may be difficult, may find it more engaging and ultimately enjoy learning those ICT materials/skills. Thus:

> **H2**: A student's *ATT* toward learning the ICT-based course materials/skills will influence his/her perception of how enjoyable (*PNJ*) the ICT-based materials/skills are to learn.

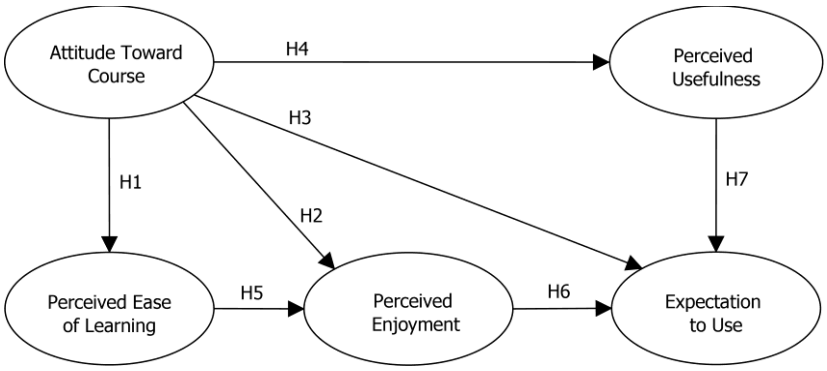


Figure 1: An ICT-adoption Model of Learning Experiences

We define expectation to use (*EXP*) as the extent to which the course materials/ICT-based skills are expected to be used in future; and it reflects the BI construct of TAM. According to TAM, an individual's behaviour is driven by BI, where intention is a function of an individual's ATT toward the behaviour. Many studies provide empirical evidence for the relationship between ATT and BI (e.g. Alismaiel et al., 2022; Barclay et al., 2018; Lin & Yu, 2023). We propose that an individual's *EXP* learnt ICT-based materials/skills in the future may be influenced by the student's *ATT*. This premise is supported since research finds that academic expectations could lead to students' arduous efforts for achievements (Levi, Einav, Ziv, Raskind, & Margalit, 2014). Hence:

> **H3**: A student's *ATT* toward learning the ICT-based course materials/skills will influence his/her *EXP* the ICT-based materials/skills in future.

We posit that *ATT* will influence *Perceived Usefulness (PUF)*. If an individual has a positive ATT towards a technology, s/he would be more likely to focus on its positive aspects and overlook any potential drawbacks. This may lead to an inflated perception of its usefulness. For example, if an individual is excited about a fitness tracker, s/he may be more likely to believe it will help him/her achieve their fitness goals, even if there is limited evidence to support this. Following on from this example, we suggest that *ATT* plays a crucial role in shaping a student's perception of the usefulness of ICT-based course materials/skills. We conjecture that a positive *ATT* can enhance *PUF* while a negative attitude may lead to a diminished perception of the ICT-based course skill's value. Thus:

> **H4**: A student's *ATT* toward learning the ICT-based course materials/skills will influence his/her *PUF* of the ICT-based materials/skills.

2.2.2 Perceived Ease of Learning (PEL)

We adapt the definition of *PEL* from Davis's (1989) conceptualization of perceived ease of use. Perceived ease of use is defined as the degree to which an individual believes that using a particular technology would be relatively free from effort. Since the focus of this study is not the actual ICT adoption, but rather, the learning experiences of ICT adoption, PEL is defined

as the extent to which it is perceived that learning the course materials/ICT-based skills does not require much effort.

We expect that *PEL* will influence perceived enjoyment (*PNJ*). Since most studies focus on actual ICT system adoption, and *PEL* is a new construct, we did not find any explicit literature to support the relationship between *PEL* and *PNJ*. However, we posit that it is reasonable to expect that if a student perceives the ICT-based course material/skills as being easy to learn then s/he will find the learning experience of the course to be more enjoyable than if s/he perceives the course materials/skills as being difficult to learn. Hence:

> **H5**: A student who perceives learning the ICT-based course materials/skills as easy to learn (*PEL*) will find the learning experience enjoyable (*PNJ*).

2.2.3 Perceived Enjoyment (PNJ)

We adapt the definition of *PNJ* by Davis, Bagozzi, and Warshaw (1992) to this study and define *PNJ* as the extent to which the course content is perceived to be enjoyable/interesting in its own right, apart from any performance consequences that may be anticipated. Previous research highlights the importance of *PNJ* in predicting users' intention to adopt technology. For instance, El-Said and Aziz (2022) find that *PNJ* of virtual tours (VTs) for touristic sites has a positive impact on intention to adopt VTs. Davis et al. (1992) report that *PNJ* is a major determinant and predictor of intentions to use computers in the workplace. We therefore suggest that when a student enjoys learning the ICT-based course materials/skills, s/he is more likely to actively participate, retain information, and see the practical value of the skills s/he is developing. This could then translate into an increased expectation to use these skills in future endeavors. Therefore:

> **H6**: A student's *PNJ* in learning the course materials/skills will influence their *EXP* the ICT-based course materials/skills in future.

2.2.4 Perceived Usefulness (PUF)

PUF is generally defined as the degree to which users believe that a particular system will enhance his/her performance (Davis, 1989). In the context of this study, *PUF* is defined as the extent to which it is perceived that taking the course is useful for the field of study and/or should enhance performance in another course(s). Empirically *PUF* is found to be a significant antecedent of technology adoption intention (e.g. Barclay et al., 2018; Mohammadi, 2015; Venkatesh & Bala, 2008; Zhang et al., 2023) and computer-based course evaluations (To & Tang, 2019). According to Rakoczy, Harks, Klieme, Blum, and Hochweber (2013), the *PUF* of feedback on student achievement is statistically significant.

Research further suggests that the degree of acceptance/rejection of a technology tool by subjects, depends on the level to which the user feels it will be of any benefit at present and/or in the future (Shim & Viswanathan, 2007), i.e. *PUF*. Analogous to this finding, we argue that the level of acceptance/rejection of learning the ICT-based course materials/skills will depend on the level to which the student feels the course materials/skills will be of benefit to him/her in their field of study and or in their future career. We therefore theorize that *PUF* should influence a student's *EXP* the ICT-based course materials/skills in future. Hence:

H7: A student's *PUF* in learning the course materials/skills will influence their *EXP* the ICT-based course materials/skills in future.

3. Research Methodology, Item Development, and Data

To assess our research model, we use a survey approach. We adapted prior validated items from the extant literature to maximize validity and reliability of our measurement instrument. Information systems (IS) faculty members as well as undergraduate and graduate students pre-tested our initial instrument. Using their feedback we revised several items. Measurement items for each construct in our research model are based on a 5-point Likert like scale. Each construct is measured reflectively, with multiple items. Our survey instrument also includes items to capture demographic data of the respondents as well as other background information such as academic major and year of study.

To test our research model, data is collected by administering an online questionnaire to business students enrolled for an undergraduate ICT-based course at a large university in the Global South. To encourage participation, the survey is included as a class activity during the final session of the semester. Approximately 160 registered students of the ICT-based course are asked to complete the survey in class (in a lab). Of the 160 students, 135 completed the survey, yielding a response rate of approximately 84%. Majority of the respondents are females (69.6%) and approximately 89% range in age between 20 – 24 years. The academic majors of the respondents vary to include Marketing, Banking and Finance, Accounting, Management Studies, and Human Resource Management, to name some.

3.1 Data Analyses and Results

Partial least squares (PLS) is used as our analytic technique. The PLS approach is suitable for our study as it appropriately assesses complex predictive models (Henseler, Ringle, & Sinkovics, 2009). PLS also avoids the problem of factor indeterminacy and inadmissible solution (Fornell & Bookstein, 1982). Smart PLS 3 (Ringle, Wende, & Becker, 2015) is the statistical tool used to analyse the measurement as well as the path models.

3.2 Measurement Model Analyses

To assess our outer model, we conduct tests of reliability and validity. As shown in Table 1 each construct demonstrates satisfactory internal consistency reliability as the composite reliability values range from 0.887 to 0.943, exceeding the recommended threshold value of 0.7. Another test of reliability is assessing the indicator loadings. As shown in Table 1, all loadings except one exceed 0.7, a common cut-off value, and are significant at the 0.01 level, based on their t-values. The indicator with the low loading is dropped from further analysis.

To test for convergent validity we assess the average variance extracted (AVE) values (Chin, 1998; Fornell & Larcker, 1981). As seen in Table 1, each AVE value exceeds 0.5, indicating that each latent construct can account for at least 50% of the variance in the items, on average. According to Hair, Hult, Ringle, and Sarstedt (2014) the Fornell-Larcker criterion (i.e. the square root of the AVE) and cross loadings are appropriate to assess discriminant validity. As shown in Table 2 discriminant validity is achieved as the square root of the AVE for each construct is larger on itself than on any number in the same row and column. Upon inspection

of the cross loadings, the loading of each item loads higher on its own construct than all its cross loadings on other construct, also indicating that discriminant validity is achieved (see Table 3).

Latent Construct	Item	Loading	t-value	AVE	Composite Reliability	Cronbach's Alpha
Perceived Usefulness	PUF1	0.758	12.986	0.686	0.938	0.923
	PUF2	0.881	16.232			
	PUF3	0.758	11.272			
	PUF4	0.819	14.073			
	PUF5	0.848	14.887			
	PUF6	0.856	14.563			
	PUF7	0.870	15.440			
Perceived Ease of Learning	PEL1	0.873	30.044	0.752	0.938	0.918
	PEL2	0.911	53.588			
	PEL3	0.882	20.987			
	PEL4	0.871	37.531			
	PEL5	0.848	34.889			
Attitude Towards Course	ATT1	0.809	14.651	0.724	0. 887	0.808
	ATT2	0.622	5.648			
	ATT3	0.896	33.960			
	ATT4	0.837	27.254			
Perceived Enjoyment	PNJ1	0.874	47.177	0.769	0.943	0.925
	PNJ2	0.914	54.890			
	PNJ3	0.885	41.164			
	PNJ4	0.870	33.220			
	PNJ5	0.841	25.689			
Expectation to Use	EXP1	0.908	49.479	0.805	0.925	0.878
	EXP2	0.854	24.835			
	EXP3	0.928	55.351			

Table 1: Measurement Model Indicators

	ATT	EXP	PEL	PNJ	PUF
Attitude Toward Course (ATT)	**0.850**				
Expectation to Use (EXP)	0.644	**0.897**			
Perceived Ease of Learning (PEL)	0.120	0.341	**0.867**		
Perceived Enjoyment (PNJ)	0.373	0.608	0.573	**0.877**	
Perceived Usefulness (PUF)	0.454	0.793	0.550	0.677	**0.827**

Table 2: Latent Variable Correlations.
Note: The square root of the AVEs are on the diagonal.

3.3 Structural Model Analyses

Figure 2 presents the estimates obtained from the PLS analyses. The R^2 value of 0.73 indicates that the model explains a substantial amount of variance for *EXP*. Additionally, the results show that approximately 0.21 of the variance in *PUF*, 0.42 of the variance in *PNJ* and 0.01 of the variance in *PEL* are explained by the variables in the model.

The results show that the paths from *ATT* → *PNJ*, *ATT* → *EXP*, and *ATT* → *PUF* are statistically significant, supporting hypotheses H2, H3 and H4. However, Hypothesis H1 is not supported since the path from *ATT* → *PEL* is not statistically significant. Figure 2 also provides

statistically significant evidence for the links between *PEL* and *PNJ* and *PUF* and *EXP*, supporting hypotheses H5 and H7. As for hypothesis H6, the link from *PNJ* → *EXP* is not statistically significant, providing no support for the hypothesis.

	ATT	EXP	PNJ	PEL	PUF
ATT1	**0.793516**	0.460846	0.351273	0.174187	0.335821
ATT3	**0.904132**	0.538995	0.306911	0.088638	0.395985
ATT4	**0.851335**	0.630102	0.297545	0.054264	0.419188
EXP1	0.597667	**0.908417**	0.519600	0.275895	0.676508
EXP2	0.586284	**0.853514**	0.509698	0.331569	0.707747
EXP3	0.550006	**0.928078**	0.603573	0.309128	0.746512
PNJ1	0.385569	0.608501	**0.873578**	0.538526	0.700370
PNJ2	0.231700	0.485113	**0.911118**	0.563180	0.573613
PNJ3	0.286200	0.501751	**0.882465**	0.507389	0.584290
PNJ4	0.316311	0.489095	**0.871136**	0.467302	0.509397
PNJ5	0.399151	0.562653	**0.844817**	0.431501	0.578046
PEL1	0.124246	0.332331	0.468801	**0.853364**	0.530146
PEL2	0.041604	0.249336	0.469320	**0.914580**	0.435882
PEL3	0.029547	0.223554	0.460090	**0.827838**	0.352380
PEL4	0.171328	0.368291	0.583810	**0.865405**	0.610813
PEL5	0.130397	0.280606	0.478755	**0.873392**	0.412517
PUF1	0.468225	0.627964	0.510367	0.426863	**0.758126**
PUF2	0.370038	0.687970	0.631608	0.504319	**0.881074**
PUF3	0.360999	0.636253	0.623908	0.543066	**0.757995**
PUF4	0.417229	0.691680	0.541723	0.503385	**0.818688**
PUF5	0.359785	0.658401	0.580284	0.410527	**0.847466**
PUF6	0.293713	0.627181	0.499606	0.375267	**0.855441**
PUF7	0.340154	0.650510	0.525147	0.407102	**0.870034**

Table 3: Item Loadings and Cross Loadings

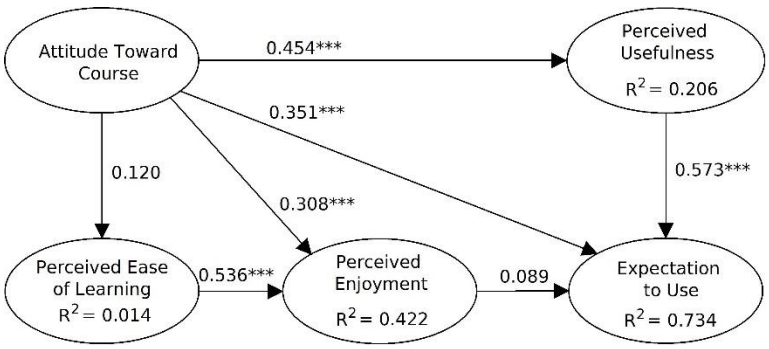


Figure 2: PLS Analysis of Results (*** $p < 0.001$)

4. Discussion & Conclusion

The reader may recall that this study's aim is to investigate what factors influence university students' learning experiences of ICT-based skills, with focus on the *Expectation to Use* (*EXP*) the relevant ICT-based course materials/skills in future. To the best of our knowledge, this is the first study that investigates the learning experiences of ICT adoption and not the actual adoption of the ICT itself. We consider two direct (i.e. *Attitude – ATT*; *Perceived Usefulnes*s - *PUF*) and two indirect antecedents (i.e. *Perceived Ease of Learning – PEL*; *Perceived Enjoyment - PNJ*) of *EXP*. We note that no previous model involved these four variables and only these four variables as possible antecedents of *EXP*.

The empirical evaluation of our model reveals that students' *EXP* is only influenced by *ATT* and *PUF*, and that together this pair of variables have a very strong impact on *EXP*, as evidenced by the magnitudes of the coefficients of the associated causal links (i.e. *ATT* to *EXP* = 0.351and *PUF* to *EXP* = 0.573) as well as the R-squared value of 0.734. These relationships are consistent with findings in the TAM literature which suggest that the BI to use/adopt a system is influenced by *ATT* and *PUF* (e.g. Alismaiel et al., 2022; Barclay et al., 2018; Lin & Yu, 2023; Zhang et al., 2023). Further, our results suggest that the assumed impact of enjoyment on learning may be exaggerated since there is no statistically significant relationship between *PNJ* and *EXP*. This result is consistent with prior research that found no statistically significant relationship between *PNJ* and mobile payment adoption intention (Koenig-Lewis, Marquet, Palmer, & Zhao, 2015). Also, while *PEL* has a statistically significant impact on *PNJ*, since *PNJ* does not impact *EXP*, then neither does *PEL*. This suggests that the focus should be on the incoming *ATT* of the students rather than the corresponding *PEL* or *PNJ*.

This study offers several theoretical contributions. While the issue of ICT adoption has been well investigated in the literature, what is less understood is the learning experiences of ICT adoption by individuals. This represents the first key contribution of our study to theory. To our knowledge, this is the first study that investigates the learning experiences of ICT adoption and not the actual adoption of the ICT itself. Second, our study offers new theoretical constructs, specifically, *PEL* and *EXP*, that can be used by others. These constructs are not only theorized but are also empirically validated. Third, this study offers a new and novel conceptual model of ICT adoption learning experiences for future investigation. Fourth, this study extends the TAM framework by integrating new constructs that may influence not only ICT adoption but also ICT adoption learning experiences.

The findings of this study offer practical guidance to ICT educators and practitioners in the Global South and in other developed jurisdictions. For instance, the finding that *PUF* influences *EXP* ICT-based skills in future indicates that educators as well as practitioners should highlight the importance of and demonstrate the value of learning ICT-based skills in today's technological world. Educators and practitioners should also use real world cases or scenarios to highlight the usefulness of ICT-based skills to organizations.

In future we could evaluate the model on non-ICT courses. Additionally, we could compare ICT values for *PEL*, *PNJ*, and other constructs to corresponding values for non-ICT courses. We could further develop and explore the direct relationship between *PEL* and *EXP*. Given data for *ATT* is collected at the end of the course, we could administer the survey close to mid-

semester to see if the results of this study are affected. i.e. to determine whether the learning experience biased students' attitude toward the course. The aim of this step is to see if the relationships that are statistically significant based on the survey administered in a final session also hold with the new data.

This study is one of the first, if not the first empirical study to investigate the experiences of leaners' ICT adoption. Additionally, it advances a long stream of research employing the TAM framework to evaluate individual BI. Further, this study integrates new/emerging factors, as well as others from the TAM framework, into a single comprehensive model which is empirically validated. The study also offers new relationships for addition to the nomological net of ICT adoption and for future ICT adoption research.

References

Al-Qaysi, N., Mohamad-Nordin, N., & Al-Emran, M. (2023). Developing a comprehensive theoretical model for adopting social media in higher education. *Interactive Learning Environments, 31*(7), 4324-4345.

Alismaiel, O. A., Cifuentes-Faura, J., & Al-Rahmi, W. M. (2022). Social Media Technologies Used for Education: An Empirical Study on TAM Model During the COVID-19 Pandemic. *Frontiers in Education, 7*, 1-12.

Bagozzi, R. P. (2007). The Legacy of the Technology Acceptance Model and a Proposal for a Paradigm Shift. *Journal of the Association for Information Systems, 8*(4), 244-254.

Barclay, C., Donalds, C., & Osei-Bryson, K.-M. (2018). Investigating critical success factors in online learning environments in higher education systems in the Caribbean. *Information Technology for Development, 24*(3), 582-611.

Brown, H. D. (2007). *Principles of language learning and teaching* (5th Ed.). White Plains, NY: Pearson Education, Inc.

Chin, W. W. (1998). "The Partial Least Squares Approach to Structural Equation Modeling". In G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 295-358). Mahwah, NJ: Lawrence Erlbaum Associates.

Cornell, R. M., Eining, M. M., & Hu, P. J. H. (2011). The effects of process accountability on individuals' use of a familiar technology. *Journal of Information Systems, 25*(1), 109–128.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 318-340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. *Journal of Applied Social Psychology, 22*(14), 1111-1132.

El-Said, O., & Aziz, H. (2022). Virtual Tours a Means to an End: An Analysis of Virtual Tours' Role in Tourism Recovery Post COVID-19. *Journal of Travel Research, 61*(3), 528–548. doi:10.1177/0047287521997567

Fathema, N., Shannon, D. M., & Ross, M. (2015). Expanding The Technology Acceptance Model (TAM) to Examine Faculty Use of Learning Management Systems (LMSs) In Higher Education Institutions. *MERLOT Journal of Online Learning and Teaching, 11*(2).

Fornell, C., & Bookstein, F. L. (1982). Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory. *Journal of Marketing Research, 19*(4), 440-452.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research, 18*(1), 39-50.

Graham, M. J., Frederick, J., Byars-Winston, A., Hunter, A. B., & Handelsman, J. (2013). Increasing persistence of college students in STEM. *Science, 341*(6153), 1455–1456.

Granić, A., & Marangunić, N. (2019). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology, 50*(5), 2572-2593.

Hair, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2014). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Thousand Oaks, CA: Sage.

Henseler, J., Ringle, C., & Sinkovics, R. R. (2009). "The Use of Partial Least Squares Path Modeling in International Marketing". In R. R. Sinkovics & P. N. Ghauri (Eds.), *Advances in International Marketing* (Vol. 20, pp. 277-319): Emerald Group Publishing Limited.

Koenig-Lewis, N., Marquet, M., Palmer, A., & Zhao, A. L. (2015). Enjoyment and social influence: predicting mobile payment adoption. *The Service Industries Journal, 35*(10), 537-554.

Levi, U., Einav, M., Ziv, O., Raskind, I., & Margalit, M. (2014). Academic expectations and actual achievements: The roles of hope and effort. *European Journal of Psychology of Education, 29*, 367–386.

Li, Y., Duan, Y., Fu, Z., & Alford, P. (2012). An empirical study on behavioural intention to reuse e-learning systems in rural China. *British Journal of Educational Technology, 43*(6), 933–948.

Lin, Y., & Yu, Z. (2023). Extending Technology Acceptance Model to higher-education students' use of digital academic reading tools on computers. *International Journal of Educational Technology in Higher Education, 20*(34), 1-24.

Mohammadi, H. (2015). Investigating users' perspectives on e-learning: An integration of TAM and IS success model. *Computers in Human Behavior, 45*, 359-374.

Moussa, N. M., & Saali, T. (2022). Factors Affecting Attitude Toward Learning Mathematics: A Case of Higher Education Institutions in the Gulf Region. *Sage Open, 12*(3), 1-13. doi:10.1177/21582440221123023

OECD. (2016). *Skills for a Digital World: Policy Brief on The Future of Work*. Retrieved from OECD Publishing, Paris: https://www.oecd.org/els/emp/Skills-for-a-Digital-World.pdf

OECD. (2022). *Education at a Glance 2022: OECD Indicators*. Retrieved from OECD Publishing, Paris: https://doi.org/10.1787/3197152b-en

Rakoczy, K., Harks, B., Klieme, E., Blum, W., & Hochweber, J. (2013). Written feedback in mathematics: Mediated by students' perception, moderated by goal orientation. *Learning and Instruction, 27*, 63-73.

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). SmartPLS 3: Boenningstedt: SmartPLS GmbH Retrieved from http://www.smartpls.com

Russo, J., & Minas, M. (2020). Student Attitudes Towards Learning Mathematics Through Challenging, Problem Solving Tasks: "It's so Hard–in a Good Way". *International Electronic Journal of Elementary Education, 13*(2), 215-225.

Shim, S. J., & Viswanathan, V. (2007). User Assessment of Personal Digital Assistants Used in Pharmaceutical Detailing: System Features, Usefulness and Ease of Use. *Journal of Computer Information Systems, 48*(1), 14-21.

Spencer, L. M., Alison, J. A., & McKeough, Z. J. (2013). A Survey of Opinions and Attitudes Toward Exercise Following a 12-month Maintenance Exercise Program for People with COPD. *Cardiopulmonary Physical Therapy Journal, 24*(3), 30-35.

Subramaniam, P. R., & Silverman, S. (2002). Using complimentary data: An investigation of student attitude in physical education. *Journal of Sport Pedagogy, 8*(1), 74–91.

To, W. M., & Tang, M. N. F. (2019). Computer-based course evaluation: an extended technology acceptance model. *Educational Studies, 45*(2), 131-144. doi:10.1080/03055698.2018.1443797

Triandis, H. C. (1971). *Attitude and Attitude Change*. New York, NY: Wiley.

Vargas-Sánchez, A., Plaza-Mejía, M. Á., & Porras-Bueno, N. (2016). "Attitude". In J. Jafari & H. Xiao (Eds.), *Encyclopedia of Tourism* (pp. 58-62). Cham: Springer International Publishing.

Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences, 39*(2), 273-315.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Management Science, 46*(2), 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly, 27*(3), 425-478.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly, 36*(1), 157-178.

Wang, C. (2010). Toward a second language socialization perspective: Issues in study abroad research. *Foreign Language Annals, 43*, 50-63.

Wu, X., Deshler, J., & Fuller, E. (2018). The effects of different versions of a gateway STEM course on student attitudes and beliefs. *International Journal of STEM Education, 5*(44), 1-12.

Zhang, C., Schießl, J., Plößl, L., Hofmann, F., & Gläser-Zikuda, M. (2023). Acceptance of artifcial intelligence among pre-service teachers: a multigroup analysis. *International Journal of Educational Technology in Higher Education, 20*(49), 1-22.

13. IT Executive Dream Teams

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

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

Abstract

The study delves into the strategic design and implementation of optimal IT executive team structures, emphasizing that the ideal configuration varies depending on each organization's specific context. It thoroughly examines the roles and definitions of key IT executive positions, such as CIO, CTO, and CDO, highlighting the challenges posed by ambiguous role definitions and the absence of standardization in IT management. A significant contribution of this paper is the development of a comprehensive framework. This framework seeks to strategically align the relationship between an organization's strategic formulation, planning, and deployment with the formation of its IT executive staff. The methodology of the study includes semi-structured interviews, offering a qualitative insight into the corporate dynamics that shape IT department compositions across different industries.

Keywords: IT Governance, Strategic IT Management, IT Executives, Performance Measurement.

1. Introduction

The fourth industrial revolution is in progress, changing economic systems, social structures, and organizational performance (Dalenogare, Benitez, & al., 2018). The success of today's organizations is largely based on the quality of their IT infrastructure, the IT intangibles, the ability of people to work with IT (Bayraktar, Tatoglu, Salih Aydiner, & Delen, 2023) and the value obtained by processing information (Cao, Duan, & Cadden, 2019). To better manage the different dimensions of IT use and exploitation, companies have defined new executive-level positions for the information technology (IT) management area that are super specialized and complement the classic roles of the CIO (Chief Information Officer) or of the CTO (Chief Technology Officer) (Bendig, Wagner, & Jung, 2022). Positions such as CDO (Chief Data Officer), CDO (Chief Digital Officer), CISO (Chief Information Security Officer), CinO (Chief Innovation Officer) or CTO (Chief Transformation Officer) are defined by the need to emphasize different technological resources and capabilities over the use and management of information in specific business environments.

Some attempts to explain and standardize the definitions of executive roles in IT have been published (La Paz, 2017; Gonzalez & Laurence, 2021), but despite the efforts, ambiguities persist about the commonalities and differences between them (Zimmermann, Schulz, & Gewald, 2020), and the design and configuration of the IT department and its executives is usually seen as a black box.

Achieving the optimal configuration of the IT executive team, which we call “the IT Dream Team” is something that does not happen spontaneously. On the contrary, it requires a process

of strategic planning, organizational design, and implementation. It is important to note that there is not just one IT team configuration that can succeed in all organizations or for all business strategies, but each organization will require a particular IT team configuration. This may seem like common sense, but alignment between organizational strategies and IT remains an elusive goal (Amarilli, van den Hooff, & van Vliet, 2023) and many factors such as policy, a poor top management team vision, lack of clarity in role definition, conflicting cultures, inappropriate IT architectures, and incorrect methods divert organizational structures from the ways in which they are required.

Whether it is an operational or strategic approach, it is necessary to have a framework that identifies and orders the main elements of the relationship between strategic deployment and the definition of the executive staff of the IT department, which allows organizations to have a practical and applicable perspective at their disposal to optimize the value they obtain from IT.

To better understand and subsequently propose a conceptual framework, the authors explored through semi-structured interviews, as the main qualitative research method (Patton, 2002), the corporate dynamics that define the design and composition of IT Departments in a sample of companies in different industries. Sampling does not presume randomness, is not stratified, and is made up of a series of questions previously established in a protocol.

2. Conceptualizing the process of IT unit design

Managerial skills are not equally distributed among all companies (Castanias & Helfat, 2001). The organization's leadership needs to evolve along with the organizational strategy, mutually adapting and iterating to avoid misalignment and achieve success in performance (Bytheway, 2011), but what is an IT Dream Team?

An IT Dream Team is the most-optimized combination of structures, leadership styles, professionals, roles, responsibilities, and skills that, in a coordinated manner, can efficiently and effectively deploy corporate strategies and produce the expected results through the investment and use of IT and its resources. IT executive teams differ in their purpose, composition, objectives, goals, and metrics by which their performance is evaluated. These teams have the responsibility of generating, communicating, and demonstrating competitive advantages in accordance with their own organizational strategy.

2.1 Building a framework for the design of IT dream teams.

The corporate strategy establishes the vision and direction of the organization and defines the quality of service expected from the IT department (Aboobucker, 2021). It considers the current state of the company and its projection of where it wants to be, including the level of dependence on IT assets (Amarilli, van den Hooff, & van Vliet, 2023). Xu and Quaddus in their book: “Managing Information Systems”, associate the basic competitive strategies first identified by Porter and reviewed by O'Brian and Marakas with the role that the IT function and information systems have in the use of these strategies. There are new business strategies such as “digital business” or “digital transformation” that cannot be classified as any of the previous categories exactly. These strategies are rather a combination of classic corporate strategies that increase the company's dependence on its technological assets and use them,

together with its technological capabilities, as fundamental elements to build competitive advantages (Hagiu & Wright, 2020).

Strategy deployment planning. In order to deploy the corporate strategy, the organization must plan and design its organizational structure, including the scope of its actions and responsibilities, the decision-making power of each functional area, the reporting structure and the corporate governance of all its operations and functional areas, including the IT department (Preston, 2008). The human resources literature adopted the term "strategic staffing" (a synonym for workforce planning) to represent the process that identifies and addresses the staffing needs of business strategies and plans and includes: recruiting, hiring, transferring, promoting, training, motivating, outsourcing, withdrawing, retaining and disengaging personnel (Oehlhorn, 2020). In so doing, a few scientific articles studied the evolution of IT executive's phenomenon, and differentiated types of executives, in an array of executive roles related to IT (Chun & Mooney, 2009, La Paz, 2017).

Performance measurement. The measurement and monitoring of performance through specific instruments, aims to generate a cybernetic cycle that provides feedback to the early phases of strategic planning, since the final intention of outcome measurement will be to provide reliable and valid information on performance (Behn, 2003).

Despite the academic recommendations, when looking at the performance indicators established in COBIT 2019 (the most widely used framework for the definition of a corporate IT governance and management system), we can see areas where the indicators are distributed among several executive positions (often the same indicator), as shown in Table 1. It is not possible to determine directly whether these overlaps are due to an ambiguity in the definition of roles, or to the consideration of implicit dependencies between positions not declared in the framework.

A top-down perspective emphasizes the strategic alignment, while a bottom-up approach allows individual's performance to be aligned and optimized with the overall objectives of the group. The literature suggests that measuring results is key to determining goals and objectives, planning activities to meet those goals, allocating resources in programs, monitoring and evaluating results to determine whether they are making progress toward goals, and verifying whether plans are being modified to ensure their performance (Akin & Karagozogiu, 2017).

The conceptual framework for the IT Dream Team is based on the references and relationships presented and on a previous model called the business transformation framework (Lambert & Peppard, 1993), which was one o the first to correlate the relationship of IT resources to the business strategy. Our frameworks is synthesized in Table 2.

The conceptual framework for the design of IT executive teams is presented in Figure 1, and describes how the definition of a particular business strategy that pursues results will require a formulation of organizational design, among the units of the structure, including the design of the IT unit, identifying roles, responsibilities, and that enable the strategic endowment, with people who meet the defined profiles and who will execute activities of some level of performance and achieve the results.

Family of Indicators	CDO	CIO	CISO	CTO	Total
Innovation Management		5		2	7
Stakeholder Management		42		35	77
People Management		17		12	29
Project Management		11		6	17
Service Management	1	7		8	16
Change Management		11	1	9	21
Document management	1	4		2	7
Financial management	5	35	1	15	56
Risk Management	1	31	1	8	41
Information Security	10	65	35	59	169
Asset Management		3		11	14
Stakeholder satisfaction	4	48	2	34	88
Communication Management	1	12		7	20
Education Management & Training	1	19	5	13	38
Compliance Management	2	22		11	35
Implementation of the strategy	5	48		31	84
Effectiveness, measurement, & evaluation	2	24	3	19	48
Operations Management	37	38	2	31	108
Grand Total	**70**	**452**	**50**	**319**	**891**

Table 1. Number of indicators associated to the IT executive's function from COBIT 2019

Stage	Component	Variables	Impact on deployment
Vision	Business Strategy	Generic Strategy: - Cost Leadership - Differentiation - Innovation - Growth - Strategic Alliance	Corporate strategies adopted by organizations vary in their IT dependence, and the way in which they use them to create competitive advantages
Planning	Formulation of Organizational Structure	Contingency Factors: - Size - Innovation - Diversification - Geographic Diversity - Strategy deployment plans	Incoherence between corporate strategy and the contingency factors leads to administrative inefficiency and detriment to value realization. The corporate strategy should reflect the intensity of IT use and the influence of the IT executive
	Design of Functional Areas	CIO	Strategist, Innovator, Enabler, Technologist
		CTO	Top Management Advisor, Strategist, Financial, Innovator, Processes, Other
		CISO	Incident Responder, Strategist, High Level Management
		CDO	Coordinator, Reporter, Architect, Ambassador, Analyst, Marketer, Developer, Experimenter

Deployment	Performance	Performance Indicators	Measuring performance is key to determining progress toward goals and whether the IT context that makes sense of the metrics used is considered.
	Results	Results Indicators	The indicators used to measure results should reflect the fulfilment of the strategic objectives

Table 2. Synthesis of elements necessary for the design of the IT executive level

The framework describes the logic for how companies should face the process of selecting the positions, roles, and responsibilities of the IT executives, how to fill such positions and finally how to assess their performance, according with the business objectives derived from and aligned with particular business strategies.

3. Methods and Results

To carry out an empirical validation to indicate how the dynamics between the components of the framework proposed manifest in real life cases, and to detect how to correct or improve the IT executive team's design process, this research conducts an exploratory, qualitative, and descriptive study of the current situation of IT executives. Data is collected through semi-structured interviews as the primary method of qualitative research. This method is used in studies evaluating the same target population and thematic analysis. The sampling does not assume randomness, is not stratified, and consists of a series of questions previously established in an interview protocol.

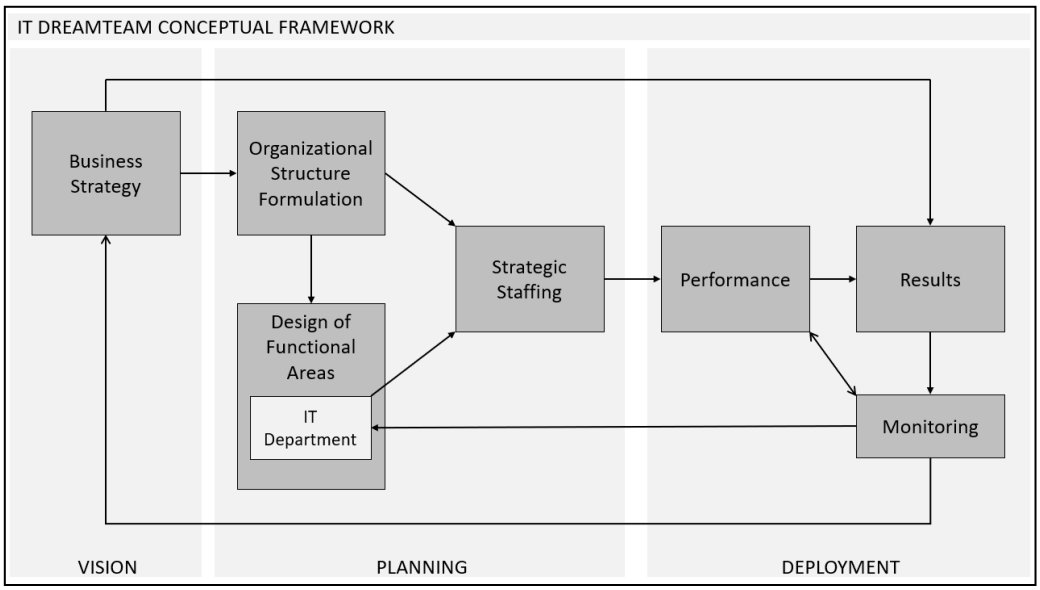


Figure 2. IT Dream Team Framework.
(Source: Author's Work)

The interview protocol was applied on IT executives from spanish-speaking organizations in Chile, aiming to identify and analyze the characteristics of each role and evidence of a planning process that supports their choice. Participation in the interview is voluntary, following an invitation letter. To ensure the validity of the study and its findings, the goal is to cover as many industries as possible. The total sample of IT executives participating in the interview consists of 10 high-level managerial profiles, achieving a satisfactory level of saturation.

Saturation is understood as the phase in which the researcher considers it not possible to extract new concepts or additional ideas from empirical information when a new observation is added to the sample. For heterogeneous groups, saturation is reached with a range of twelve to twenty informants, while in a homogeneous group, saturation is approximately twelve informants, with a base of ten for an effective qualitative study. The sample in this study corresponds to high-level individuals who are scarce resources and difficult to access due to the high demand of their responsibilities and the necessary privacy in managing information.

The interview transcripts were reviewed and analyzed in search of patterns following the thematic analysis process. The results analysis was conducted through open coding, an approach that relies on the transparency and accessibility of data. Open coding allows sharing research data in an organized manner, finding potential explanations amongst respondents' answers related to the same theme. In the first open coding process, this research generated 126 codes and five in vivo codes. A code reduction process was then carried out, eventually generating a total of 111 codes. The produced codes were grouped into categories in the axial coding process, generating a more abstract product from the codes and an articulation with the theoretical background.

- The sample was collected between December 2022 and January 2023.
- All interviews were conducted via videoconference, prior contact via e-mail or telephone to invite participants.
- Two interviewees did not give permission to record, so notes and written records are obtained.
- A total of 219 minutes of audio was collected between the eight recorded interviews, with an average of 28 minutes per interview.

3.1 Interview highlights

A first set of questions assesses the degree of knowledge about the strategic vision. The answer to this question is an indicator of how aligned the IT department can become with corporate governance and its degree of autonomy. Through these questions, we seek to understand whether corporate strategy is a guideline for IT decision-making. In the sample, most of the interviewees recognize the formal definition of a corporate strategy as an instrument for the alignment, but not necessarily participate in the strategic definitions. One interviewee declared to know the corporate strategy, but seemed himself as not responsible for it:

> *"I know it, but you'd rather ask to the person who really knows about it and in this case it's not me."*

A second set of questions in the interview protocol, focused on the knowledge, involvement, and influence of the IT executives in the organizational structure and design of the IT unit. IT executives typically possess significant autonomy and decision-making power, regardless of their department's position within an organization. In tech-dependent firms, those with high IT risks, or critical industries, there's increased formalization due to regulatory compliance and the necessity for risk and benefit management best practices.

A very important finding in this section is the relationship with other executives in IT related areas. Most organizations are prevalent in the "classic roles" of IT management (CIO – CTO), and a specialist role such as the CISO or a COO exclusively for IT operations. In addition, the relationship dynamics between the CIO, CTO and specialist executives appear to be functional and not hierarchical. From this it can be inferred that, in most organizations, there is a chief IT executive who holds the real C-level, supported by a set of managers or chiefs who fulfill a more tactical or operational role and not a plural executive body of specialists. As noted by one interviewee:

"We have a CTO and the general manager. IT departments support each other in generating and optimizing product development."

Whether by design or spontaneously, IT executives act like high-level executives (hence a high level of responsibility and autonomy) even when they are at lower hierarchical levels. One cause that can be ventured is the technical language required to interpret and understand IT's own processes, which requires IT executives to participate in high-level instances to generate common spaces that benefit decision-making (Amarilli, van den Hooff, & van Vliet, 2023).

The last section in the interview protocol dealt with questions around the formal measurement of results and determination of objectives. One participant declared:

"…we obviously work with indicators at the company level, also the different areas can have them, or not, it will depend just on the individual leadership…"

Confusing relationships are drawn between the levels of autonomy delivered and the way in which performance is measured. There is a great deal of freedom to develop and implement strategies, but measurement about how these efforts aim at meeting the strategic objectives of the organization seems disconnected. This is even more evident in cases where IT executives themselves define their own performance indicators. This scenario disrupts the feedback loop and the flow of information needed to determine whether and at what level the IT function contributes to the achievement of strategic objectives.

The IT executives interviewed tend to position themselves to the right of Figure 2, pointing to the "strategic focus" over the operational, in terms of attributes and self-perception of their roles; however, the mechanisms of performance and results measurement is positioned more to the "operational" side. Such types of contradictions in the strategic planning and organizational design reflects improvisation in the deployment of the corporate IT strategy, and an operational rather strategic approach towards technology. The Figure assists organizations

in the definition of IT executives' teams, where they can choose from, considering the conceptual framework introduced earlier.

3.2 Application of the conceptual framework

As revealed in the interviews and the literature, the arrangement of positions, structure, objectives, and performance metrics do not seem to reach yet a maturity or good practices leading to the conformation of IT dream teams. As the authors could not identify optimized teams that respond to a strategic vision, decided to introduce composite cases, as shown in table 3, to illustrate the application of the framework's logic, and contrast how different corporate contexts require a specific composition of IT executives.

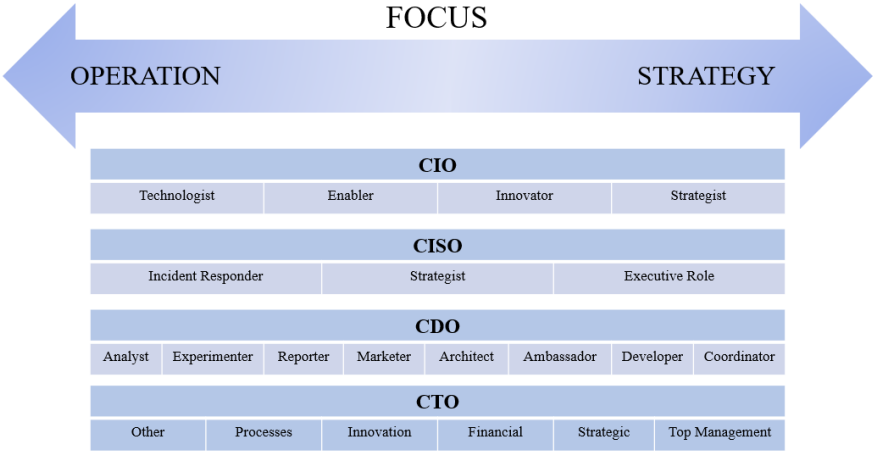


Figure 2. IT Leadership Orientations

The composite cases, highlight the differences in determining the optimal composition of the IT team and the main metrics to measure its performance, given the nature of a business strategy and industrial context.

4. Conclusions

Answering the question "How can companies design and implement an IT dream team maximize the value for the organization?" it is found that, for the organizations represented in the interviews, there is no well-defined leadership combination that can be called an "IT Dream Team". On the contrary, the entire IT function, including its various processes, often falls under one executive who relies on middle management. This leads to silos and a loss of potential collaborative benefits for organizational performance. Regardless of their title (CTO, CIO, CISO, CDO), IT executives have nearly identical responsibilities with minor variations in approach but not in substance. They all handle the overall budget, define and implement area-level policies and projects, and oversee IT processes. The implications of the discussion are as follows: First, academic research indicates that IT executive roles are becoming more specialized, with diverse technological functions (like production, operations, intelligence, security, strategy) at various organizational levels (strategic vs. operational). However, many organizations fail to fully utilize these professionals' capabilities. This underutilization could

be due to misunderstandings about technology reliance, cyber risks, or the importance of proper professional placement. As a result, IT tasks are frequently assigned to just one individual, limiting the potential benefits and risk management effectiveness.

	Case 1: Financial Institution	Case 2: Chain of Sports Centers
Business Strategy	**Differentiation**: Development of unique characteristics. **Strategic Alliance**: Inter-Operational Information **Risk aversion**: Moderate risk, focus on risk minimization and compliance management. **IT Goals** ▪ Reduction of Cybersecurity Incidents ▪ Uptime	**Differentiation**: Development of unique characteristics or reduction of differentiation from competitors. **IT Goals** ▪ Supplier Management ▪ Infrastructure Assurance
Strategic Endowment	The IT function is designed to meet operational needs. In this particular case, its departmentalization must manage risks, control and respond to incidents, execute IT projects and provide information for operational continuity.	Matrix, decentralized organization with few levels of authority. Its main characteristic is the existence of several business units, replicated for operational functions and that consume central services (machinery maintenance, IT, marketing)
Functional Area	**CISO:** The need to maintain various associated management systems, requires a Strategist type of CISO **CTO:** The main role of the CTO will be to ensure interoperability processes. The required CTO will be Strategist. **CIO:** Responsible for providing information and vision on new strategic development. The CIO should take the form of Innovator.	**CTO:** Ensures that ongoing maintenance plans are met, manage external IT service providers, and service level agreements are met. A Process CTO is required. **CDO**: A coordinator CDO focuses on obtaining value from the data that can be used by the marketing and commercial teams.
Performance	▪ Number of Incidents Detected ▪ Number of Unpatched Vulnerabilities ▪ Uptime of the Services ▪ Provisioning Time ▪ Percentage of Reports Delivered on Time	▪ Rate of incidents reported to help desk with timely response. ▪ % of Maintenance Programs carried out on time. ▪ Percentage change in the budget executed.
Result	▪ Achieving IT goals is moving toward strategic goals. ▪ Average Cost Per Incident ▪ % of Incidents Resolved at 1° Contact ▪ % SLA Compliance ▪ Cost of Downtime ▪ Impact of Information on Strategic Decisions	▪ IT Service Provider Turnover ▪ Breach of SLA ▪ % of Maintenance Programs carried out on time. ▪ Percentage change in the budget executed.
Monitoring	As a monitoring tool, the organization uses a Balanced Scorecard deployed at the Strategic, Tactical and Operational levels.	Indicators associated with the operation, based on short-term indicators and financial instruments like Cost Management and Operational Continuity

Table 3. Composite cases for the design of IT executive teams

Second, if organizations persist in choosing a single leader for IT, selecting the right individual, and designing their job profile is crucial, as they must fulfill various roles and risk becoming a "single point of failure.

Third, the interviews reveal unstructured executive role formulation. Therefore, organizations should adopt a structured approach to minimize deviations in formulating IT executive positions for strategic deployment. This method should align with strategic planning, considering IT dependence, each professional's unique traits, and the organization's value perception.

Future research will broaden this study by including more organizations, individuals, and industries to validate the conceptual framework. Cultural and developmental differences might affect the composition of the 'Dream Team' and emphasize various framework components differently. Additionally, incorporating IT objectives as a model input, rather than solely relying on outcome and performance metrics, is identified as a key enhancement in the framework's design.

References

Aboobucker, I. (2021). Information technology governance and IT capability to realize firm performance. *Benchmarking An International Journal*.

Akin, Z., & Karagozogiu, E. (2017). The Role of Goals and Feedback in Incentivizing Performance. *Managerial and Decision Economics* .

Amarilli, F., van den Hooff, B., & van Vliet, M. (2023). Business-IT alignment as a coevolution process: An empirical study. *The Journal of Strategic Information Systems*.

Bayraktar, E., Tatoglu, E., Salih Aydiner, A., & Delen, D. (2023). Business Analytics Adoption and Technological Intensity: An Efficiency. *Information Systems Frontiers*.

Behn, R. (2003). Why Measure Performance? Different Purposes Require Different Measures. *Public Administration Review.*

Bytheway, A. (2011). Assessing information management competencies in organisations. *Electric Journal of Information Systems Evaluation*.

Cao, G., Duan, Y., & Cadden, T. (2019). The link between information processing capability and competitive advantage. *International Journal of Information Management*, 121-131.

Chun, M., & Mooney, J. (2009). CIO roles and responsibilities: Twenty-five years of evolution and change. *Information & management*, *46*(6), 323-334.

Castanias, R., & Helfat, C. (2001). The managerial rents model: Theory and empirical analysis. *Journal of management*, 661-678.

Dalenogare, L. S., Benitez, G. B., & al., e. (2018). The expected contribution of Industry 4.0 technologies for industrial performance. *International Journal of Production Economics*.

Gonzalez, P., & Laurence, A. (2021). Not Your Typical Leader? CIOs Compare to Other Memebers of the Top Management Team. *Communications of the Association for Information Systems*.

Hagiu, A., & Wright, J. (2020). When data creates competitive advantage. . *Harvard Business Review*, 94-101.

La Paz, A. (2017). How to Become a Strategist CIO. *IT Professional*, 45-55.

Lambert, R., & Peppard, J. (1993). Information technology and new organizational forms: destination but no road map? *The Journal of Strategic Information Systems, 3*, 180-206.

Oehlhorn, C. M. (2020). Human resource management and its impact on strategic business-IT alignment. *The Journal of Strategic Information*.

Patton, M. (2002). *Qualitative research & evaluation methods.* California: Sage Publications.

Preston, D. C. (2008). Examining the antecedents and consequences of CIO strategic decision-making authority. *Decision sciences*.

Zimmermann, S., Schulz, T., & Gewald, H. (2020). Conflicting Roles of CIOs and their Negative Effects on the Workplace of the Future. *Hawaii International Conference on System Sciences.* Hawaii.

14. Sprinting toward the future: transition to industry 4.0 in the context of small and medium-sized firms
(Research-in-Progress)

Redouane El Amrani
Audencia Business School
relamrani@audencia.com

Abstract
The study explores digitalization in industrial small and medium-sized enterprises (SMEs) and the transition to Industry 4.0 through a dynamic capabilities framework. The research question aims to identify the unique capabilities of companies to cope with Industry 4.0 in the new digitalization environment. The study takes a qualitative approach and draws on nine interviews with key respondents such as the owner, executives and Industry 4.0 consultants who helped the company manage the transition to Industry 4.0. The study reveals that distinctive dynamic capabilities such as technological, organizational and human capabilities are necessary for the digitalization process and the transition to Industry 4.0. The study aims to contribute theoretically and empirically to new meanings of dynamic capabilities for digital transformation in the context of industrial SMEs.

Keywords: Digitalization, Industry 4.0, SMEs, dynamic capabilities

1. Introduction
Digital transformation through the implementation of the Industry 4.0 paradigm is becoming a strategic priority for SMEs (Soluk & Kammerlander, 2021; Volberda et al., 2021). They are experiencing a massive shift in the dynamics of their value chain and a major advancement in the way their products are designed and marketed (Marino-Romero et al., 2024). Yet SMEs have diverse perspectives on what Industry 4.0 means and are eager to adopt new digital technologies and managerial approaches (Hanelt et al., 2021). For most SMEs, it is difficult to leverage existing capabilities while exploring and developing new digital capabilities. In fact, compared to large companies, they are often characterized by poorly formalized processes, fragmented existing hardware and software systems, and fewer economic resources. Additionally, they typically lack in-house IT skills and associated technological knowledge. Thus, without transformation actions, traditional SMEs would lose market share to new companies entering traditional sectors with business models based on the use of new digital technologies.

While digital disruption is hitting their business model and generating a major change in their core businesses, how can we understand and explain the fundamental changes taking place with the transition to digital transformation in the context of industrial SMEs? The scope of impact remains unclear and we do not know if we are in the early phase of the transformation or if we have already passed the turning point where there is no going back? These questions concern

many researchers and business leaders and are currently being discussed in different disciplines.

In this study, we aim to explore digitalization process of an industrial SMEs and the specific transition to Industry 4.0 through dynamic capabilities framework (Ghosh et al., 2022). We argue that dynamic capabilities are likely to provide a better understanding of how SMEs make this transition to Industry 4.0, in particular, how they perceive new digital opportunities to modify existing capabilities, resource configurations, business model components and get the entire organization in a journey of transformation. However, knowledge on this topic is scattered across various research streams, so insights into the process are both limited and fragmented (Escoz Barragan et al., 2024). According to Moeuf et al. (2018), little is known whether SMEs can successfully transition to Industry 4.0 and what types of dynamic capabilities are specific to their contextual size. Knowledge of this process is important because traditional organizations may fail to achieve digital transformation due to their lack of understanding of Industry 4.0 principles. From an information systems perspective, the development of useful dynamic capabilities in the context of Industry 4.0 has received limited attention, although the dynamic capabilities framework has been one of the most important topics in the field of strategic management (Warner & Wager, 2019). This literature gap is particularly present in the context of SMEs (Marino-Romero et al., 2024). To the best of our knowledge, understanding how SMEs operating in traditional industries engage in the transition to Industry 4.0 has not received sufficient attention from IS scholars and requires further investigation (Chauhan et al., 2022). Although this topic has attracted the attention of researchers beyond the IS discipline, the size of organizations has not been sufficiently studied and most previous research has considered large companies rather than SMEs.

Our goal in this paper is to propose a conceptual model for the new dynamic capabilities that SMEs must build and develop to respond to the serious threats and challenges posed by digital transformation. It is difficult for most SMEs to develop appropriate capabilities, which would consist of a mix of digital capabilities and sector-specific capabilities. SMEs often struggle to put Industry 4.0 paradigms into practice and many are unable to develop such new capabilities on their own. To address this topic and answer this question, this study is exploratory research, conducted based on inductive reasoning, which aims to provide new insights on the subject and to propose prospective reflection for future research. This paper is organized as follows. First, we review the relevant literature underlying our theoretical framework. We then present our research methodology in terms of data collection and analysis, before presenting and discussing the first results of our study.

2. Conceptual Background

2.1. Digital transformation of SMEs: the importance of Industry 4.0

Digital transformation is a process that aims to improve organizations by triggering significant changes in their properties through combinations of information, computing, communication and connectivity technologies (Vial 2019). Due to its potentially immense organizational impact, digital transformation has been one of the top management priorities over the past few years. Coping with these changing business dynamics requires organizations to reinvent their

business models and the way they interact with their customers, employees and other stakeholders (Kraus et al., 2022).

In the context of SMEs, this is not the first time that SMEs have had to face a transformation project. Over the last decade, many manufacturing SMEs have launched transformation projects such as Enterprise Systems (ERP, CRM and MES) considered as important competitive tools for them. With Industry 4.0, SMEs are taking a leap into another technological register where they must equip themselves with a portfolio of new digital technologies that must be integrated in real time. Industry 4.0 is understood in this paper as a revolutionary industrial concept of the production process in the manufacturing sector, focused on new digital technologies that interconnect machines and equipment with digital data in autonomous and intelligent systems. It is associated with many opportunities and advantages, such as highly flexible mass production, reduced complexity costs, the emergence of entirely new services and business models or real-time coordination and optimization of value chains (Moeuf et al., 2018).

This revolution is a game changer for the survival of many businesses and industries. Firstly, because it meets the requirements of the ongoing digital transformation (Vial, 2019). Second, it has many potentials and should boost industrial competitiveness and put labor productivity back on the path to growth. Third, the increased flexibility brought by Industry 4.0 technologies may incentivize companies to locate production close to their customers and regain some of the responsiveness lost in fine-sliced global supply chains (Veile et al., 2020). SMEs have generally recognized that if they do not take into account the possibilities offered by Industry 4.0 by carrying out a wide range of complex business transformations, more technologically advanced competitors will soon drive them out of the market. Specific concerns for SMEs during Industry 4.0 transition are related to their intrinsic characteristics such as small size, lack of organizational specialization and resources and generally unsophisticated internal and information technology infrastructure. Many studies (Soluk & Kammerlander, 2021; Li et al., 2018) point out some deficiencies that SMEs face during Industry 4.0 transition such as lack of competency and know-how, insufficient qualified human resources, lack of appropriate support technologies or organizational structures and problems in gaining sufficient participation from different organizational functions during digitalization implementation.

However, the question remains whether they are actually capable of implementing such a modern paradigm and achieving digital transformation. Meet the future needs of their customers; they must transform their business model by developing a wide range of new capabilities and specific skills. These capabilities should cover transformational capabilities that include innovative thinking, formulating a digital leadership vision and strategy, developing digital talent and managing change. At the same time, they must assimilate new digital technologies and know how to integrate them into their industry's value chain. However, the literature has shown that SMEs face increased difficulties in adopting new technologies due to high capital requirements, lack of resources, skills, commitment and a good understanding of digital opportunities.

Faller and Feldmüller (2015) found that SMEs particularly have great difficulty acquiring skills in Industry 4.0 solutions and technologies. These technologies have been recognized as one of

the main key elements of Industry 4.0 and neglecting them is risky and can seriously compromise the survival of the company. A large of literature refers to digital technologies as an important contributor to Industry 4.0 and can be considered the most important dimension of the most popular Industry 4.0 readiness models (Hizam Hanafiah et al., 2020). This implies that they support value creation, organizational agility and customer engagement, which are essential success factors for SMEs. As a result, they need to significantly improve their technology readiness to strengthen their readiness for Industry 4.0.

The transformation also comes with various challenges that SMEs must face. While this transition to Industry 4.0 leads to new value propositions for markets, it requires new value chain activities, new partnerships and new revenue models. The objective of Industry 4.0 is to manage the entire value chain process by improving the efficiency of the production process. The scale of the impact is therefore expected to be transformational and disruptive for almost all SMEs. The analytical thinking behind the Industry 4.0 approach is not limited to rapid digitalization of production processes and systems. Industry 4.0 involves the need to leverage existing capabilities while exploring new transformational capabilities. As a result, the impacts on the business model, the work system, knowledge management and people's skills and expertise are very significant. It is difficult for most companies to develop appropriate capabilities, which would consist of a mix of digital capabilities and industry-specific capabilities. SMEs often struggle to put Industry 4.0 paradigms into practice and many are not capable of developing such new capabilities on their own or it will take them years to build and implement proprietary internal solutions.

Dynamic capabilities in the digital area

Dynamic capabilities for Industry 4.0 can be an enabler of continuous adaptation through sensing changes, seizing opportunities, and transforming resources. This perspective contrasts with the usual universalistic approach (Qaiyum and Wang, 2018). Drawing on resource-based views, the dynamic capabilities perspective describes how a firm’s competencies can be transformed to fit new environmental circumstances (Teece et al., 1997; Eisenhardt and Martin, 2000). They are described as higher-order organizational capabilities that support companies in changing their strategy, adapting their organizational structures, processes, and company cultures (Karimi and Walter, 2015). As a part of organizational capabilities theory, dynamic capabilities are distinguished from a firm’s ordinary capabilities (Qaiyum and Wang, 2018). Whereas ordinary capabilities involve the exploitation of current resources to achieve the desired result, the literature describes dynamic capabilities as an organization’s ability to change and modify the current resource-base by means of exploration. Dynamic capabilities are a “meta-competence” that transcends operational competence (Teece, 2007) and represent an organization’s ability to transform its resources portfolio to have a positive influence on performance. In this sense, dynamic capabilities involve long-term commitment to specialized resources (Helfat et al., 2007). This continuous adaptation of internal resources and competencies will determine the long-term profitability of enterprises.

The actual digital disruption is creating a significant “capability gap” for incumbent SMEs (Volberda *et al.,* 2021). It introduces new digital knowledge and alternatives, stimulating new ways of performing and organizing the value chain. This capability gap represents the distance between a firm’s actual configuration of capabilities and the most valuable configuration

potentially available to offer tangible responses to this digital disruption (Warner and Wäger, 2019). Such gaps require SMEs to develop a portfolio of capabilities to face new types of competition, innovate new products, and integrate the new digital technologies as an asset in their new business models. Therefore, dynamic capabilities are directed toward strategic change and are essential in responding to disruptive technological shifts (Helfat *et al.,* 2007).

Prior studies emphasize dynamic capabilities as the key driver of digital transformation (Yeow *et al.,* 2017). They argue that dynamic capabilities are innovation-based and provide an organization with the capacity to purposefully create, extend, or modify its resource base. By helping firms to deal with environmental change, dynamic capabilities transcend the barriers created by the transition to Industry 4.0 and circular economy as it enables firms to overcome core rigidities and misalignments (Teece, 2007) and effectively allocate necessary resources. They are also indispensable to smoothly integrate the new digital technologies with the existing business operation, and connect with new customers' expectations (Magistretti *et al.,* 2021). Recent research on dynamic capabilities has suggested different frameworks where companies sense and seize valuable growth opportunities and then transform their capabilities in the pursuit of these opportunities (Dong *et al.,* 2016). Employing dynamic capabilities as an analytical lens, Canhoto *et al.,* (2021) identify the specific combinations of sensing, seizing, and reorganizing capabilities associated with each phase of digitalization.

Warner and Wäger (2019) examine in their qualitative study how incumbent firms in traditional industries build dynamic capabilities for digital transformation. Their model conceptualizes digital transformation as a process of building dynamic capabilities for the ongoing strategic renewal of an organization's business model, its collaborative approach, and eventually its culture. Moreover, because digital disruption is an ongoing process, the organization should recalibrate the need for sensing and seizing new opportunities, which involves ongoing refreshments and replacements of business models, collaborative approaches, and cultures. In the same register, but this time in relation to a specific technology,

Vrchota et al., (2019) showed that the readiness and the capability to meet this challenge exist in parts within SMEs, they are more agile, flexible and innovative in comparison with large and well-established firms (Dibrell et al., 2008). The same research identified typical factors that influence the implementation of Industry 4.0 in SMEs by emphasizing specific capabilities such as social capital development, business team building and organizational capability building. If some SMEs, especially in highly regulated sectors as healthcare, electronics and automotive, are forced to move to meet the Industry 4.0 principles, others are developing an awareness concerning the relevance of the topic among SMEs. Finally, Karimi and Walter (2015) structure a hierarchical framework of dynamic capabilities into resources, processes and values.

3. Methodology

Given the exploratory nature of the research and the relative novelty of the phenomena studied, we specifically chose a case study design. This choice adapts to the study of contemporary phenomena in a real context and allows us to reinforce our results and lead to conclusions that are more robust. To address the call for higher methodological rigor in IS case research (Dubé

and Paré, 2003) and concerns about the validity of this method, we used the procedures and quality criteria proposed by Yin (2018).

3.1 Data collection

We used purposive sampling (Yin, 2018) to identify the context and SMEs providing rich empirical data on the phenomena studied. We selected our case based on two main criteria. First, the company concerned had to be considered as one of the SMEs that worked on digital transformation and Industry 4.0 in order to provide empirical evidence to our research questions. Second, the company must have been actively engaged in digitalization projects in recent years and must have a deep understanding of the interaction between the concepts of Industry 4.0 and digitalization. Following these guidelines, we decided to conduct a case study on a French SME that has greater practice in implementing Industry 4.0 for more than five years in the manufacturing industry.

For the data collection process, we adopted a combination of primary and secondary sources. A total of nine interviews constituted the main source of data. We conducted semi-structured interviews with five members of SMEs adopting the digital transformation and four interviews with consultants with concrete expertise in managing Industry 4.0. These consultants were involved in the company's Industry 4.0 project and had complementary points of view enriching this study. The interviews were conducted face-to-face and online interviews and took place between June and December 2023. We asked the interviewees to share their experience and personal views on how digitalization has helped achieve the transition to Industry 4.0. During the interviews, we collected comments, reactions, suggestions and additional information on events and changes observed before and during the projects. This helped to understand the challenges of Industry 4.0 in an SME and to review the logic and the coherence of digitalization strategies in the development of new capabilities specifically for Industry 4.0. Additionally, we collected secondary data from the company website, official press releases, and media coverage. These sources were used and analyzed throughout the research process for the triangulation of our results (Yin 2018).

Expert	**Current Designation**	**Date**	**Duration** (minutes)
1	CEO	12- June-23	45
2	Production Manager	17-June-23	72
3	Project Director	23-Aug-23	54
4	Sales Manager	07-Sept-23	45
5	Quality Manager	12-Sept-23	54
6	Consultant	30-Oct-23	72
7	Consultant Industry 4.0	30-Nov-23	45
8	Consultant Industry 4.0	17-Dec-23	63
9	Consultant Digital Transformation	17-Dec-23	81

Table I: *Details of interviews & experts*

3.2 Data Analysis

To analyze our data, we relied on NVivo software 11. We organized it into sub-topics. For each sub-topic, we analyzed information from multiple data sources in order to identify common themes across the data sources. The data analysis concentrated on reviewing interview transcripts to uncover both the hidden stories and apparent storylines surrounding different projects. We identified quotes that are particularly representative in capturing the essence of the insights highlighted in discussing each topic. The identification of storytelling patterns relied primarily on the interviews as a means of eliciting information about the studied phenomenon. We reviewed and labelled the interview data via detailed reading and rereading of the transcripts.

To analyze our data, we relied on the NVivo software 11. We organized it into sub-topics. For each sub-topic, we analyzed information from multiple data sources to identify common themes across data sources. Data analysis focused on examining interview transcripts to uncover both the hidden stories and apparent storylines surrounding the different projects. We identified particularly representative quotes because they capture the essence of the ideas highlighted during the discussion of each topic. We examined and labeled the interview data through detailed reading and re-reading of the transcripts. To maximize research credibility, we conducted the methodological triangulation of the data by complementing interviews with the field notes and an analysis of the secondary data.

4. Case Study and Initial Results

Located in the north of France, Klarex is a pump manufacturer specializing in the engineering and manufacturing of corrosion resistant centrifugal pumps. Manufactured from engineered thermoplastics, these pumps are designed for complex industries such as nuclear power, chemical and petroleum to handle corrosive and abrasive fluids.

The company's business model is based on the engineering, manufacturing and sale of pumps and associated services in more than 25 countries. To maintain and strengthen its international presence, the company relies on an international network of distributors and direct agents. This sales and after-sales process involves a significant volume of data and this area was one of the reasons why we started the process of digitizing the company. Moreover, the market analysis clearly shows new general trends in the international market, such as better designed machines at lower prices with greater integration of digital technologies and environmentally friendly machines. This new context pushed the CEO to engage a digital transformation process by developing an Industry 4.0 program. The goal was to transform the small manufacturer of anti-corrosion centrifugal pumps and build a new business model based on digital and servitization. Klarex has focused its transition to Industry 4.0 around three main axes:

1. Increase value creation by paying a particular attention to the customer experience and differentiating its offer via the integration of new digital technologies
2. Develop a new strategy called “Product as a Service: PaaS”
3. Involve and Integrate all stakeholders in this transition to Industry 4.0

Klarex has redefined its strategy by creating new business units called “Engineering and New Projects” and “Services” in order to structure the company's offering to its customers and

manufacture more products with better performance. Today, Klarex builds new machines from a standard catalog and custom machines specifically sized to fit customers. In this new configuration, the company offers new pre-sales engineering services where digital simulation brings a new advantage to the company's customers. During this phase, senior management was exploring the opportunity to expand its expertise from a manufacturer making its product to a new industrial company capable of providing engineering services and selling studies on complex technical solutions. By connecting the machines and collecting data, the company will be able to personalize the machine and therefore to provide different services related to different usage of the product: real time analysis of vibrations, energy consumption and operational performances of machines, predictive maintenance actions specific to each technical environment, etc. Machine optimization is one of the results of intelligent predictive maintenance. Compared to more conventional data acquisition, the adaptability of the algorithms developed by Klarex allows them to process different types of data (structured and unstructured) for constantly renewed functions according to customer needs. The pumps are capable of providing physical data, but they can also self-regulate. This new approach allows the company to change the value creation process and position itself as a future major player in the French Industry 4.0 ecosystem.

To achieve this transformation, the company launched a reskilling and upskilling training program targeting customers and the distributor network. This program offers theoretical training on pump structures and practical training related to the repair and maintenance of equipment. Some internal engineers first carried out this activity before developing a partnership with an external training company which set up a specific training program around the company's profession.

The first round of data analysis allowed us to identify the different capabilities necessary for the adoption of digital transformation and the transition to Industry 4.0 for an SME in an industrial context. These capabilities are classified into three main categories: technological, organizational and human.

Technological capabilities: technological capabilities that Klarex is developing are a combination of common emerging digital technologies such as cloud computing, IoT, analytics and AI algorithms. According to different interviewed experts, it is clear that most of the technologies of the Industry 4.0 are scattered in different aspects. For example, sensor devices are embedded with smart algorithms to sense and report data with accuracy and powerful analytics and AI algorithms are correlated with simulation, automation and optimization. These prove that the implementation of analytical models plays an important role in the smart factory. These technologies are both disruptive and offering great opportunities for the company to enhance its competitiveness. They could be considered as a cornerstone of any dynamic capability framework for digital transformation.

Organizational capabilities: in the dynamic capability paradigm, organizational capabilities have been identified as one major source for the generation and development of sustainable competitive advantages. In the context of Klarex, the CEO and the team in charge of Industry 4.0 initiated an organizational learning process in which a specific way of ‘selecting and linking’ resources gradually develops. For example, to change their familiar ways of doing, the

team installed a new digital service and initiated different training workshops to overcome resistance and develop new data and digital culture. For the project director, one of the greatest challenges was finding a way for re-imagining the employees' experience and bringing their data digital literacy up to date. Moreover, as Klarex started partnering with start-ups for access to innovative solutions and creating new digital ecosystems and industry alliances

Human capabilities: in the context of SMEs, exploring digital transformation and Industry 4.0 is a consecutive process of trial, error and decision-making. This study confirms that capabilities reside in people, particularly those responsible for orchestrating, combining and organizing resources to create value. Certain capabilities (detecting, searching and selecting the right source of digital knowledge) reside either only within the entrepreneur or are shared among his or her closest collaborators; the others (seize and orchestrate internal and external assets, physical and digital resources) reside in the executive team; others are distributed granularly among many people, as is the case with manufacturing processes.

5. Conclusion

Digital transformation and Industry 4.0 are pushing industrial SMEs around the world to adopt digital technologies and new managerial approaches to develop their competitiveness. SMEs are most likely to be the big winners from the shift; they are often able to implement the digital transformation more rapidly than large enterprises, because they can develop and implement new IT structures from scratch more easily.

However, if research has widely acknowledged that the adoption of digital technologies is positively linked to development of Industry 4.0 capabilities, we have been lacking insights from SMEs (Chauhan et al., 2022). This study addresses this issue from a dynamic capabilities perspective and identify three main categories (technology, organization and human) that are specific to an industrial SME.

By investigating these three capabilities required for the successful adoption of digital transformation and the implementation of Industry 4.0 in an industrial SME, this study shows that it is not easy for an SME to change the business model by integrating new digital technologies and managerial approaches. Transition to Industry 4.0 is about sensing and responding to the digital disruption changing industrial environment of SMEs. At the heart of this continuous transformation, SMEs has to update their resources portfolios and develop new capabilities to deal with the challenges imposed by the digital disruption ongoing in their industry.

References

Canhoto, A, Quinton, S, Pera, R, Molinillo, S, Simkin, L. (2021), Digital strategy aligning in SMEs: A dynamic capabilities perspective, *Journal of Strategic Information Systems*, 30 (3), 101682.

Chauhan, C.; Parida, V.; Dhir, A. Linking circular economy and digitalization technologies: A systematic literature review of past achievements and future promises. *Technological Forecasting and Social Change*, *177*, 121508.

Dong, A, Garbuio, M, Lovallo, D, (2016), "Generative sensing: a design perspective on the microfoundations of sensing capabilities**",** *California Management Review*, 58/4: 97-117.

Dubé, L. Paré, G. (2003), Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and Recommendations, *MIS Quarterly*, 27(4), 597–636.

Eisenhardt K.M, Martin J.A (2000), Dynamic Capabilities: what are they? *Strategic Management Journal*, 21 (10-11), 1105-1121.

Escoz Barragan, K., Hassan, S.S., Meisner, K. and Bzhalava, L. (2024), "Dynamics of digital change – measuring the digital transformation and its impacts on the innovation activities of SMEs", *European Journal of Innovation Management*. https://doi.org/10.1108/EJIM-05-2023-0432

Faller, C. Feldmüller D, (2015), « *Industry 4.0 learning factory for regional SMEs*", Procedia CIRP, 32, pp. 88-91

Ghosh, M. Hughes, J. Hodgkinson, P. Hughes (2022), "Digital transformation of industrial businesses: a dynamic capability approach" *Technov*., 102414.

Hanelt, A. Bohnsack, B. Marz, D. Antunes, C (2021), "A systematic review of the literature on digital transformation", *Journal of Management Studies*, 1-36

Helfat CE, Finkelstein S, Mitchell W, Peteraf M, Singh H, Teece D, Winter S. (2007), *Dynamic Capabilities: Understanding Strategic Change in Organizations*. Blackwell Publishing: Malden, MA.

Hizam-Hanafiah, M, Soomro, M.A., Abdullah N.L. (2020), "Industry 4.0 readiness models: A systematic literature review of model dimensions", *Information*, 11 (7).

Karimi, J. Walter, Z, (2015), "The role of dynamic capabilities in responding to digital disruption: a factor-based study of the newspaper industry", *Journal of Management Information Systems*, 32 (1), 39-81.

Kraus, S. Durst, J. Ferreira, P. Veiga, N. Kailer, A. Weinmann (2022), "Digital transformation in business and management research: An overview of the current status quo", *International Journal of Information Management*, 63, 102466.

Li, L. Su, F. Zhang, W. MaoJY, (2018), "Digital transformation by SME entrepreneurs: A capabiltiy Perspective », *Information Systems Journal*, 28(6), pp. 1129-1157.

Magistretti, S. Pham, C. Dell'Era, C (2021), Enlightening the dynamic capabilities of design thinking in fostering digital transformation, *Industrial Marketing Management*, 97, 59-70.

Marino-Romero, J.A., Palos-S´ anchez, P.R., Velicia-Martín, F., (2024), "Evolution of digital transformation in SMEs". *Technological Forecasting and Social Change*, 199.

Moeuf, A.; Lamouri, S.; Pellerin, R.; Eburdy, R.; Tamayo, S. (2018), *Industry 4.0 and the SME: A technology-focused review of the empirical literature*, Process Safety and Environmental Protection, 12, 408-425.

Qaiyum, S. Wang, C.L. (2018), Understanding internal conditions driving ordinary and dynamic capabilities in Indian high-tech firms, *Journal of Business Research*, 90, 206-214.

Soluk, J & Kammerlander, N, (2021), "Digital transformation in family-owned Mittelstand firms: A dynamic capabilities perspective," *European Journal of Information Systems*, 30(6), pp. 676-711.

Teece, D. J. (2007), Explicating dynamic capabilities: The nature and microfoundations of enterprise performance, *Strategic Management Journal*, 28: 1319-1350.

Teece, D.J.; Pisano, G; Shuen, A (1997), Dynamic Capabilities and Strategic Management, *Strategic Management Journal*, 18, pp. 509-533.

Veile, J.W.; Kiel, D.; Müller, J.M.; (2020), "Lessons Learned from Industry 4.0 Implementation in the German Industry", *J. Manu. Tech. Manag.*, *31*, 977–997.

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

Volberda, H. W., Khanagha, S., Baden-Fuller, C., Mihalache, O. R., & Birkinshaw, J. (2021), "Strategizing in a digital world: Overcoming cognitive barriers, reconfiguring routines and introducing new organizational forms". *Long Range Planning*, 54(5).

Vrchota, J., Maˇrikova, ´ M., Reho ˇ ˇr, P., Rolínek, L., (2020), "Human resources readiness for industry 4.0", *Journal of Open Innovation: Technology, Market, and Complexity*, 6(1)

Warner, K.S., Wäger, M. (2019), "Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal", *Long Range Planning*, 52 (3), pp. 326-349

Wimelius, H. Mathiassen, L. Holmström, J. Keil, M, (2021), "A paradoxical perspective on technology renewal in digital transformation", *Information System Journal.*, 10.1111.

Yeow, A. Soh, C. Hansen, R. (2017), "Aligning with new digital strategy: a dynamic capabilities approach", *Journal Strategic Information Systems*, 27 (1), 43-58.

Yin, R. K. (2018). *Case study research and applications: Design and methods*, 6th ed; SAGE.

16. The Impact of Generative Artificial Intelligence on Employees' Creativity in the Workplace: A Literature Review and Research Agenda

Kyrillous George
German University in Cairo
Kyrillous.naguib@dba.guc.edu.eg

Nada M. ElSehemy
German University in Cairo
Nada.elsehemy@dba.guc.edu.eg

Abstract

Two of the main factors that impact businesses in today's world are Artificial Intelligence and Creativity. Workplace creativity is important for the success of any organization especially in today's competitive environments. Artificial Intelligence -AI, is one of the most pressing issues in the 21st century with lots of literature discussing its impact on organizations and employees. However, there is little research on how AI can influence employee workplace creativity.

This paper is trying to understand the impact of AI and especially a type of AI that can be used to create new content which is Generative AI- Gen AI on employees' workplace creativity by conducting a literature review on Generative AI and factors that influence employee workplace creativity.

A hypothesized model is presented from the findings that suggests that there are three different ways in which Gen AI can affect employees' workplace creativity. The three ways suggested by the model include a relationship that is fully mediated by efficiency and enabling employees to have more free time to create and innovate, the second is through collaborative AI and is moderated by employee skills and the third is by using Gen AI to keep the employees engaged in a creative process and customizing their workplace (through impacting the factors like daily work experiences, workplace fun, happy music) having both, psychological safety and employee aspects as moderators in the process.

This model can be used to facilitate future research and empirical studies that can be conducted to validate the findings and answer the research question about whether Gen AI can make the workplace more human; by empowering and enabling employees to create and add more value, or less human; by moving the creative activities from humans to machines or replacing humans with machines.

Keywords: Artificial Intelligence, AI, Generative Artificial Intelligence, Gen AI, Creativity, Employee Workplace Creativity, Organizational Behavior.

1. Introduction

In its simplest form, Creativity can be defined as the production of ideas that are novel and useful (Amabile, 1996). Creativity in the workplace has been considered as an important ingredient in the recipe of success for any organization (STAŠKEVIČA, 2019). Because of the

increasing competition, creativity and innovation help organizations stand out from their competitors. In today's modern business environment, the success of an organization in today's business environment relies largely on employees bringing forth creative and useful ideas for problem solving and for product or process innovation (Thompson, 2003).

While Artificial Intelligence, or AI, as defined by McCarthy "the science and technology of creating intelligent machines, especially intelligent computer programs" (Weber, 2023)" where machines can conduct decision-making tasks that require high level of intelligence, like reasoning and problem solving. Generative Artificial Intelligence, or Gen AI, is also defined as the machine technology that can create human-like novel content (Grimes, 2023), it can generate texts, images or other media and is thus mostly related to creativity. It is obvious that there are two reasons that make the Gen AI an important area of concern in today's economy, namely; impact and growth rate. The first reason is the expected strong impact and the value added to the business due to adopting Gen AI in the workplace. For example, (PwC, 2018) forecasts that AI will lead the growth of the global economy by 9.1 trillion USD by 2030 according to one report and up to around 13 trillion USD in another report (Bughin, 2017). This impact can be either on the employment, efficiency and productivity, the quality of work, or even the ethical implications resulting from adopting Gen AI in the workplace as mentioned by (BELKACEMI, 2022). The other reason Gen AI is a main concern is due to the dramatic growth rate; Gen AI is expected to have unprecedented adoption in a very short time unlike any other technology, as it is predicted to be adopted by up to 75% of companies in the coming 5 years (World Economic Forum, 2023). Hence, it is important to discuss the role Gen AI will play in the workplace focusing on Employees.

There is literature on the impact of AI on specific jobs/industries, there is also some general research on the impact of Gen AI on other Organizational Behavior factors, however, there was little research on how Gen AI can impact employees' creativity at the workplace. The contribution this paper makes is the presentation of a literature review on two concepts that are important to business success, i.e. Creativity and AI namely Gen AI, and trying to fill the research gap in literature by exploring how Gen AI can positively or negatively Impact the Employees' creativity in the Workplace. It goes further beyond that by providing a theoretical model that suggests how Gen AI can be used to enhance employee creativity, a model that is suggested to be further validated with empirical research.

This paper is divided into six sections. The next section is dedicated to shed light on methodology and the research plan. It is followed by section three and four, each independently discussing the literature available on the two main concepts of the paper namely; Creativity, and Gen AI; literature is used in these sections to define Creativity and highlight the main factors and theories that influence it and to explain AI and Gen AI as well as how they have developed over time. In section five, the two concepts are brought together by trying to highlight the relationship between Generative Artificial Intelligence and the factors that influence Creativity in an attempt to find how Gen AI can impact employees' creativity in the workplace and define factors that, if improved can leverage the benefits of this tool for the good of humanity. These factors are plotted in a theoretical model that is presented at the end of the section. The paper then is concluded with a recommendation on future research agenda.

2. Methodology and Research Plan

EBSCOhost Research Platform and Google Scholar were used to search for several keywords, including Artificial Intelligence, Generative Artificial Intelligence, Creativity, Workplace Creativity, Innovation, Creativity and Artificial Intelligence or Gen AI. The search was limited to academic, peer reviewed, recently published articles and included studies that either explain the concepts used i.e. Creativity, Gen AI, etc., or examine how Gen AI can directly or indirectly impact employee's creativity by impacting some of the variables that can accordingly impact creativity. Papers were narrowed down after reading the abstract and conclusion to 10 articles that discussed creativity, 21 that discussed AI and 28 that discussed the direct or indirect impact of AI on employee's creativity, few selected business sources and/or industry reports were added later whenever relevant. In addition, some relevant research papers that were referenced in the pool of articles selected were referenced to make sure the topic is covered from several aspects.

This paper aims at tackling how Gen AI can positively or negatively Impact the Employees' Creativity in the Workplace and to prepare for a research journey that answers the following research question, will Gen AI make the workplace more human? By enabling humans and freeing employees from the manual work that can be automated to focus on more value-adding and creative work? Or will the workplace be less human, by replacing humans with machines that can provide a creative novel content. To answer that question a new model is presented to hypothetically explain the relationship between Gen AI and the Employees' Creativity in Workplace based on existing literature provided in this paper that may facilitate the scientific process in this area.

3. Creativity Concept Definition and Factors that Influence Creativity

In this section, the concept and definition of Creativity will be explained as a prerequisite before discussing how AI impact Employee's creativity in section 5. Findings from literature presented in this section is used to build the model presented in section 5 below.

Creativity and intellectual abilities were considered distinct to human beings (Korinek, 2023). In its simplest forms, Creativity can be defined as the production of ideas that are novel and useful (Amabile, 1996). It also often entails the identification of problems that are worth solving (Baer, Dirks, & Nickerson, 2013; Unsworth, 2001). The authors of this paper used the first definition of Creativity which is the production of ideas that are novel and useful and it is explored among employees in their workplace.

What influences Creativity? And how to be more creative? Has been the subject of lots of research. Within literature, creativity has been studied both as a stable individual difference and as a transitory state of mind prompted by situational and/or affective states (Helzer et al., 2019). The belief that increased creativity will yield better organizational performance and that certain situational factors can increase creativity has impelled researchers to look into certain factors or theories that will make employees more creative. In this section, the paper highlights some of the research that has tackled this area with a focus on some factors that may impact employees' creativity which are work role identity conflict, daily workplace experiences, music, workplace fun and other employees' aspects. It is important to note that other factors

exist but only the main factors that Gen AI can be used to augment are discussed, and by this, the way is paved to section 5, where Gen AI impact on creativity will be studied using those factors.

3.1 Work Role Identity Conflict and Workplace Creativity

Multitasking is very common in organizations and it can result in conflicting roles. A study conducted by (Akkan & Guzman, 2022) argues that these conflicting roles or in other words; work role identity conflict can put people in a problem-solving mode, driving them to deliberately and thoroughly ideate to alter the resources in their workplace. They hypothesized that work role identity conflict is positively related to creative process engagement which shall boost creativity. They also stressed on the importance of having a positive relationship between employees and their supervisors to help them make use of the creative ideas in ways that benefit the organization. The results suggest that work role identity conflict does not directly drive workplace creativity. However, work role identity conflict should trigger creative process engagement (acting as a mediator) and employees should have high levels of relational identification with their supervisors (acting as a moderator), for creativity to increase (Akkan & Guzman, 2022).

3.2 Daily Workplace Experiences and Workplace Creativity

Creative climate research shows that an employee's perception of their work environment impacts their creative performance (McKay et al., 2022). A study was conducted with the goal to better understand how people perceive their daily work environment across the lifespan of organizational projects and how this impacts their creativity in the workplace. The researchers (McKay et al., 2022) used a person-centered approach to creative climate and identified five different workplace experiences namely; Toxic Day, Disengaged Day, Typical Day, Ideal Day and Crisis Day. They then conducted a study using a diary experiment – where data is reported by participants over a span of time – to validate the different workplace experiences identified and evaluate their impact on workplace creativity across the span of organizational projects. The aforementioned results validated the different workplace experiences and supported that Ideal Days are important for creative performance in comparison to other days.

The researchers included nine indicators in their survey to rate these workplace experiences, namely; Freedom, Challenging work, Resources, Work group support, Supervisor encouragement, Organizational encouragement, Time pressure, Political problems, Low risk – conservative attitude, those indicators were either considered stimulant factors or obstacle factors (McKay et al., 2022). Although this research didn't explain exactly how to mimic an Ideal Day, yet it gives us insights about how important it is to manage employee perception on the daily workplace experiences if the aim is to improve their creativity.

3.3 Music and Workplace Creativity

Music is recognized as an important environmental condition for work performance (Landay & Harms, 2019). A study by (Zhu et al., 2022) draws upon this by trying to examine the effect of listening to happy music in the workplace on the creative performance of Chinese employees. Acknowledging that creative behavior is inherently risky and that to engage in creative activities, employees need to perceive that they have psychological safety support (Liu and Ge, 2020), the researchers draw upon the concept of psychological safety in the workplace which is an emotional state in which people feel safe when taking interpersonal risks in the

workplace (Edmondson, 1999). Results of the study showed that employee's perception of happy music enhances their creative performance both directly and also indirectly via enhanced perception of psychological safety (acting as a partial mediator). These findings reinforce the value of employees' perception of happy music for stimulating their creativity (Zhu et al., 2022).

3.4 Workplace Fun and Workplace Creativity

Having fun at work can increase the bonds that employees have with each other and with the organization, creating happy moments and reducing stress, all of which can improve the performance of the employees and increase their loyalty to their organizations. Workplace fun is defined as "playful social, interpersonal, recreational, or task activities intended to provide amusement, enjoyment, or pleasure" (Lamm & Meeks, 2009). According to (Ford et al., 2004), researchers have suggested that workplace fun is a crucial precursor of employee creativity, and have pointed out that having fun in the workplace may increase creativity among employees. A paper by (Yang G., 2020) draws upon a study that was conducted on Chinese employees to explain how workplace fun is effective in facilitating employee creativity, with a focus on the mediating role of psychological safety in this relationship. The results of the study validated that workplace fun is positively related to employee creativity, and that psychological safety was a mediator in this relationship (Yang G., 2020).

3.5 Employee Aspects and Workplace Creativity

In an attempt to understand the theories on creativity that has been introduced in literature, several theories have been presented in a literature review paper by (Ameen et al., 2022). A number of these theories can be grouped under the umbrella of employee aspects, namely; Big Five Personality Traits, Intrinsic Motivation theory and Empowerment theory. The idea is that sometimes creativity is triggered by certain employee or leaders' aspects and not by external factors. According to the big five personality traits theory, conscientiousness and agreeableness have a negative association with creativity (Kwang & Rodrigues, 2002). According to the Intrinsic Motivation Theory, intrinsically motivated employees are more likely to search for new and novel ways to create products and services; and take greater risks (Amabile et al., 1990). Finally, the Empowerment Theory focuses on how leaders can encourage creativity (Cheng et al., 2019). It comprises processes as well as outcomes, which implies that certain acts, activities, and structures can be empowering and that the conclusion of such structure results in a state of empowerment in terms of creativity (Ameen et al., 2022).

4. Artificial Intelligence Concept Definition, History and Evolution

In this section the concept and evolution of AI will be discussed in order to answer the research question, how AI, and more precisely Gen AI, can Impact the workplace creativity? The aim is also to consider how Gen AI can result in producing a creative novel content which is a main topic in the proposed theoretical model in section 5.5 in this paper.

In1950, a groundbreaking breakthrough was announced, when a small mouse-shaped robot could successfully not only finish a maze, but also remember the course taken all the way through the end (World Economic Forum, 2022). The year 1977 marked another important milestone, when an IBM machine could beat Kasparov, the legendary all-time best chess player (Iglesias and Pablo, 2020). Another example was Google's AlphaGo software that could beat

the world champion of the board game in 2016 (Moyer C., 2016). All these incidents rang a bell that AI could one day outperform the human Intelligence. (McKinsey, 2023) has stated a definition that shed light on the contribution AI can make in conducting creative tasks as McKinsey defines Gen AI as "algorithms - such as ChatGPT - that can be used to create new content, including audio, code, images, text, simulations, and videos." This definition emphasizes how machines create different forms of novel content. However, most of the literature and reports discussing the implications of AI in the workplace have tackled the powerful impact of automation of manual activities done by AI, leaving behind the innovative and creative tasks that can be done by AI and might have a different type of impact on the workplace.

4.1 Artificial Intelligence Creativity

Gen AI, was found able to conduct creative tasks that require even high level of thinking and creativity (Korinek, 2023). In this article the term "AI Creativity" will be used to refer to the Creative tasks conducted by a machine using AI or Gen AI and to differentiate it from the Human Creativity. Many efforts are exerted to replicate the Human capabilities in Intelligence by machines

Since the topic is relatively new, only few recent papers were found to cover the creative output of the Gen AI. For example, in a study conducted by (Phillips and Robie, 2023), Gen AI GPT-4 version of Open AI was found superior in faking the personality test question and not only to the other Large Language Models (LLM) like Google's bard and even the previous version introduced by the same company, but also to a sample taken by humans who faked the same questions. In addition, the difference between the answers provided by humans were found similar to those provided by the LLM. While faking is a task that required creativity according to (Levashina et al., 2009), (Chen, 2023) still argues that the current Capabilities of Artificial Intelligence is currently capped at the "Imitation Barrier" where AI can conduct tasks that show a significant level of Pattern, Prediction, and Imitation and that up till now it cannot create, as the author argues that creation requires inspiration which is still a human-only task till now. However, there are still some concerns regarding the AI novel content reliability and factual mistakes found. (Monteith et al., 2023) highlighted that "*AI models include factual errors, inappropriate or dangerous advice, nonsense, fabricated sources and arithmetical errors*". (Koivisto, 2023) argues that humans can still "outperform AI in creative divergent thinking tasks". In conclusion, Gen AI is found to be capable of producing creative novel content. But capped by a certain ceiling which is the Imitation Barrier and in the following section, the collaboration between the AI creativity and human creativity will be discussed further.

4.2 Gen AI and Customized Customer Experiences

So far, Gen AI has been used to create immersive experiences for customers, making companies with such applications stand out and reach for more adventurous customers who are shopping for the new limits that are constantly being broken. Very little research is conducted though on how these experiences can impact employees' creativity if conducted at the workplace.

The year 2023 has seen press releases from companies expanding on the use of Gen AI to create customized experiences for their customers without taking away their customers' needs to

create. SundaySky, a leading software platform for efficient video creation and personalization, recently announced the release of its Gen AI assistant, Copilot, for its platform users, in the press release the company mentioned it does not intend to take video creation out of the hands of its customers, but fundamentally believes that AI is an efficiency accelerator for its business users to stimulate their creativity, scale the quality of their video content and widen the diversity of their media mix (Eisenhauer, 2023). A similar announcement was made by Kaleidoscope, an interactive record label, when releasing their "Space for Sleep (Kaleidoscope Remix), the release mentioned that by using Gen AI technologies to extend and remix original compositions, the company brings listeners an array of immersive content, designed to accompany specific moods and experiences (LifeScore music, 2023). The year has also seen the announcement by PopSockets, a brand leader in innovative phone accessory solutions, about the launch of its groundbreaking AI Customizer that integrates generative AI by introducing a tool that redefines personalization and self-expression. Among other things, the tool is advertised to have an AI customizer that offers ideas and suggestions to jumpstart the creative process (PopSockets, 2023). All these and similar announcements show how generative AI has successfully been used to make customers' experiences more customized and fun.

5. The Impact of AI on Factors that Influence Employee's workplace Creativity

So far, the paper has discussed the main factors that can influence Creativity as proven from previous literature as well as how AI and Gen AI have developed and are being used to create customized and unique customers' experiences. In this section, these findings are used to explore how Gen AI can influence employee creativity by directly affecting the factors that influence creativity outlined above. A model is plotted at the end with the findings.

5.1 Gen AI and Creative Process Engagement

The relationship between work role identity conflict and creativity was found to be fully mediated by the creative process engagement. Emphasis was placed on the relationship between the employees and their supervisors to ensure the positive impact on employee creativity (Akkan & Guzman, 2022). Hence, it can be deduced that the actual factor that triggers employees to be more creative is the creative process engagement.

By using the recent practical Gen AI inventions, like the (PopSockets, 2023) AI customizer that offers ideas and suggestions to jumpstart the creative process, employees can be kept engaged in the creative process with the help of Gen AI which will lead to higher employees' creativity. Research needs to be conducted to validate that and to find ways to keep employees engaged and connected to their workplace and their supervisors to ensure that better employee creativity at the workplace is in fact achieved.

5.2 Gen AI and Employee Customization

Previous studies conducted has shown that;

- Employees' daily work experiences has a direct impact in increasing the level of their creativity. (McKay et al., 2022).

- Employees' perception of happy music enhances their creative performance both directly and also indirectly via enhanced perception of psychological safety (acting as a partial mediator) (Zhu et al., 2022).
- Workplace fun is positively related to employee creativity, and that psychological safety was a mediator in this relationship (Yang G., 2020).

All of these 3 factors can be combined into one consolidated factor that can be called employee customization. Whereby, recent Gen AI inventions can be used to create a customizable and fun experience for employees enabling a more employee-centered workplace and allowing for the individualized or group happy integrative music playlists that can trigger employees' creativity and increase the bonds among colleagues. It can also be used to identify the Ideal days for each employee and try to maximize them by replicating the factors that cause them for each individual. Noting that each employee's personal traits or "Employee Aspects" as defined in section 3.5 above will play an important role in the receptivity of each employee to any Gen AI intervention.

Those individualized integrative interactions are not far from reality and have already been introduced to the market and used for providing an immersive experience for customers by companies like Kaleidoscope (LifeScore music, 2023) and SundaySky (Eisenhauer, 2023), they just need to be extended to employees in the workplace.

An important identified aspect that causes the increased creativity in the above-mentioned cases was the psychological safety of employees. Because AI in general has been viewed as a source of stress and frustration due to the threats it imposes on employees who are constantly fearing that they might be replaced, it can be argued whether creativity can be reached eventually if employees are feeling threatened. This assumption needs to be validated by empirical studies while considering that the relationship between stress and creativity is bidirectional. On one hand, increased stress will lead to less psychological safety and less employee creativity, but on the other, research has found that increased creativity can be considered as a resource that can promote well-being by promoting flexible stress responses (Helzer et al., 2019).

5.3 Collaborative AI

On another front, AI can be used in an integrative or collaborative way to serve the human creativity. (Kayoed, 2023) argued that a kind of collaborative relationship may be shaped between the Human Intelligence and the AI Creativity, while AI can create creative content, the human intelligence is still needed to add context to this content, and shape how it fits in the big picture, in addition, the human can also add an emotional touch to the creative novel content prepared using the AI creativity, and accordingly the Human creativity and AI creativity can collaborate.

Another aspect that may be considered is the factual mistakes provided in LLM's Gen AI answers and content that foster the collaborative relationship between the employee in the workplace and the AI tools that cannot, up till now, provide a standalone creative solution according to (Monteith et al., 2023), while (Jian, 2024) provided an empirical evidence that AI can support increasing the impact of employee's creativity in a telemarketing company which

is moderated by the level of employees' skills this synergetic effect is called "AI-augmented employee creativity".

5.4 Efficiency and time spare

A third aspect that needs to be studied is the spare time that employees have due to certain tasks being handed over to AI. Gen AI can be used to create certain content that would have otherwise been created by humans. The time spared for employees can be tested if it actually increases their creativity by giving them more time to excel in unique tasks that couldn't be handled by AI tools. (Lăzăroiu, 2023) supported this argument that Gen AI can change the shape of workplace employment by automating some activities, (Herm, 2022) added that AI can be used in robotic process automation which is defined in the same study as "a technology to automate already digital but yet manual tasks and sub processes as well as whole business processes rapidly". This will definitely lead to an efficient workplace.

5.5 Suggested hypothesized Gen AI – Employee's Workplace Creativity Model.

In light of the previous discussion, a new hypothesized model is introduced in Figure 1 suggesting the impact of Gen AI on the employees' workplace creativity, through a group of mediators and moderators. The model hypothesizes, according to literature, that there are three different ways by which Gen AI can affect the creativity of employees in the workplace. The first way is fully mediated by increasing the workplace efficiency. This will in turn free employees from non-value-adding manual activities, enabling them to have more time for creativity and innovation. This can be achieved by AI-powered machines taking on the manual tasks. The second way is through the collaboration between human creativity and machine creativity according to (kayoed, 2023) that is hypothesized to increase the quality of the creative novel content produced being moderated by the employees' skills (Jian, 2024). The third way by which Gen AI is hypothesized to affect Employees' Workplace Creativity is by using Gen AI to keep the employees engaged in a creative process and customizing their workplace (through impacting the factors like daily work experiences, workplace fun, happy music) having both, Psychological Safety and Employee Aspects as moderators in the process. This model should be validated by conducting empirical research.

6. Conclusion

AI is considered one of the fast-evolving topics in the workplace. This article studies the impact of Gen AI on the Employees' Workplace Creativity and suggests a new hypothesized model – derived from literature review – for how Gen AI can affect the employees' workplace creativity through 3 different ways. This model can be used to facilitate future research and empirical studies that can be conducted to validate the findings. Besides, it can answer the research question about whether Gen AI can make the workplace more human; by empowering and enabling employees to create and add more value, or less human; by moving the creative activities from humans to machines or replacing humans with machines.

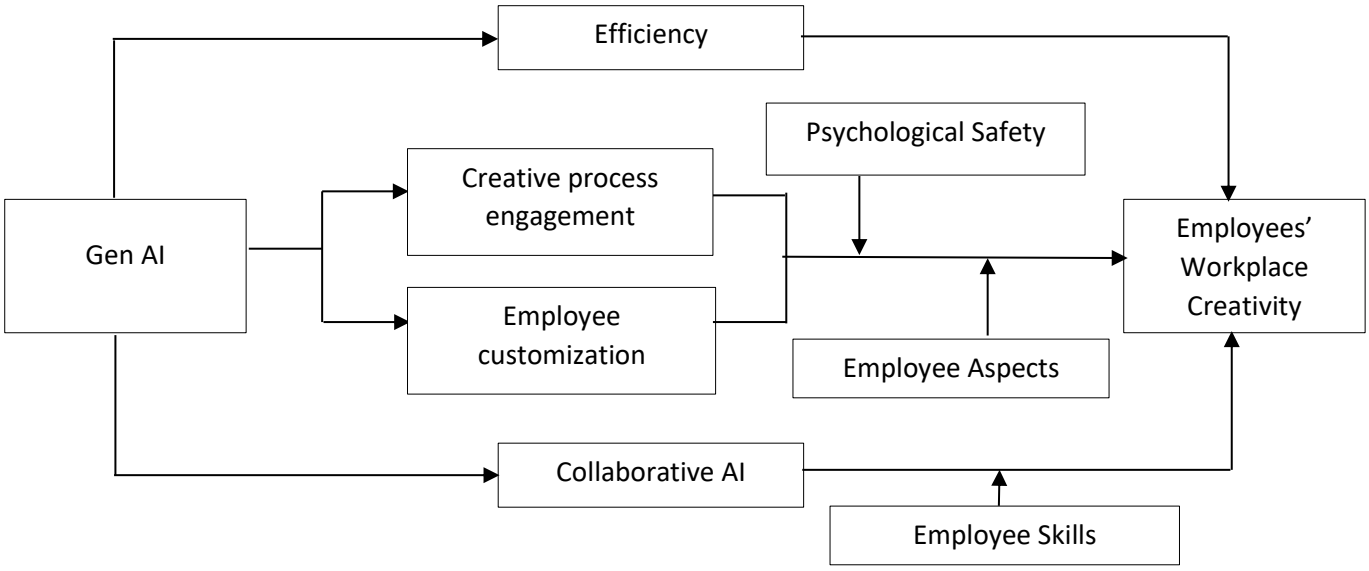


Figure 1: Gen AI-Employees' Workplace Creativity Model

The 3 ways suggested by the model include a relationship that is fully mediated by efficiency and enabling employees to have more free time to create and innovate, the second is through collaborative AI and is moderated by employee skills and the third is by using Gen AI to keep the employees engaged in a creative process and customizing their workplace (through impacting the factors like daily work experiences, workplace fun, happy music) having both, psychological safety and employee aspects as moderators in the process.

6.1 Future research agenda

Further research is recommended to empirically test each relationship between AI; Gen AI as an Independent Variable and Creativity as a dependent variable, including the suggested mediators and moderators in the model, more variables may be suggested according to the context,. Further research is also required to test whether the model applies equally for all employees; the impact on creativity may differ depending on where the employees are in the organization chart and also on the type of industry. Various considerations are to be taken like the different types of creativity and the different methods by which Creativity is measured.

References

Akkan, Eren & Guzman, Felipe. (2022). When discordant work selves yield workplace creativity: The roles of creative process engagement and relational identification with the supervisor. *Journal of Occupational and Organizational Psychology*. 95. 184–208 © 2021 British Psychological Society.

Amabile, T. M. (1996). Creativity in context: *Update to the social psychology of creativity*. Boulder, CO: Westview Press.

Amabile, T. M., Goldfarb, P., & Brackfleld, S. C. (1990). Social influences on creativity: Evaluation, coaction, and surveillance. *Creativity Research Journal*, 3(1), 6–21.

Ameen, Nisreen & Sharma, Gagan & Tarba, Shlomo & Rao, Amar & Chopra, Ritika. (2022). Toward advancing theory on creativity in marketing and artificial intelligence. *Psychology & Marketing*. 39. n/a-n/a. 10.1002/mar.21699.

Baer,M., Dirks,K. T.,&Nickerson, J.A. (2013). Micro foundations of strategic problem formulation. *Strategic Management Journal*, 34: 197–214.

Bankins, S., & Formosa, P. (2023). "The Ethical Implications of Artificial Intelligence (AI) For Meaningful Work." *Journal of Business Ethics*, 185(4), 725–740.
Belkacemi, S. (2022). "Artificial Intelligence "AI" and its impact on global Economy". *Journal of Financial, Accounting & Managerial Studies*, 9(2), 269–285.
Benbya H, Davenport TH, Pachidi S. (2020), "Special Issue Editorial: Artifificial Intelligence in Organizations: Current State and Future Opportunities." *MIS Quarterly Executive*, 19(4):ix-xxi.
Bughin, J., Seong, J., Manyika, J., Chui, M., & Joshi, R. (2018). "Notes from the AI frontier: Modeling the impact of AI on the world economy". *McKinsey Insights*, 1.
Cheng, C., Cao, L., Zhong, H., He, Y., & Qian, J. (2019). The influence of leader encouragement of creativity on innovation speed: Findings from SEM and fsQCA. Sustainability (Switzerland), 11(9), 2693. https://doi.org/10.3390/su11092693
Chen, Rowland, Roger B. Dannenberg, Bhiksha R., and Rita S. (2020). "Artificial Creative Intelligence: Breaking the Imitation Barrier." *Proceedings of the 11th International Conference on Computational Creativity (ICCC'20)*. Accessed Jan 4, 2024. https://computationalcreativity. net/iccc20/papers/026-iccc20.pdf.
de Cassia Nakano T, Wechsler SM, Campos CR, Milian QG. (2015) "Intelligence and Creativity: Relationships and their Implications for Positive Psychology". *Psico-USF*, 20(2):195-206.
Edmondson, A. C. (2004). Psychological safety, trust, and learning in organizations: A group-level lens. In R. M. Kramer & K. S. Cook (Eds.), Trust and distrust in organizations: Dilemmas and approaches (pp. 239–272). Russell Sage Foundation.
Eisenhauer, Rachel – SundaySky (2023). *SundaySky Introduces Generative AI to Assist and Accelerate Video Creation.* Retrieved 2 January, 2024, from https://sundaysky.com/resources/gen-ai-video-creation/
Escudero Iglesias P. "Deep Thinking: Where Machine Intelligence Ends and Human Creativity Begins." *Empresa y Humanismo. 2020*, 23(2):139-142.
Ford, R. C., Newstrom, J. W., & McLaughlin, F. S. (2004). Making workplace fun more functional. *Industrial and Commercial Training*, 36(3), 117–120. https://doi.org/10.1108/00197850410532131
Grimes, M., von Krogh, G., Feuerriegel, S., Rink, F., & Gruber, M. (2023). From Scarcity to Abundance: Scholars and Scholarship in an Age of Generative Artificial Intelligence. Academy of Management Journal, 66(6), 1617–1624. https://doi.org/10.5465/amj.2023.4006
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.
Harper, G. (2023). "Artificial Creative Intelligence (ACI)". Creative Industries Journal, 16(3) 263–264.
Helzer Erik G. Kim Sharon H. 2019. Creativity for Workplace well-being. *Academy of Management Perspectives*, Vol. 33, No. 2, 134–147.
Hough, L. M. (1992). The "Big Five" personality variables construct confusion: Description versus prediction. *Human Performance*, 5(1–2), 139–155. https://doi.org/10.1080/08959285. 1992.9667929
Herm, L.-V., Janiesch, C., Helm, A., Imgrund, F., Hofmann, A., & Winkelmann, A. (2023). A framework for implementing robotic process automation projects. Information Systems & E-Business Management, 21(1), 1–35. https://doi.org/10.1007/s10257-022-00553-8
Jauk, Emanuel & Benedek, Mathias & Dunst, Beate & Neubauer, Aljoscha. (2013). "The relationship between intelligence and creativity: New support for the threshold hypothesis by means of empirical breakpoint detection. Intelligence." *PubMed*, 41(4). 212-221.
Jia, N., Luo, X., Fang, Z., & Liao, C. (2024). When and How Artificial Intelligence Augments Employee Creativity. Academy of Management Journal, 67(1), 5–32. https://doi.org/10.5465/amj.2022.0426
Kayoed, S. (2023) "The Symbiotic Relationship: Exploring the Intersection of Artificial Intelligence and Human Intelligence", *Tech-E,* December 2023.
Korinek, A. (2023). "Generative AI for Economic Research: Use Cases and Implications for Economists". *Journal of Economic Literature*, 61(4), 1281–1317.
Koivisto, M., & Grassini, S. (2023). Best humans still outperform artificial intelligence in a creative divergent thinking task. Scientific Reports, 13(1), 1–10. https://doi.org/10.1038/s41598-023-40858-3
Kwang, N. A., & Rodrigues, D. (2002). A big-five personality profile of the adaptor and innovator. *Journal of Creative Behavior*, 36(4), 254–268. https://doi.org/10.1002/j.2162-6057.2002.tb01068.x
Lamm, E., & Meeks, M. D. (2009). Workplace fun: The moderating effects of generational differences. *Employee Relations*, 31(6), 613–631. https://doi.org/10.1108/01425450910991767
Landay, K., & Harms, P. D. (2019). Whistle while you work? A review of the effects of music in the workplace. *Human Resource Management Review*, 29(3), 371–385. https://doi.org/10.1016/j.hrmr.2018.06.003.

Lăzăroiu, G., & Rogalska, E. (2023). How generative artificial intelligence technologies shape partial job displacement and labor productivity growth. Oeconomia Copernicana, 14(3), 703–706. https://doi.org/10.24136/oc.2023.020

Levashina J, Morgeson FP, Campion MA.(2009). "They don't do it often, but they do it well: Exploring the relationship between applicant mental abilities and faking." *International Journal of Selection and Assessment*, 17(3):271-281.

LifeScore Music. (10 C.E. 2023). *Kaleidoscope Announces First Artist Collaboration - 8 Hour Sleep Album With Sleeping At Last. Business Wire (English)*. Retrieved 2 January 2024, from https://www.businesswire.com/news/home/20231030926221/en/

Liu, K., & Ge, Y., (2020). How psychological safety influences employee creativity in China: Work engagement as a mediator. *Social Behavior and Personality: An international journal*, 48(8), Article 9211.ttps://doi.org/10.2224/sbp.9211

McKay, Alexander & Mohan, Mayoor & Reina, Christopher. (2022). Another Day, Another Chance: Daily Workplace Experiences and Their Impact on Creativity. *Journal of Product Innovation Management*. 39. 292-311. 10.1111/jpim.12573.

McKinsey Insights (2023 January 19) *What is generative AI? January 2023*: N.PAG. Retrieved 28 December 2023, from https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai

Monteith S, Glenn T, Geddes JR, Whybrow PC, Achtyes E, Bauer M. (2023), Artificial intelligence and increasing misinformation. *The British journal of psychiatry: the journal of mental science*. doi:10.1192/bjp.2023.136

Moyer, C. (2016, March 28). How Google's AlphaGo beat a Go world champion [Atlantic]. Retrieved from https://www.theatlantic.com/technology/archive/2016/03/the-invisible-opponent/475611/

Phillips J, Robie C, (2023) "Can a computer outfake a human?" *Personality and Individual Differences*. 217, (112434), ISSN 0191-8869,

PopSockets (2023). *POPSOCKETS REVOLUTIONIZES ACCESSORY CUSTOMIZATION WITH NEW AI TOOL. PR Newswire*. Retrieved 2 January 2024, from https://www.prnewswire.com/news-releases/popsockets-revolutionizes-accessory-customization-with-new-ai-tool-301925883.html

PwC, (2018), "The macroeconomic impact of artificial intelligence", *PwC*. Retrieved 4 January 2024 from https://www.pwc.co.uk/economic-services/assets/macroeconomic-impact-of-ai-technical-report-feb-18.pdf

Qu, R., Janssen, O., & Shi, K. (2015). Transformational leadership and follower creativity: The mediating role of follower relational identification and the moderating role of leader creativity expectations. *The Leadership Quarterly*, 26(2), 286–299. https://doi.org/10.1016/j.leaqua. 2014.12.004

Staškeviča, A., & Dunska, M. (2019). "The Importance of Creativity for Business Development in Latvia", *Journal of Economics & Management Research*, 8, 36–54.

Thompson, L. (2003). Improving the creativity of organizational work groups. *Academy of Management Perspectives*, 17(1), 966–109. https://doi.org/10.5465/AME.2003.9474814.

Unsworth, K. 2001. Unpacking creativity. *Academy of Management Review*, 26: 289–297.

- Weber, F. (2023). Artificial Intelligence. In: Artificial Intelligence for Business Analytics. Springer Vieweg, Wiesbaden. https://doi.org/10.1007/978-3-658-37599-7_2

World Economic Forum (2022). *How has AI developed over the years and what's next?* Retrieved 12 December 2023, from https://shorturl.at/gqy45

World Economic Forum (2023). *Future of Jobs Report 2023*. Retrieved 12 December 2023, from https://www.weforum.org/publications/the-future-of-jobs-report-2023/

Yang, G. (2020). Workplace fun and employee creativity: The mediating role of psychological safety. *Social Behavior and Personality: An international journal*, 48(11), e9510

Zhang, X., & Bartol, K. M. (2010). Linking empowering leadership and employee creativity: The influence of psychological empowerment, intrinsic motivation, and creative process engagement. *Academy of Management Journal*, 53(1), 107–128. https://doi.org/10.5465/ amj.2010.48037118

Zhu, Q., Li, W., & Chen, Y. (2022). Happy music and employee creativity in the workplace: Psychological safety as a mediator. *Social Behavior and Personality: An international journal*, 50(4), e11390.

17. The Impact of Seeing Human Involvement in Sensor-Based Journalism on Reader Appreciation
(Research-in-Progress)

Irina Boboschko
University of Cologne
irina.boboschko@uni-koeln.de

Abstract
Datafication disrupts journalistic processes and intensifies economic challenges for the journalism industry. The increasing use of vast amounts of sensor data in journalism makes journalistic processes opaque, thereby decreasing the visibility of human involvement within the inherent sensor-based processes, such as sensor data collection or analysis. While consumers are averse towards opaque data-based products, process transparency gained growing research interest to increase consumer appreciation by enhancing sensemaking. Previous research investigating physical product appreciation finds that particularly the visibility of human work implied in the production processes increases consumer appreciation. For sensor-based contexts, current research still lacks insights into the impact of seeing human involvement in sensor-based processes. To address this gap, we use sensor-based journalism as research context and draw on the literature on process transparency and social translucence. Hence, in this research-in-progress paper, we hypothesize the impact of seeing human involvement in sensor-based processes on reader appreciation for sensor-based articles. We propose a research model and suggest forthcoming research steps to test the model by conducting an online experiment among more than 200 participants based in Germany.

Keywords: Appreciation, Process Transparency, Sensor-Based Journalism, Social Translucence.

1. Introduction
In the past decades, digitization posed economic challenges for the journalism industry as it reduced subscriptions, advertising earnings, and readers' willingness-to-pay for journalistic content (Lewis & Westlund, 2015). Due to vast amounts of data and thereof resulting datafication, sensor-based journalism gained momentum by accelerating journalistic processes, like capturing phenomena in (near) real-time (Diakopoulos, 2019; Loebbecke & Boboschko, 2020). However, sensor-based journalism became increasingly opaque, for instance, through the reduced visibility of human involvement within the sensor-based processes, such as sensor data collection or analysis (Diakopoulos, 2019). In this context, Information Systems (IS) research emphasizes the 'invisible' work behind a data-based product as the so-called 'backrooms of data science' (Parmiggiani et al., 2022).

Recent literature finds opaque data-based products to reduce consumer appreciation (Koenig et al., 2022; Mahmud et al., 2022). As a counteracting mechanism, making the data-based processes transparent for consumers gained growing scholarly attention (Fernandez-Loria et

al., 2022; Haque et al., 2023; Rai, 2020). For example, previous research on data-based journalism uncovers that explaining details on algorithms and input data increases readers' trust (Shin, 2021a; Shin, 2021b).

In this context, recent computer science research adds that when explaining data-based products, particularly the visibility of contextual, human interactions with a data-based product increases consumer trust and enhances decision-making, holistic explainability, and sensemaking (Ehsan et al., 2021). However, given the diminished visibility of human involvement in sensor-based journalism, investigations into the effects of making human involvement within sensor-based processes visible remain still scarce. With this regard, prior marketing research in the context of physical products finds, for instance, that making human work visible through so-called 'handmade' labels signals high production efforts and hence increases consumers' willingness-to-pay (Fuchs et al., 2015; Kruger et al., 2004).

By addressing this research gap, we ask how the visibility of human involvement in sensor-based journalism affects reader appreciation for sensor-based articles. Thereby, we leverage literature related to process transparency by emphasizing additional aspects suitable for explanations beyond solely technical insights (Haque et al., 2023; Rai, 2020). Also, we refer to insights on social translucence (Ehsan et al., 2021; Erickson & Kellogg, 2000) by highlighting the influence of making human involvement inherent in the sensor-based processes visible when explaining those sensor-based processes to consumers. In this research-in-progress paper, we hence derive hypotheses and develop a research model. Finally, we present forthcoming research steps to test the model by conducting an online experiment.

2. Research Context

2.1 Sensor-Based Journalism

Sensor-based journalism enables the deployment of sensors as 'walking-data-generators' (Loebbecke & Boboschko, 2020; Newell & Marabelli, 2015). These sensors collect sensor data from moving objects or living creatures that are, for instance, either beyond human access or barely perceptible to the human eye (Diakopoulos, 2019; Loebbecke & Boboschko, 2020). They enable continuous and comprehensive real-time collection of sensor data - the so-called 'N=all' scenario (Jones, 2019). Thereby, they convert stimuli of physical phenomena into electronic signals (Bardhan et al., 2020; Monteiro & Parmiggiani, 2019).

In sensor-based journalism, the collection of sensor data changes journalistic knowledge generation (Diakopoulos, 2019; Kitchin, 2014). While journalists traditionally referred to testimony-based knowledge (qualitative epistemologies), deploying sensors and collecting sensor data facilitates evidence-based investigations (positivistic epistemologies). It enhances the generation and cross-verification of journalistically relevant information (Godler et al., 2020) and enables efficient, content creation at scale (Diakopoulos, 2019; Lewis & Westlund, 2015).

While journalism used sensor data for decades, e.g., for weather reports, recent computational advancements such as Natural Language Processing (NLG) have facilitated the immediate generation of sensor-based articles based on collected sensor data (Diakopoulos, 2019). Moreover, contemporary sensor-based articles often rely on sensor data gathered through

crowdsourcing techniques, wherein citizens participate in data collection relevant to journalism and public interests (Loebbecke & Boboschko, 2020). Examples of such sensor-based articles include real-time reports or live-tickers on continuously changing conditions, e.g., 'live' animal diaries or traffic updates (Loebbecke & Boboschko, 2020). Thus, we understand sensor-based articles as those that integrate vast amounts of sensor data, supported by real-time, continuous, and comprehensive sensor data collection processes, and are complemented by computational mechanisms like NLG. Consequently, the inherent human involvement in sensor-based processes increasingly fades from the readers' perspective (Diakopoulos, 2019).

2.2 Human Involvement in Sensor-Based Processes

In line with the literature on sociomateriality, human involvement in the sensor-based processes impacts the materiality and meaning of a final sensor-based product (Jones, 2019; Monteiro & Parmiggiani, 2019; Oberlaender et al., 2018). Thereby, the interplay between humans and sensor data mutually shapes the final sensor-based product (Diakopoulos, 2019; Godler et al., 2020; Jones, 2019; Oberlaender et al., 2018; Parmiggiani et al., 2022). Moreover, recent IS research sheds light on the importance of 'human-in-the-loop' tasks, pointing to human augmentation work in sensor-based processes, i.e., auditing and altering sensor-based processes (Gronsund & Aanestad, 2020; Parmiggiani et al., 2022).

For instance, contrary to assumptions about sensor data objectivity and neutrality (Kitchin, 2014), humans integrate their expertise, values, interests, and also their biases into sensor-based processes, like sensor data collection or analysis. In particular, humans choose phenomena to be represented by the sensor data, operationalize the variables of interest, and determine the spectrum for sensor data collection. Additionally, they participate in the curation of sensor data, including transformation, editing, cleaning, assembly and finally the analysis of the sensor data (Gronsund & Aanestad, 2020; Parmiggiani et al., 2022).

3. Theoretical Background

3.1 Process Transparency

While past research emphasizes consumer aversion towards opaque data-based products and services (Koenig et al., 2022; Mahmud et al., 2022), explaining and making data-based processes transparent gained growing importance as a counteracting mechanism in IS research (Fernandez-Loria et al., 2022; Haque et al., 2023; Rai, 2020). Such process transparency allows for communicating technical insights, e.g., into the collection and sampling of the data, operationalizations of the variables, the statistical model, and the logical rules embedded in the algorithmic analyses (Fernandez-Loria et al., 2022; Haque et al., 2023). Process transparency impacts how consumers perceive a product or a service (Buell et al., 2016; Haque et al., 2023). Thereby, process transparency helps reduce consumers' aversion towards opaque data-based products by increasing sensemaking (Haque et al., 2023). Furthermore, process transparency enhances trust, understandability, and observability; and reduces perceptions of bias underlying the data-based product (Haque et al., 2023; Shin, 2021a).

3.2 Social Translucence

Drawing on social translucence literature (Erickson & Kellogg, 2000), recent research suggests that making technical processes transparent is not sufficient. Instead, explanations for data-based products should further highlight contextual and human interactions with the data-based

products to enhance consumer appreciation (Ehsan et al., 2021). Human interactions can be communicated through so-called 'social cues' in text- or image-based formats. These cues include details on particular human identity attributes, e.g., profession, related to specific task responsibilities within a process (Erickson & Kellogg, 2000; Gefen & Straub, 2004; Ehsan et al., 2021). Seeing social cues may hence address knowledge gaps and consequently enhance consumer sensemaking, especially when data-based products and the inherent processes are perceived as complex or opaque (Ehsan et al., 2021; Gefen & Straub, 2004; Haque et al., 2023).

4. Hypotheses Development

According to the 'law of contagion', objects that are connected to humans may influence each other through the transfer of their attributes (Nemeroff & Rozin, 1994). In this context, consumers' association with those transferred attributes affects how they perceive those objects (Job et al., 2017). For instance, consumers pay more for a product they assume to be used by a celebrity but are disgusted by a product which they assume to be used by a person they dislike (Nemeroff & Rozin, 1994). However, it is not only a specific positive attribute that may shape consumer perceptions, but it is also the mere visibility of socialness, e.g., implying warmth, that is associated with humans in general (Job et al., 2017). Hence, consumers may associate products that are linked to even generic humans with positive attributes (Job et al., 2017).

In the context of physical products, past marketing research indicates a positive effect of 'handmade' labels on consumers' willingness-to-pay, compared to products lacking indications of human work (Fuchs et al., 2015; Kruger et al., 2004). It suggests that seeing a product made by a human triggers perceptions of high production efforts, increased product quality, and satisfaction of unique needs (Fuchs et al., 2015; Kruger et al., 2004). In contrast, consumers' willingness-to-pay decreases for products perceived as easy to replicate (Fuchs et al., 2015).

While previous research in digital contexts finds that seeing human interaction with a data-based product advances common sensemaking and increases trust (Ehsan et al., 2021; Haque et al., 2023; Shin, 2021b), we argue that highlighting human involvement particularly related to the sensor-based processes, like sensor data collection or analysis, may further enhance reader appreciation. Hence, we hypothesize:

H1: Readers show higher levels of appreciation for sensor-based articles when human involvement is visible within explanations for sensor-based processes (as opposed to no visible human involvement).

Early IS research shows that seeing human faces in website interfaces impacts emotions and drives consumption (Cyr et al., 2009). Visual information and particularly images of human faces have a stronger effect on the perception of human involvement and hence emotions, trust, perceived quality, and adoption behavior than text-based explanations (Buell et al., 2016; Cyr et al., 2009; Haque et al., 2023). Hence, we hypothesize:

H2: Readers show higher levels of appreciation for sensor-based articles when human involvement is visually depicted through images (as opposed to being described via text).

5. Proposed Research Model

In our research model (Figure 1), we consider two independent variables. The first independent variable is the visibility of human involvement in sensor-based processes, varying according to its presence or absence. The second independent variable is the format, varying according to image-based or text-based explanations conveying human involvement. For the dependent variable, we consider the level of reader appreciation for sensor-based articles. As payment intents generally tend to decrease for journalistic content, we consider reader appreciation as a suitable dependent variable over, for instance, willingness-to-pay. Reader appreciation is associated more strongly with reflective gratification and the sensation of gaining insights than with responses to hedonic media, such as enjoyment or fun (Oliver & Bartsch, 2011). We further consider control variables, such as age, gender, average income, country of residence, trust in (sensor-based) journalism, consumption patterns of (sensor-based) journalism, and general appreciation for journalism.

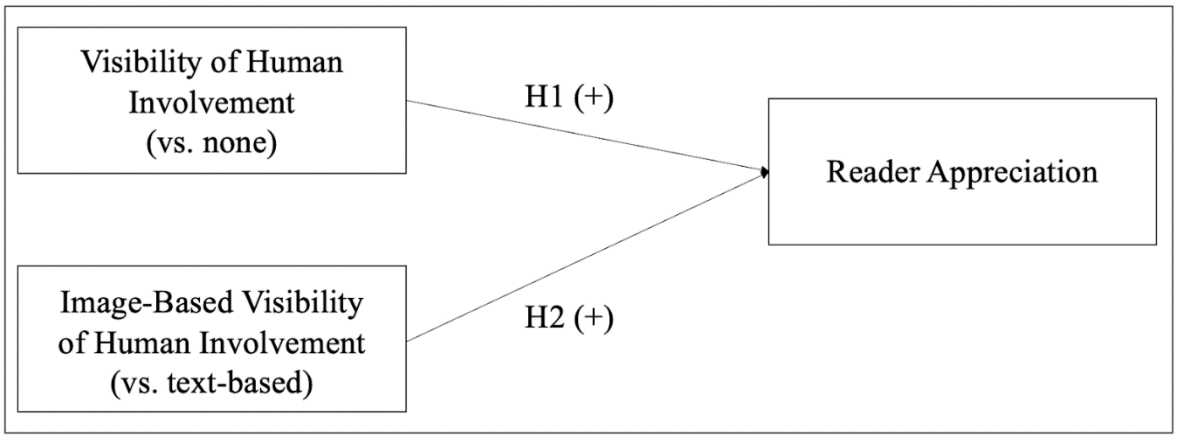


Figure 1: Proposed Research Model

6. Forthcoming Research Steps

To test our research model, we will conduct an online experiment in a controlled setting. We will recruit the participants (at least 200) through random sampling via social media and internet forums, targeting news readers, at least on an occasional basis, residing in Germany. We will randomly divide the participants into three groups, each for one of three conditions. Participants will then be exposed to the inherent condition based on their assigned group. While the first condition contains no cues on human involvement (1), the remaining conditions contain either image-based (2) or text-based (3) cues on human involvement. Every condition will contain three short sensor-based articles (each max. 200 words) on three different topics (traffic, nature, and sports) to exclude topic-specific effects. Each of those articles contains an explanation for the sensor-based processes, e.g., the collection procedure of the sensor data, the operationalizations of the variables, and the analysis methods applied to the sensor data. For the sensor-based articles, we will use realistic screenshots of slightly modified versions of articles already published in newspapers. After reading each of the three sensor-based articles, participants will be asked whether they have already seen or read a similar article. Participants who are familiar with these sensor-based articles will be excluded from the sample. Then,

participants will be asked to indicate their appreciation for each sensor-based article on a scale ranging from 1 (very low appreciation) to 10 (very high appreciation) (Oliver & Bartsch, 2011). The collected data will be analyzed according to multiple one-way analyses of variance (ANOVA) using the statistical program R.

7. Conclusion and Expected Contributions

In this research-in-progress paper, we build upon the diminished visibility of human involvement in the sensor-based processes in sensor-based journalism. Grounding our study on the literature on process transparency and social translucence, we develop a research model by deriving two hypotheses on the impact of seeing human involvement in sensor-based processes on reader appreciation for sensor-based articles. We also consider the influence of the format, varying according to image-based or text-based explanations conveying human involvement, on reader appreciation for sensor-based articles. In the forthcoming steps, we will test the model via an online experiment.

We hope to broaden IS research to journalism - an industry underexplored in IS, but pivotal in our economies and democracies (Diakopoulos, 2019). Additionally, we aim to contribute to process transparency literature (Haque et al., 2023; Rai, 2020) by extending the scope of process-related information that can be communicated to readers to influence their appreciation for sensor-based articles. Furthermore, we aim to add to social translucence literature (Erickson & Kellogg, 2000; Ehsan et al., 2021) by applying the 'handmade' labels previously explored in marketing literature (Fuchs et al.. 2015) also to sensor-based products.

References

Bardhan, I., Chen, H., and E. Karahanna (2020). "Connecting Systems, Data, and People: A Multidisciplinary Research Roadmap for Chronic Disease Management." *Management Information Systems Quarterly, 44*(1), 185-200.

Buell, R., Kim, T., and C. Tsay (2016). "Creating Reciprocal Value Through Operational Transparency." *Management Science, 63*(6), 1673-1695.

Cyr, D., Head, M., Larios, H., and B. Pan (2009). "Exploring Human Images in Website Design: A Multi-Method Approach." *Management Information Systems Quarterly, 33*(3), 539-566.

Diakopoulos, N. (2019). *Automating the News: How Algorithms Are Rewriting the Media*, Cambridge, MA: Harvard University Press.

Ehsan, U., Liao, Q., Muller, M., Riedl, M., and J. Weisz (2021). "Expanding Explainability: Towards Social Transparency in AI Systems." ACM CHI Conference on Human Factors in Computing Systems, Yokohama, Japan.

Erickson, T. and W. Kellogg (2000). "Social Translucence: An Approach to Designing Systems that Support Social Processes." *Periodical ACM Transactions on Computer Human Interaction, 7*(1), 59-83.

Fernandez-Loria, C., Provost, F., and X. Han (2022). "Explaining Data-Driven Decisions Made by AI Systems: The Counterfactual Approach." *Management Information Systems Quarterly, 46*(3), 1635-1660.

Fuchs, C., Schreier, M., and S. Van Osselaer (2015). "The Handmade Effect: What's Love Got to Do with It?" *Journal of Marketing, 79*(2), 98-110.

Gefen, D. and D.W. Straub (2004). "Consumer Trust in B2C e-Commerce and the Importance of Social Presence: Experiments in e-Products and e-Services." *Omega, 32*(6), 407-424.

Godler, Y., Reich, Z., and B. Miller (2020). "Social Epistemology as a New Paradigm for Journalism and Media Studies." *New Media & Society, 22*(2), 213-229.

Gronsund, T. and M. Aanestad (2020). "Augmenting the Algorithm: Emerging Human-in-the-Loop Work Configurations." *Journal of Strategic Information Systems, 29*(2), 1-16.

Haque, A., Islam, N., and P. Mikalef (2023). "Explainable Artificial Intelligence (XAI) from a User Perspective: A Synthesis of Prior Literature and Problematizing Avenues for Future Research." *Technological Forecasting and Social Change, 186*(2), 1-19.

Job, V., Nikitin, J., Zhang, S., Carr, P., and G. Walton (2017). "Social Traces of Generic Humans Increase the Value of Everyday Objects." *Personality and Social Psychology Bulletin, 43*(6), 785-792.

Jones, M. (2019). "What We Talk About When We Talk About (Big) Data." *Journal of Strategic Information Systems, 28*(1), 3-16.

Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*, London, UK: SAGE.

Koenig, P., Wurster, S., and M. Siewert (2022). "Consumers are Willing to Pay a Price for Explainable, but not for Green AI. Evidence from a Choice-Based Conjoint Analysis." *Big Data & Society, 9*(1), 1-13.

Kruger, J., Wirtz, D., Van Boven, L., and T. Altermatt (2004). "The Effort Heuristic." *Journal of Experimental Social Psychology, 40*(1), 91-98.

Lewis, S. and O. Westlund (2015). "Big Data and Journalism Epistemology, Expertise, Economics, and Ethics." *Digital Journalism, 3*(3), 447-466.

Loebbecke, C. and I. Boboschko (2020). "Reflecting Upon Sensor-Based Data Collection to Improve Decision Making." *Journal of Decision Systems, 29*(Sup1), 1-14.

Mahmud, H., Islam, A., Ahmed, S., and K. Smolander (2022). "What Influences Algorithmic Decision-Making? A Systematic Literature Review on Algorithm Aversion." *Technological Forecasting and Social Change, 175*(121390), 1-26.

Monteiro, E. and E. Parmiggiani (2019). "Synthetic Knowing: The Politics of the Internet of Things." *Management Information Systems Quarterly, 43*(1), 167-184.

Nemeroff, C. and P. Rozin, (1994). "The Contagion Concept in Adult Thinking in the United States: Transmission of Germs and Inter-Personal Influence." *Ethos, 22*(1), 158-186.

Newell, S. and M. Marabelli (2015). "Strategic Opportunities (and Challenges) of Algorithmic Decision-Making: A Call for Action on the Long-Term Societal Effects of "Datification"." *Journal of Strategic Information Systems, 24*(1), 3-14.

Oberlaender, A., Roeglinger, M., Rosemann, M., and A. Kees (2018). "Conceptualizing Business-to-Thing Interactions - A Sociomaterial Perspective on the Internet of Things." *European Journal of Information Systems, 27*(4), 486-502.

Oliver, M. and A. Bartsch. (2011). "Appreciation of Entertainment: The Importance of Meaningfulness via Virtue and Wisdom." *Journal of Media Psychology, 23*(1), 29-33.

Parmiggiani, E., Osterlie, T., and P. Almklov (2022). "In the Backrooms of Data Science." *Journal of the Association for Information Systems, 23*(1), 139-164.

Rai, A. (2020). "Explainable AI: From Black Box to Glass Box." *Journal of the Academy of Marketing Science, 48*(1), 137-141.

Shin, D. (2021a). "The Effects of Explainability and Causability on Perception, Trust, and Acceptance: Implications for Explainable AI." *International Journal of Human-Computer Studies, 146*(Feb), 1-10.

Shin, D. (2021b). "The Perception of Humanness in Conversational Journalism: An Algorithmic Information-Processing Perspective." *New Media & Society, 24*(12), 2680-2704.

18. The Mediating Role of Government Support in E-wallet Acceptance Intention in Malaysia
(Research-in-Progress)

Kamarul Faizal Hashim
University of Dubai, UAE
kbinhashim@ud.ac.ae

Felix B. Tan
Auckland University of Technology
felix.tan@aut.ac.nz

Jairo A. Gutierrez
Auckland University of Technology
jairo.gutierrez@aut.ac.nz

Abstract
The aim of this research is to examine the factors that predict e-wallet acceptance intention and especially the mediating role of government support that can lead to better intention to accept this technology in Malaysia. The contribution of this research is to explore the nuanced relationships that can help expand the understanding of e-wallet acceptance intention in Malaysia. Furthermore, this study is among the few that investigates the influence of government support as a mediating construct to predict e-wallet adoption intention. This research in progress also presents the study's research model and method.

Keywords*:* E-wallet acceptance, Government support, Mediation analysis

1. Introduction
Malaysia has established itself as one of the competitive financial service industries in Asia and have been rapidly transforming its financial systems (Komba & Abd Razak, 2021). The Malaysian central bank implemented strategic policies to facilitate seamless financial transactions, serving as a cornerstone for nurturing a cashless society (Chan, Leong, Lim, & Chiu Yiong, 2020). In response to the Malaysia Government's vision for digitalizing its financial infrastructure, a plethora of e-wallet applications flooded the market (Abdull Rahman et al., 2022; Chan, Leong, Lim, & Chiu Yiong, 2020). Both local and international e-wallet service providers introduced e-wallet services such as GrabPay, WeChat Pay, Touch n Go Wallet, Duit Now, Boost, Fave, Samsung Pay, BigPay, Apple Pay, and many more.

Consequently, e-wallet service providers in Malaysia have made substantial investment towards improving their applications and at the same time trying to achieve a competitive advantage over competitors (Chan et al., 2020). Companies use financial incentives, rewards, and paybacks to entice users to adopt their e-wallet services. In addition, e-wallet service providers also include gamification as part of the application to lure younger e-wallet users (Sanny et al., 2022). For e-wallet service providers, it is important to widen their user base of the application to help ensure the company's sustainability and survival (Tian et al., 2023). Despite these efforts, the adoption of e-wallet in Malaysia remains relatively low (Kiew et al., 2022; Tian et al., 2023).

Previous studies examining this topic within the Malaysia context mainly focus on identifying factors predicting e-wallet acceptance behaviour. Factors such as perceived usefulness and perceived ease of use were used extensively to predict e-wallet acceptance behaviour (Chan, Leong, Lim, & Chiu Yiong, 2020; Teo et al., 2020). In addition, other factors like perceived trust, security, privacy and service quality also received equal attention as predictors of this behaviour (Karim et al., 2020; Teng Tenk Teoh et al., 2020; Teo et al., 2020).

To extend the understanding of e-wallet acceptance in Malaysia, this study examines the role government support (GS) plays in influencing this behaviour. GS refers to support from government agencies that help open new opportunities for users and businesses to take part in using or offering the use of a new technology (Hussain et al. 2021). Examples of GS can be in the form of financial support, provision of infrastructure, upgrading human resource skills, advisory services and many others. This support can help roll out new technology to the wider public more effectively (Yusoff et al., 2021). GS was also identified as an important factor that can help increase the acceptance of new technology especially in developing countries (Hussain et., 2021; Yusoff et al., 2021).

In this study, GS is conceptualised as a mediating factor that mediates the relationship between determinants of e-wallet acceptance and one's acceptance intention of the technology. The mediating role of GS has been examined in prior research; for instance, in user's acceptance of e-commerce (Yusoff et al., 2021), electronic banking (Hussain et al., 2021) and online shopping (Hai & Alam Kazmi, 2017). However few studies have examined the mediating role of GS on e-wallet acceptance behaviour especially in Malaysia. Furthermore, by exploring this relationship, the researchers hope to enrich the existing understanding of e-wallet acceptance by investigating the nuanced interconnections between determinant factors, enabling a better explanation of e-wallet acceptance behaviour (Tian et al., 2023).

2. Literature Review

2.1. E-wallet Adoption in Malaysia

An e-wallet refers to a mobile application that enables users to conduct financial activities by linking the application to their respective bank (Tian et al., 2023). In Malaysia, e-wallets are offered not only by banks but also third-party financial institutions. In Malaysia the availability of e-wallet platforms is growing steadily but the adoption level is still at infancy level.

As a consequence of the low rate of adoption of e-wallet platforms in Malaysia, the topic of e-wallet adoption in Malaysia has received substantial attention by Malaysian researchers. Table 1 lists previous studies conducted in Malaysia. Generally, majority of previous studies used factors derived from technology acceptance theories to predict e-wallet acceptance behaviour.

Author	Focus of Study	Mediating/Moderating variable	Respondents
Tian et al., (2023)	Direct influence of PUSE, PEOU, attitude, subjective norms & perceived behavioural control on e-wallet adoption intention.	Examine the moderating influence of perceived trust & service quality on e-wallet adoption behaviour	378 respondents using e-wallets in Malaysia
Abdul Halim et al., (2021)	Influence of price benefit, trust, habit, satisfaction, operational constraints, satisfaction, PUSE, PEOU, attitude & confirmation on e-wallet continuous adoption intention.	None	379 respondents using e-wallets in Malaysia
Kiew et al., (2022)	Influence of perceived trust, convenience & promotion on e-wallet adoption intention.	None	249 respondents using e-wallets in Sarawak, Malaysia.
Chan, Leong, Lim, & Yiong, (2020)	Influence of mobile PUSE, PEOU, perceived compatibility, risk, security & trust on e-wallet adoption intention.	None	233 respondents using mobile payment apps in Malaysia
Teoh et al., (2020)	Influence of performance expectation, effort expectation, social influence, perceived risk & cost affecting e-wallet adoption intention.	None	210 respondents using e-wallets in Malaysia
Karim et al., (2020)	Direct influence of PUSE, PEOU, privacy & security on e-wallet adoption intention.	None	289 youths using e-wallets in Malaysia
Teo et al., (2020)	Perceived security, PUSE, PEOU & social influence on e-wallet adoption intention.	None	200 Malaysian using e-wallets in Malaysia
Chiew Yi & Perera, (2019)	Trust, PUSE & security influence on e-wallet adoption intention.	Examine the mediating influence of PEOU on e-wallet adoption behaviour	140 university students in a private university using e-wallets in Malaysia

Table 1: Summary of Previous Studies on E-wallet in Malaysia

Based on the literature analysis, it can be concluded that:

1. Most studies used technology related factors from adoption theories to determine e-wallet acceptance behaviour. Technology adoption behaviour evolves over time and focusing on these factors at the early stage of the technology adoption phase can be relevant. Over time, a user might have different expectations that might affect their acceptance behaviour (Hashim & Tan, 2018). Thus, focusing on non-technology related factors may provide additional perspectives and as such a better understanding of e-wallet acceptance.
2. Very few studies have tried to extend the conceptualization beyond examining direct relationships between determinant factors and the acceptance behaviour. For example, Tian et al., (2023) examined how perceived trust and service quality can moderate the relationship between a user's intention and actual behaviour to adopt e-wallet. Chiew Yi & Perera (2019) predicted that perceived ease of use (PEOU) mediated the relationship between trust, perceived usefulness (PUSE) and security on intention to adopt e-wallet. Hence, to expand the understanding of e-wallet acceptance behaviour in Malaysia, more

research is needed to extend the applicability of identified factors beyond the existing conceptualization. Further research can help to explore nuanced interconnections between determinant factors, enabling a better explanation of e-wallet acceptance behaviour (Tian et al., 2023).

Therefore, for this study three constructs are selected to predict e-wallet acceptance intention. The factors are perceived convenience, compatibility, and security. These constructs were selected due to their relevance and remain a significant concern among e-wallet users in Malaysia (Andrew et al., 2021). Meanwhile, GS is selected as the mediating variable to help further the understanding of e-wallet acceptance in Malaysia. GS was selected as a factor because in Malaysia, there is a general expectation that the government will always play an important role in supporting and facilitating the introduction of new technology (Hussain et al., 2021; Yussof et al., 2021). According to Husain et al. (2021), government support plays a linking role between the users' perception on the technology and their acceptance behaviour.

2. Government Support as a Mediating Variable

According to Yussof et al. (2021), GS acts as a catalyst that help users and businesses to venture into using new technology. GS also lends credibility and feasibility leading to the acceptance of new technology by potential users (Marakarkandy et al., 2017). For example, in 2018 the government of Malaysia gave MYR30 to every Malaysian aged 16 years and above to promote the initial uptake of e-wallet. Meanwhile, in Pakistan, the government helped to promote the use of electronic banking by formulating relevant policies that allow banks to develop strategies based on users' preferences and demands (Hussain et al., 2021).

The examination of the role of GS as a mediating variable has been found in previous studies that examine technology adoption. Hussain et al. (2021) examined the mediating role of GS on users' e-banking adoption in Pakistan. According to Hussain et al. (2021), users who find the performance of the e-banking system as beneficial and effective will be more likely to adopt the technology with the present of appropriate GS such as having easy to access internet or other infrastructure.

However, the impact of GS as a mediating factor within e-wallet adoption remains largely unexplored, particularly within the Malaysia context. In a developing country like Malaysia, GS is very important especially when introducing a new technology to the public (Hussain et al., 2021; Yussof et al., 2021). Users are considerably more confident to adopt innovative technologies when they are assured of the presence of appropriate laws and regulations to safeguard their interests. Hence based on these arguments, this study aims to explore the mediating role of GS impacting e-wallet acceptance intentions in Malaysia.

3. Research Model

Figure 1 illustrates the proposed research model for this study. In the model, e-wallet acceptance intention (EW_INT) is positively influenced by perceived convenience (PER_CON), perceived security (PER_SEC) and perceived compatibility (PER_COM). Meanwhile, government support (GS) is predicted to mediate the relationship between

PER_CON, PER_SEC, PER_COM and EW_INT. The following section discuss the hypotheses for this study.

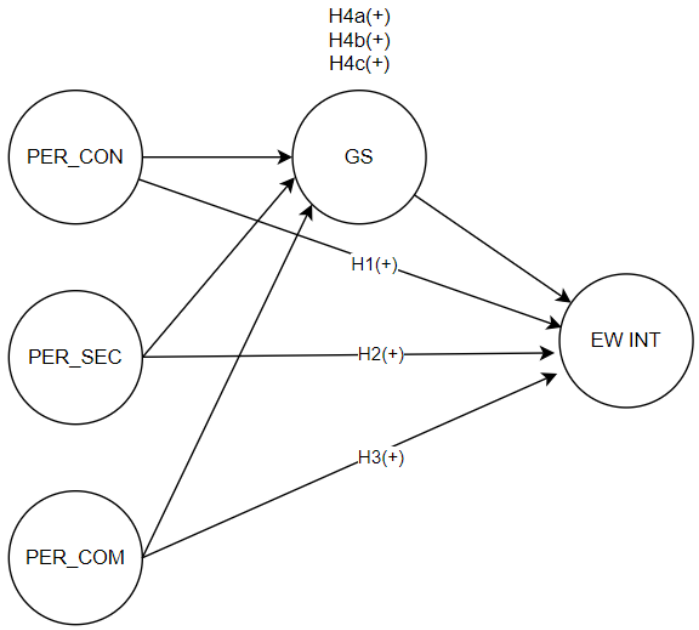


Figure 1: Research Model

4. Research Hypothesis

Based on the research model proposed in the previous section, this study hypothesized the following hypotheses:

Hypothesis 1: Perceived convenience positively influences e-wallet acceptance intention.

Hypothesis 2: Perceived security positively influences e-wallet acceptance intention.

Hypothesis 3: Perceived compatibility positively influences e-wallet acceptance intention.

Hypothesis 4a: Government support mediates the relationship between perceived convenience and e-wallet acceptance intention.

Hypothesis 4b: Government support mediates the relationship between perceived security and e-wallet acceptance intention.

Hypothesis 4c: Government support mediates the relationship between perceived compatibility and e-wallet acceptance intention.

5. Research Methodology

This study will administer a web survey to collect the data from respondents. The respondents of this study are Malaysians who have experience using any e-wallet services in Malaysia. This study will adopt a purposive sampling technique. The link for this survey will be promoted in the researchers' social media platforms such as WhatsApp, Facebook and Twitter. The

collected data will then be analysed using the partial least square - structural equation modelling (PLS-SEM) technique.

6. Conclusion

This research in currently at the data collection stage. By completing the data collection and analyses phases, this study aims to predict 1) the direct influence of perceived convenience, perceived security and perceived compatibility; and 2) the mediating influence of GS on user's intention to accept e-wallets in Malaysia.

References

Abdul Halim, N. A., Zadeh, A. V., Hanifah, H., Teoh, A. P., & Nawaser, K. (2021). Understanding the Determinants of e-wallet Continuance Usage Intention in Malaysia. *Quality & Quantity*.

Abdull Rahman, N. L., Abd Mutalib, H., Md Sabri, S., Annuar, N., Abdul Mutalib, S. K. M. S., & Abdull Rahman, Z. S. (2022). Factors Influencing E-Wallet Adoption among Adults During Covid-19 Pandemic in Malaysia: Extending the TAM Model. *International Journal of Academic Research in Business and Social Sciences*, *12*(7). https://doi.org/10.6007/ijarbss/v12-i7/14327

Andrew, J.V., Ambad, S.A., Wong Abdullah, N.S., Nordin, S., & Tan, K.E. (2021). A Systematic Review of E-wallet Usage Intention: Integrating UTAUT2 with Perceived Security. *Journal Intelek, 16*(1), 124-134.

Chan, K.-L., Leong, C.-M., Lim, B., & Chiu Yiong, B. L. (2020). Sharing Economy through E-Wallet: Understanding the Determinants of User Intention in Malaysia. *Journal of Marketing Advances and Practices, 2*(2), 1–18.

Chan, K.-L., Leong, C.-M., Lim, B., & Yiong, C. (2020). Sharing Economy through E-Wallet: Understanding the Determinants of User Intention in Malaysia. *Journal of Marketing Advances and Practices, 2*(2), 1–18.

Chiew Yi, P. K., & Perera, S. (2019). Addressing E-wallet Adoption among Malaysians. *Taylor's Business Review, 8*(2), 41–57.

Hai, L.C., & Alam Kazmi, S.H. (2017). Dynamic Support of Government in Online Shopping. *Asian Social Science, 14*(22), 1-9.

Hashim, K.F., & Tan, F.B. (2018). Examining the Determinant Factors of Perceived Online Community Usefulness using the Expectancy Value Model. *Journal of Systems and Information Technology, 20*(2), 152-167.

Hussain, A., Hussain, M.S., Khan Mari, M.Y., & Zafar, A. (2021). Acceptance of Electronic Banking among University Students in Pakistan: An Application of Technology Acceptance Model. *Pakistan Journal of Humanities and Social Sciences, 9*(2), 101-113.

Karim, M. W., Haque, A., Ulfy, M. A., Hossain, M. A., & Anis, M. Z. (2020). Factors Influencing the Use of E-wallet as a Payment Method among Malaysian Young Adults. *Journal of International Business and Management, 01*(12), 1–11. https://doi.org/10.37227/jibm-2020-2-21

Kiew, J. P., Toh, E. T. L., Ngian, E. T., & Law, S. H. C. (2022). Perceived Trust, Convenience and Promotion for the Adoption of E-wallet. *International Journal of Academic Research in Business and Social Sciences*, *12*(9). https://doi.org/10.6007/ijarbss/v12-i9/14591

Komba, K. J., & Abd Razak, K. (2021). Factors Influencing Customers Retention for Electronic Wallet Services in Malaysia. *International Journal of Social Science and Humanity*, *11*(2), 44–48.

Marakarkandy, B., Yajnik, M., & Dasgupta, M. (2017). Enabling Internet Banking Adoption: An Empirical Examination with an Augmented Technology Acceptance Model (TAM). *Journal of Enterprise Information Management, 20*(2), 263-294.

Sanny, L., Iskandar, I., & Yus Kelana, B. W. (2022). Gamification Reward in Indonesia's e-wallet Market. *2022 10th International Conference on Cyber and IT Service Management, CITSM 2022*. https://doi.org/10.1109/CITSM56380.2022.9935942

Teoh, M. T. T., Chin Yew, H., & Teck Heang, L. (2020). E-wallet Adoption: A Case in Malaysia. *International Journal of Research in Commerce and Management Studies*, *2*(02), 1–18. http://ijrcms.com

Teo, S. C., Law, P. L., & Koo, A. C. (2020). Factors Affecting Adoption of E-Wallets Among Youths in Malaysia. *Journal of Information System and Technology Management*, *5*(19), 39–50. https://doi.org/10.35631/jistm.519004

Tian, Y., Chan, T. J., Suki, N. M., & Kasim, M. A. (2023). Moderating Role of Perceived Trust and Perceived Service Quality on Consumers' Use Behavior of Alipay e-wallet System: The Perspectives of Technology Acceptance Model and Theory of Planned Behavior. *Human Behavior and Emerging Technologies*, *2023*. https://doi.org/10.1155/2023/5276406

Quin Seng, W., & Choon Hee, O. (2021). Factors Influencing the Intention to Use E-Wallet: An Extended Hedonic-Motivation System Adoption Model. *International Journal of Academic Research in Business and Social Sciences*, *11*(11), 2265–2278. https://doi.org/10.6007/ijarbss/v11-i11/11379

Yusoff, M. N. H., Zainol, F. A., Abdul Rahim Merican, R., Hassan, H., & Afthanorhan, A. (2021). Modeling the Micro-Entrepreneurs' E-commerce Acceptance: The Role of Government Support Programme. *Journal of Positive School Psychology*, *2022*(3), 1910–1925. http://journalppw.com

19. The Role of SOCs for Critical Infrastructure Providers

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

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

Abstract

Cyber security is a challenge for all organizations and companies, especially for those classified as critical infrastructure providers, such as hospitals and energy providers. Attempts to prevent or fight against cyberattacks are various., Therefore, critical infrastructure providers remain searching for an institutionalized solution. Security operation centers (SOCs) seem to offer an opportunity to tackle this issue, but to date, it remains unclear, what influenced the implementation of a SOC. Directly bound to cyber security and business continuity management, the strategic role of SOCs has been addressed, previously. The role that SOCs play for critical infrastructure providers still needs to be clarified. Our case study shows, that decisions for implementing a SOC are influenced by different perspectives, i.e., strategic, socio-technical, and solution-implementation perspectives. All three perspectives share factors with existing cyber security research but also give new insights, particularly regarding the social aspects of SOC staff and the fact that businesses are in favor of an outsourced SOC to benefit from shared resources, reduced costs and skilled staff.

Keywords: Cyber Security, Security Operation Center, Critical Infrastructure, Case Study Research.

1. Introduction

Cyber threats and cyber security have become well-discussed topics in academia and business since the digital transformation forced most organizations – public or private - to adopt digital communication and participate in cyberspace. Besides the benefits evolving from the Internet, organizations need to secure their information systems and protect the data they collect and store. In particular, organizations classified as critical infrastructure providers (CRITIS providers) are obliged to implement security in a structured and holistic way (Bederna et al., 2020; Napolitano et al., 2018). CRITIS providers are characterized by their meaning for society, like hospitals, and emergency organizations but also energy and telecommunication providers. Hence, their failure would lead to massive negative effects on society (Mikhalevich & Trapeznikov, 2019). The importance of CRITIS providers for society makes them an attractive target for any kind of aggressors (Bederna et al., 2020). For example, the infiltration of the uranium enrichment plant in Natanz with Stuxnet (Chen & Abu-Nimeh, 2011; Langner, 2011), or the cyberattack of the Ukrainian power grid which caused a blackout (Liang et al., 2017; Whitehead et al., 2017), to name just two examples. Therefore, it is indispensable to ensure that critical infrastructures and their providers can deliver a constant quality of service. To guarantee this service, governments, governmental organizations and political institutions such as the European Union created some management frameworks and regulations to safeguard the operation of CRITIS providers (Krumay et al., 2018; Materese, 2016; NIST,

2023). However, these approaches mainly define issue areas to protect but do not clearly state, how the measures are managed within the company (Krumay et al., 2018). Since the issues are manifold (Krumay et al., 2022), critical infrastructure providers are looking for a structured and stable way. Security Operation Centers (SOCs) have been named among the possibilities to achieve cyber security goals. A SOC is a business service or internal business unit that aims to deal with cyber-attacks, fulfilling different tasks (Kelley & Moritz, 2006; Vielberth et al., 2020). The literature on this topic mainly focuses on technical and organizational issues (Jacobs et al., 2013; Schinagl et al., 2015a), but lacks a clear understanding of the role of a SOC and its implementation. Since CRITIS providers need to guarantee a certain level of reliability and resilience, SOCs may contribute to achieving this goal (Martin & Ludek, 2012). This study addresses the area of influencing factors for a CRITIS provider to decide whether to implement a SOC. We aim to close this gap by investigating the implementation of a SOC as an institutionalization of cyber security. Therefore, we ask: how do critical infrastructure providers perceive the role of SOCs? We adopt a case study approach in collaboration with three different central European CRITIS providers from the communication sector. We collected data using six interviews, combined it with additional sources (i.e., policies from the case companies and internal training material) and applied qualitative content analysis, particularly coding techniques to deduct the relevant information from these sources to build higher-level categories. The remainder of this paper is structured as follows. First, we provide an overview of the topics involved in this research. Next, we describe the methodological approach, followed by the results. Finally, we discuss the results and provide a conclusion, limitations, and future research possibilities.

2. Background Information

2.1. Critical Infrastructure

The term critical infrastructure has gained a lot of attention lately – not only due to the Covid-19 pandemic (Scholz et al., 2022) but also due to scenarios like blackouts that have been discussed in the media (Liang et al., 2017; Shull et al., 2016). Although there are different definitions, in general, critical infrastructure is defined by its essential contribution to society (Bederna et al., 2020; European Commission, 2024; Mikhalevich & Trapeznikov, 2019). Many governments agree to 16 sectors fulfilling this characteristic: chemical commercial facilities, communications, critical manufacturing, dams, defence industrial base, emergency services, energy, financial services, food and agriculture, government facilities, healthcare and public health, information technology, nuclear reactors (incl. materials and waste), transportation systems, water and wastewater (CISA, 2024). Although most resources address critical infrastructures from a general point of view, some authors differentiate between physical and information infrastructure (Mikhalevich & Trapeznikov, 2019).

The main goals of existing frameworks and regulations focus on the reliability and resilience of critical infrastructures and their providers (Martin & Ludek, 2012; Neumannova et al., 2023). Reliability is considering single components in the infrastructure and their probability of working properly (Murray & Grubesic, 2007). The reciprocal concept is vulnerability, i.e. the probability of the failure of an element in the system (Murray & Grubesic, 2007). Reliability needs also to be granted in times of unforeseen events or threats (Murray & Grubesic, 2007).

The higher the reliability, the better the resilience of the infrastructure (Martin & Ludek, 2012). Resilience in general is often referred to as the ability to "bounce back" (Burnard & Bhamra, 2011; Garmezy & Masten, 1986; Heeks & Opina, 2019; Holling, 1973). Resilient organizations are faster in coming back after a disaster (Rothrock, 2018) and they even may find a way to learn from the situation and "bounce forward" (Heeks & Opina, 2019). In the context of critical infrastructures, resilience is more referred to the ability to maintain the service even in times of crisis (Cernan et al., 2020; Martin & Ludek, 2012). To achieve this ability, organizations in general and CRITIS providers in particular need to establish appropriate measures (Boin & McConnell, 2007; Cernan et al., 2020; Martin & Ludek, 2012). From the organizational point of view, these measures include approaches like business continuity management, crisis management and cyber security management (Boin & McConnell, 2007; Hecht, 2002; Krumay et al., 2022). Business continuity management (BCM) focuses on the time-to-recover after a disaster, but also the responsibilities, prioritization and resources used (Burnard & Bhamra, 2011; Tompkins, 2007). Crisis management, by contrast, mainly aims at safeguarding preparedness as a precondition for resilience (Boin & McConnell, 2007). Both – BCM and crisis management – are organization-wide schemes that address all disturbances on the physical and digital level (Smith, 2003). The different disturbances are classified based on their impact on the organization (Smith, 2003). It is important to mention, that BCM and crisis management are not solely related to cyber security, but are implemented for all disturbances (Burnard & Bhamra, 2011; Smith, 2003). The two are closely related and crisis management is seen as a precondition for BCM (Herbane et al., 2004).

2.2. Cyber Security

With the rise of the internet, cyber security has become vital due to the increasing number of attacks targeting networks and computers (Bederna et al., 2020; Humayun et al., 2020). Cybercrime such as attacks from cyberspace, challenges all organizations, particularly CRITIS providers since their operation is at stake when being attacked (Choras et al., 2015; Krumay et al., 2022). Cybercrime occurs on a global level and affects everyone, thus, measures against it are seen to be critical (Humayun et al., 2020). The ongoing digitalization and adoption of new technologies such as the Internet of Things (IoT) and artificial intelligence (AI) forces more and different means of prevention and protection against cybercrime. Traditional crimes such as fraud, theft, discrimination but also espionage (Banerjee et al., 2008; Promnick, 2017; Trierweiler & Krumay, 2023) using computers as tools or targets of criminal activity constitute and foster cybercrime (Humayun et al., 2020; Kraemer-Mbula et al., 2013). Besides this, attacks like ransomware, distributed denial of service (DDoS) or phishing solely exist in cyberspace (Huang et al., 2019; Wright et al., 2014). Reasons for cyberattacks are various – besides monetary motives or revenge, attackers might also have political reasons or want to harm the reputation of a company (Huang et al., 2019; Kraemer-Mbula et al., 2013). Since attacks are various, measures to fight them need to be as multifaced as the crime, making cybersecurity a complex and challenging task on various levels (Aoyama et al., 2017). Therefore, different management frameworks have been developed to support companies in their efforts to set up cyber security from a holistic point of view (ISO/IEC 27001, 2013; NIST, 2023). Although cyber security is often discussed from a more technological point of view, the

human factor is important (Abroshan et al., 2021; Bauer et al., 2017). Directly related to cybersecurity is the assessment of risks to decide upon the appropriate measures (Harkins, 2013). Risk is mainly seen as the ratio between threats including vulnerabilities, likeliness to be affected and asset value (Harkins, 2013), also referred to as probabilistic risk analysis (Napolitano et al., 2018). As a result, companies may develop a risk matrix to identify the highest risks and decide on investments to prevent negative consequences (Katsumata et al., 2010; Lee, 2021). Furthermore, CRITIS providers are obliged to follow specific regulations to ensure high-level cyber security (Bederna et al., 2020; Krumay et al., 2018, 2022).

2.3. Security Operation Centers (SOCs)

Due to the complexity of cyber security management, the implementation of stable structures in companies has already been discussed intensively (Aoyama et al., 2017; Krumay et al., 2018; Lee, 2021). SOCs are such a stable and structured approach, because they are operated 24/7 and can be seen as an institutionalization of cyber security measures. They fulfill at least three different but connected tasks: logging (incl. monitoring), vulnerability detection and response to incidents (Kelley & Moritz, 2006; Onwubiko, 2015). SOCs depend on a stable budget, thus they need to be integrated into the company's strategy (Mikhalevich & Trapeznikov, 2019). SOCs have both – technological and social components, as they can be seen as a facility (e.g., in a room) but also as a team of specialists to achieve security goals such as availability, integrity and confidentiality in the company (Onwubiko, 2015; Wall & Rodrick, 2021). Every SOC requires staff with specific skills and knowledge, that are hard to find on the job market (Wall & Rodrick, 2021). A SOC provides services for cybersecurity concerns in a company, mainly built as a business or service unit (Onwubiko, 2015). However, this requires some structural changes, staff with specific expertise as well as huge investments (Hale & Brusil, 2007; Schinagl et al., 2015b). Some account for the analysis of network traffic to identify irregularities as the main task for SOCs (Jacobs et al., 2013). Due to the constantly growing amount of data traffic, SOCs rely on machine-learning techniques to identify irregularities and avoid false positives (Karim Ganame et al., 2008). Other technologies, such as virtual reality, have also been used to support the tasks of SOC staff (Munsinger et al., 2023). Yet, there are some characteristics, that differentiate SOCs from other, unstructured approaches to achieving information security goals. First, it is a team of cyber security analysts who try to ensure the functional, security and safety-related integrity of a company's IT infrastructure. Second, SOCs have some specific tasks and properties, including collecting and analyzing data as well as responding with appropriate countermeasures (Onwubiko, 2015). There are two main ways of implementing SOCs – as an internal business unit or as an outsourced service. Especially when SOCs are seen as a business unit, responsibilities and tasks have to be assigned, often in a hierarchical structure. A common hierarchical structure consists of four tiers, represented by requirements for employees' skills (Vielberth et al., 2020). Tier 1 (triage specialist) is characterized by analysts focusing on irregularities in the data traffic, partly based on machine-learning techniques. Tier 2 analysts responsible for ongoing security attacks and reporting on them after the attacks to document countermeasures and weaknesses. Tier 3 analysts (threat hunters) identify weaknesses within the company's infrastructure, focusing on already existing threats. In tier 4, the leadership of the SOC is established, i.e. SOC managers, responsible for

the strategic direction of the SOC as well as the functionality provided by the SOC (Vielberth et al., 2020). The maturity of SOCs represents the security level and offers provided by it (Jacobs et al., 2013). To build and maintain or even expand the maturity level of a SOC is even more resource-intensive (Warda, 2000). Thus, outsourced SOCs have become widely accepted and involve specific providers of managed SOCs (Schinagl et al., 2015a). SOC providers set up SOCs, including infrastructure and staff, based on the requirements of their customers, also defining the level of outsourcing (Dimitrov & Syarova, 2019; Nassar et al., 2013; Vielberth et al., 2020). Lately, SOCs have also been offered as a service (SOCaaS), especially in the context of cloud computing (Alruwaili & Gulliver, 2014). Another lately discussed possibility is a virtual SOC based on software and platform solutions (Falk et al., 2017). Although SOCs are often seen as a possibility to overcome many problems evolving with information security management (Jacobs et al., 2013), their full implementation in companies seems rare.

3. Methodological Approach

The current study is based on a case study approach (Yin, 2013) to answer our research questions, i.e., identify "how do CRITIS providers perceive the role of SOCs?". As suggested for case study research, we adopted an iterative approach (Yin, 2013), running through two iterations. For the analysis, we applied coding techniques to develop higher-level categories (Charmaz, 2014; Corbin & Strauss, 1990) on all sources. Thus, we identified basic codes which we further condensed to higher-level constructs and categories (Elo et al., 2014; Neuendorf, 2017).

This case study involved three CRITIS providers from the telecommunication sector in central Europe, further referred to as case companies, to focus on this specific sector. Two of them are private companies, whereas one is partly state-owned. They have been on the market for more than 20 years and employ between 250 and 2.000 people. Among all CRITIS providers in this sector, the case companies are seen as forerunners and innovation drivers when it comes to cyber security measures. They have already implemented BCM and crisis management as well as cyber security management on a high level. Security frameworks, guidelines, procedures and policies are well documented and disseminated in these companies. In addition, information campaigns are conducted regularly (targeting all employees) and specific onboarding processes are set up for newly hired employees. For this case study, different sources related to BCM and security management were used to answer the research question. The data was collected by 2 different researchers over six months and finished before the crisis in Ukraine started, i.e., before February 2022. In this period, unstructured interviews with six experts from three case companies were conducted based on an interview guideline we developed from the literature. The interview guideline consisted of five general questions (1 - How do you define a SOC? 2 - How would you consider a SOC implementation possible in your organization? 3 - What are the benefits of a SOC? 4 - What are the preconditions for implementing a SOC? 5 - How should a SOC for your organization look like?). The interviewees were selected based on their experience and role in the case organizations. They are involved in the development and operations of BCM (2), cyber security management (2) and information management (1), all five with leadership responsibility. In addition, one interviewee is an external security consultant, who is permanently involved in the security

decisions of one of the CRITIS providers. Two of the interviewees are female and all six were working for or with the respective company for more than five years, with experience in their field for at least 8 years. Three of them hold an internationally renowned security certificate or a university degree in security management. Half of the interviews were conducted on-site, the other half via a video communication tool (i.e., Zoom). On average, the duration of the interviews was 48 minutes. They were recorded and transcribed for analysis and category building. In addition, existing documents related to the research topic from all three case companies were added. This included security policies (3), BCM guidelines (2), emergency plans (3), and security handbooks (3). These documents were analyzed in their most current format (issued no longer than 6 months before the analysis). Overall, the documents were about 300 pages long. We integrated into our analysis additional internal training material, like general security handbooks (7) and posters (15), used in the companies to raise security awareness among all employees. Furthermore, documents describing strategic considerations for the next three years from all three case companies were analyzed (one per company). The analysis of the documents as well as the interviews was based on qualitative content analysis coding techniques (Elo et al., 2014; Neuendorf, 2017), involving two experienced researchers. In this process, we first developed basic codes from all sources (e.g., notion of "control", "clear responsibilities", "hierarchy and reporting" from the interviews and the notion of a responsible person and the hierarchical information chain in case on an incident from the documents), which were subsumed under the term "governance" and further aggregated to the category "Governance structure". We ran through this process in two iterations to assure reliability of the developed categories. It is important to mention that due to the criticality of the subject matter and to assure confidentiality, no names, exact data, or other identifiable information (e.g., wordings from the documents) can be provided in this manuscript.

4. Results

Based on the above-mentioned sources, three different perspectives have been identified as the higher-level categories. First, the strategic perspective, i.e., strategic aspects of a SOC implementation and the influence on other CRITIS providers. Second, the socio-technical perspective considers the effects of a SOC implementation on the employees and the technology used in the company. Third, the solution-implementation perspective considers the current situation of the case company for the implementation of a SOC and suggests specific avenues. A summary of these three perspectives is provided in Table 1 and further described in the next sections. From the interviews, it has become evident that there is a need to institutionalize cyber security measures, for example, a SOC.

4.1. Strategic Perspective

From the strategic perspective, the study revealed some aspects that have a long-term influence on the company. This was mainly derived from answers to questions 1, 2 and 4. All interviewees described a SOC implementation as a strategic decision. Thus, the implementation of a SOC cannot be successful without high-level support from the C-level management (e.g., CEO, CIO, CFO) and direct integration into the strategy of the company. As the strategy is influenced by the specific context as a CRITIS provider, strategic management needs to

consider all official regulations and guidelines constantly (stated in the strategy documents of all case companies).

Perspectives	Aspects of influencing factors when implementing a SOC	
Strategic	Company	SOC as strategic decision Strategic integration and alignment C-level support Specific budget Relationship with BCM, crisis management and cybersecurity management Governance structure Risk mitigation
	CRITIS providers	Role model for others Clarifying benefits Obligation in regulations to implement SOCs not recommended
Socio-technical	Technological preconditions	Facility set-up (location, access control, technical equipment) Powerful infrastructure Monitoring and logging Guidelines and policies (semi-automated) Staff with proper skills Resource shortage
	People	Team spirit Good working conditions Reputation of SOC staff Incentives
Solution-implementation	Outsourced SOC	Shared resources Shared responsibilities Skills and knowledge updates SOCaaS

Table 1: Results – Perspectives and Aspects

This is well reflected in all analyzed documents and was also supported by the interviewees. It was also mentioned by all interviewees, that the SOC would need to have a specific financial budget for the long run to stabilize its function in the company. Furthermore, three of the interviewees focused on the relationship between BCM and crisis management. Two of the interviewees stressed the necessity to adapt the existing governance structures when implementing a SOC to clarify responsibilities. Although not directly addressed in the interviews, risk mitigation based on clear rules and principles was evident in the documents analyzed. Interestingly, the interviewees holding a formal cyber security degree (e.g., certification/university degree), were not only clear about strategy and governance but also about the impacts on other CRITIS providers. They assumed that the implementation of a SOC in one company would lead to a chain effect, meaning that more companies would follow their example. On the other hand, they expressed their doubts that other companies would take the investment risk before having a role model. They also mentioned that this might lead to a change in the regulations, making a SOC for CRITIS providers of a specific size obligatory. The interviewees were very careful and stressed that they find an obligation to implement a SOC the “wrong move”.

4.2. Socio-technical Perspective

What describes this perspective best is the direct relationship between people and technology. Although the interview guideline did deliberately not address technology, the interviewees

brought up the topic in answers to all questions. In particular, the technological preconditions were frequently addressed, including the set-up of a specific facility with high-level access control and appropriate technical equipment. In addition, a powerful infrastructure in terms of connectivity and computers needs to be provided. Whereas one interviewee was strongly voting for a specific room, the others were supporting the idea of building a network infrastructure for the specific working space "that can be everywhere, even virtual" (Interviewee 1). Most prominently mentioned were the facilities to fulfill the regular tasks of SOCs, i.e., monitoring and logging. All interviewees stressed the necessity to guarantee stable and reliable monitoring based on extended logfiles. They further mentioned that cyber security measures need to be tracked and integrated into a dashboard for continuous monitoring, oversight and early warnings. Two of the interviewees were bothered by the huge amount of data evolving from the log files and how to process it in real time. Furthermore, the visualization of the logging was discussed broadly, as to the interviewee's impression that "existing monitoring tools are not flexible enough" (interviewee 4). Another aspect that is more on the technological side is the integration of guidelines and policies into the monitoring system. The interviewees suggested semi-automatically integrating guidelines and policies, e.g., warning levels and crisis procedures to enssure immediate action. The documents (in particular BCM and crisis management policies) already document these procedures and may serve as a basis for such an implementation. On the other hand, the interviewees indicated that people are the most important factor for successful SOC operations. In all interviews, the need for staff with proper skills was intensively discussed. The interviewees particularly expressed their concerns on "how to find the right staff with according knowledge" (Interviewee 6). All were clear about the current situation in which a resource shortage – mainly staffing but also in terms of budget – hinders a successful implementation of staff. Further, concerns were raised targeting the continuous improvement and training of SOC staff, the high wages expected by very specialized people and the fear that well-trained staff will "run away" (interviewees 4 & 6) to other and better-paid jobs. This concern is in line with another issue raised: the need to establish and maintain a team spirit for making a seamless handover between shifts possible. Therefore, the interviewees suggested establishing good working conditions: "making it pleasant to work there" (Interviewee 3) (e.g. to allow for a good work-life balance), increase the reputation of SOC staff within the company (Interviewee 2) and offer a clear incentive scheme to bind them to the company.

4.3. Solution-implementation Perspective

The solution-implementation perspective focuses on the way, the companies would prefer a SOC to be implemented. All interviewees expressed a preference for the creation of an outsourced SOC. They addressed concerns related to the other perspectives (e.g., skilled staff) would, that would hinder the implementation of an internal SOC. Particularly in a crisis situation, they saw many advantages of an outsourced SOC. They mentioned that this SOC will be "independent from failures within the company" (interviewee 6) and more reliable due to the separated infrastructure. Furthermore, the scarcity of knowledgeable staff, the strategic budgeting and the requirement for constant training and skill development would also favor an outsourced SOC. They also mentioned that sharing resources, i.e. technical resources, is favorable. Three of the interviewees (1, 2, 5) expressed that an outsourced SOC would include a shared responsibility between the companies and the SOC provider. Based on an outsourcing contract, the SOC provider would also have responsibility for providing staff with appropriate

skills and a current knowledge base. They expected that SOC providers would need to guarantee up-to-date skills and knowledge of their staff. Two of the interviewees further suggested establishing a SOC as a Service (SOCaaS) approach, even in form of one big CRITIS provider offering the service over a platform to others within the same sector.

5. Discussion

This case study focuses on the role of SOCs for CRITIS providers and aspects influencing the implementation of a SOC in such a company. As the results show, the role of SOCs in this context is perceived from three perspectives: strategic, socio-technical, and solution-implementation perspective. The data collected and analyzed provided some interesting insights, that contribute to the current knowledge but also offer some implications for companies in their decision to implement a SOC, not only for those operating in the critical infrastructure sector.

Regarding the contribution to the current academic discussion on the implementation of SOCs, the results of this case study support prior studies in other contexts. This is true for more or less all three perspectives. For example, regarding the strategic perspective, the issues raised such as C-level support, strategic considerations or specific budget have already been discussed (Hale & Brusil, 2007; Mikhalevich & Trapeznikov, 2019). Interestingly, we found three other factors that have not been discussed in the literature so far. First, the idea of role model companies evokes more activities, i.e., the implementation of a SOC by other CRITIS providers. This is directly related to benefits, which need to be clarified. By contrast, instantiating an obligation, e.g., in a regulation, to implement a SOC in CRITIS providers has not been recommended, as it may lead to organizational and administrative overload. It can be assumed that role models and benefits based on success stories will be a stronger argument for implementing SOCs. From the socio-technical perspective, technological preconditions have already been discussed in the literature (Kelley & Moritz, 2006; Onwubiko, 2015). Even the problems of finding appropriate staff have already been discussed (Wall & Rodrick, 2021). However, we found that other social conditions, such as team spirit, working conditions, the reputation of SOC staff and incentives have hardly been looked at in previous research. Finally, from the solution-implementation perspective, it was interesting to learn that CRITIS providers would prefer an outsourced SOC or even SOCaaS. This owes to the fact that finding and paying skilled staff, but also fulfilling the technical requirements of SOCs is problematic. Furthermore, sharing resources and responsibilities seems to be attractive. For CRITIS providers and other organizations or companies, thinking about the implementation of a SOC, the results support the kind of make-or-buy decision to be made. Overall, SOCs seem to be a possibility for all organizations aiming at the institutionalization of cyber security measures. For companies, it is important to consider not only the technological but also the social factors influencing a successful implementation. The three perspectives in their richness of topics may guide the whole decision-making and implementation process from the strategic to the operational level towards the SOC integration in different facets.

6. Conclusion, Limitations and Further Research

This study investigated the role of SOCs for CRITIS providers and their specific needs. Based on a case study approach and integrating not only interviews with experts but also existing documents from the case companies, we identified three perspectives related to the implementation of SOCs in CRITIS providers. All three perspectives have aspects, already considered in the current discussion, but also provide new insights, such as the social aspects influencing the work of SOC staff and their willingness to do the job. But also, the willingness to share responsibilities and resources among different CRITIS providers has not been discussed intensively before. Yet, this case study does not provide a generalizable insight, as the experts and documents were limited and focused on Europe. Furthermore, the current job market situation in Europe influenced the results. Further research should therefore extend the study by e.g., investigating SOC implementation in CRITIS providers worldwide. In addition, the newly found aspects – particularly the social level – could lead to a new field of investigation, i.e., working conditions for SOC staff or to analyze the functioning of a SOC from a system point of view. This research has laid the ground for such future research paths

References

Abroshan, H., Devos, J., Poels, G., & Laermans, E. (2021). Phishing happens beyond technology: The effects of human behaviors and demographics on each step of a phishing process. IEEE Access, 9, 44928–44949.

Alruwaili, F. F., & Gulliver, T. A. (2014). SOCaaS: Security operations center as a service for cloud computing environments. International Journal of Cloud Computing and Services Science, 3(2), 87.

Aoyama, T., Nakano, T., Koshijima, I., Hashimoto, Y., & Watanabe, K. (2017). On the complexity of cybersecurity exercises proportional to preparedness. Journal of Disaster Research, 12(5), 1081–1090.

Banerjee, A., Barman, D., Faloutsos, M., & Bhuyan, L. N. (2008). Cyber-fraud is one typo away. IEEE INFOCOM 2008-The 27th Conference on Computer Communications, 1939–1947. https://ieeexplore.ieee.org/abstract/document/4509853/

Bauer, S., Bernroider, E. W. N., & Chudzikowski, K. (2017). Prevention is better than cure! Designing information security awareness programs to overcome users' non-compliance with information security policies in banks. Computers & Security, 68, 145–159. https://doi.org/10.1016/j.cose.2017.04.009

Bederna, Z., Rajnai, Z., & Szadeczky, T. (2020). Attacks against energy, water and other critical infrastructure in the EU. 2020 IEEE 3rd International Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE), 000125–000130. https://doi.org/10.1109/CANDO-EPE51100.2020.9337751

Boin, A., & McConnell, A. (2007). Preparing for Critical Infrastructure Breakdowns: The Limits of Crisis Management and the Need for Resilience. Journal of Contingencies and Crisis Management, 15(1), 50–59. https://doi.org/10.1111/j.1468-5973.2007.00504.x

Burnard, K., & Bhamra, R. (2011). Organisational resilience: Development of a conceptual framework for organisational responses. International Journal of Production Research, 49(18), 5581–5599. https://doi.org/10.1080/00207543.2011.563827

Cernan, M., Muller, Z., Tlusty, J., & Halaska, J. (2020). Critical Infrastructure and the Possibility of Increasing its Resilience in the Context of the Energy Sector. 2020 21st International Scientific Conference on Electric Power Engineering (EPE), 1–5. https://doi.org/10.1109/EPE51172.2020.9269175

Charmaz, K. (2014). Constructing grounded theory. SAGE Publications Ltd.

Chen, T. M., & Abu-Nimeh, S. (2011). Lessons from Stuxnet. Computer, 44(4), 91–93. https://doi.org/10.1109/MC.2011.115

Choras, M., Kozik, R., Torres Bruna, M. P., Yautsiukhin, A., Churchill, A., Maciejewska, I., Eguinoa, I., & Jomni, A. (2015). Comprehensive Approach to Increase Cyber Security and Resilience. 2015 10th International Conference on Availability, Reliability and Security, 686–692. https://doi.org/10.1109/ARES.2015.30

CISA. (2024). Critical Infrastructure Sectors CISA. https://www.cisa.gov/topics/critical-infrastructure-security-and-resilience/critical-infrastructure-sectors

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. Qualitative Sociology, 13(1), 3–21. https://doi.org/10.1007/BF00988593

Dimitrov, W., & Syarova, S. (2019). Analysis of the Functionalities of a Shared ICS Security Operations Center. 2019 Big Data, Knowledge and Control Systems Engineering (BdKCSE), 1–6. https://doi.org/10.1109/BdKCSE48644.2019.9010607

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative Content Analysis: A Focus on Trustworthiness. Sage Open, 4(1), 2158244014522633. https://doi.org/10.1177/2158244014522633

European Commission. (2024). Critical infrastructure—European Commission. https://home-affairs.ec.europa.eu/pages/page/critical-infrastructure_en

Falk, E., Repcek, S., Fiz, B., Hommes, S., State, R., & Sasnauskas, R. (2017). VSOC-a virtual security operating center. GLOBECOM 2017-2017 IEEE Global Communications Conference, 1–6. https://ieeexplore.ieee.org/abstract/document/8254427/

Garmezy, N., & Masten, A. S. (1986). Stress, competence, and resilience: Common frontiers for therapist and psychopathologist. Behavior Therapy, 17(5), 500–521.

Hale, J., & Brusil, P. (2007). Secur(e/ity) Management: A Continuing Uphill Climb. Journal of Network and Systems Management, 15(4), 525–553. https://doi.org/10.1007/s10922-007-9079-4

Harkins, M. (2013). Managing risk and information security: Protect to enable. Springer.

Hecht, J. A. (2002). Business continuity management. Communications of the Association for Information Systems, 8(1), 30.

Heeks, R., & Opina, A. V. (2019). Conceptualising the link between information systems and resilience: A developing country field study. Information Systems Journal, 29(1). https://onlinelibrary.wiley.com/doi/abs/10.1111/isj.12177

Herbane, B., Elliott, D., & Swartz, E. M. (2004). Business continuity management: Time for a strategic role? Long Range Planning, 37(5), 435–457.

Holling, C. S. (1973). Resilience and Stability of Ecological Systems. Annual Review of Ecology and Systematics, 4(1), 1–23. https://doi.org/10.1146/annurev.es.04.110173.000245

Huang, K., Siegel, M., & Madnick, S. (2019). Systematically Understanding the Cyber Attack Business: A Survey. ACM Computing Surveys, 51(4), 1–36. https://doi.org/10.1145/3199674

Humayun, M., Niazi, M., Jhanjhi, N., Alshayeb, M., & Mahmood, S. (2020). Cyber Security Threats and Vulnerabilities: A Systematic Mapping Study. Arabian Journal for Science and Engineering, 45(4), Article 4. https://doi.org/10.1007/s13369-019-04319-2

ISO/IEC 27001. (2013). Information technology—Security techniques—Information security management system—Requirements. ISO/IEC.

Jacobs, P., Arnab, A., & Irwin, B. (2013). Classification of Security Operation Centers. 1–7. https://doi.org/10.1109/ISSA.2013.6641054

Karim Ganame, A., Bourgeois, J., Bidou, R., & Spies, F. (2008). A global security architecture for intrusion detection on computer networks. Computers & Security, 27(1), 30–47. https://doi.org/10.1016/j.cose.2008.03.004

Katsumata, P., Hemenway, J., & Gavins, W. (2010). Cybersecurity risk management. 2010-MILCOM 2010 Military Communications Conference, 890–895. https://ieeexplore.ieee.org/abstract/document/5680181/

Kelley, D., & Moritz, R. (2006). Best Practices for Building a Security Operations Center. Information Systems Security, 14(6), 27–32. https://doi.org/10.1201/1086.1065898X/45782.14.6.20060101/91856.6

Kraemer-Mbula, E., Tang, P., & Rush, H. (2013). The cybercrime ecosystem: Online innovation in the shadows? Technological Forecasting and Social Change, 80(3), 541–555.

Krumay, B., Bernroider, E. W. N., & Walser, R. (2022). A Framework to Achieve Cybersecurity Accountability of Critical Infrastructure Providers – A Design Science Research Approach. In S. Za, A. Consorti, & F. Virili (Eds.), Organizing in a Digitized World (Vol. 50, pp. 233–248). Springer International Publishing. https://doi.org/10.1007/978-3-030-86858-1_14

Krumay, B., Bernroider, E. W., & Walser, R. (2018). Evaluation of Cybersecurity Management Controls and Metrics of Critical Infrastructures: A Literature Review Considering the NIST Cybersecurity Framework. 369–384.

Langner, R. (2011). Stuxnet: Dissecting a Cyberwarfare Weapon. IEEE Security & Privacy Magazine, 9(3), 49–51. https://doi.org/10.1109/MSP.2011.67

Lee, I. (2021). Cybersecurity: Risk management framework and investment cost analysis. Business Horizons, 64(5), 659–671.
Liang, G., Weller, S. R., Zhao, J., Luo, F., & Dong, Z. Y. (2017). The 2015 Ukraine Blackout: Implications for False Data Injection Attacks. IEEE Transactions on Power Systems, 32(4), 3317–3318. https://doi.org/10.1109/TPWRS.2016.2631891
Martin, H., & Ludek, L. (2012). Conceptual design of the resilience evaluation system of critical infrastructure elements and networks in selected areas in Czech republic. 2012 IEEE Conference on Technologies for Homeland Security (HST), 353–358. https://doi.org/10.1109/THS.2012.6459874
Materese, R. (2016, June 30). Information Technology. NIST. https://www.nist.gov/topics/information-technology
Mikhalevich, I. F., & Trapeznikov, V. A. (2019). Critical Infrastructure Security: Alignment of Views. 2019 Systems of Signals Generating and Processing in the Field of on Board Communications, 1–5. https://doi.org/10.1109/SOSG.2019.8706821
Munsinger, B., Beebe, N., & Richardson, T. (2023). Virtual reality for improving cyber situational awareness in security operations centers. Computers & Security, 132, 103368.
Murray, A. T., & Grubesic, T. H. (2007). Overview of Reliability and Vulnerability in Critical Infrastructure. In A. T. Murray & T. H. Grubesic (Eds.), Critical Infrastructure (pp. 1–8). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-68056-7_1
Napolitano, P., Rossi, G., Lombardi, M., Garzia, F., Ilariucci, M., & Forino, G. (2018). Threats Analysis and Security Analysis for Critical Infrastructures: Risk Analysis Vs. Game Theory. 2018 International Carnahan Conference on Security Technology (ICCST), 1–5. https://doi.org/10.1109/CCST.2018.8585725
Nassar, M., al Bouna, B., & Malluhi, Q. (2013). Secure Outsourcing of Network Flow Data Analysis. 2013 IEEE International Congress on Big Data, 431–432. https://doi.org/10.1109/BigData.Congress.2013.71
Neuendorf, K. A. (2017). The Content Analysis Guidebook (2nd ed.). SAGE Publications, Inc. https://doi.org/10.4135/9781071802878
Neumannova, A., Bernroider, E., & Obwegeser, N. (2023). Conceptualizing Digital Resilience: An Intellectual Capital Perspective. ICIS 2023 Proceedings. https://aisel.aisnet.org/icis2023/emergpand/emergpand/3
NIST. (2023). The NIST Cybersecurity Framework 2.0 (Draft) (NIST CSWP 29). U.S. Department of Commerce. https://doi.org/10.6028/NIST.CSWP.29.ipd
Onwubiko, C. (2015). Cyber security operations centre: Security monitoring for protecting business and supporting cyber defense strategy. 2015 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA), 1–10. https://doi.org/10.1109/CyberSA.2015.7166125
Promnick, G. (2017). Cyber economic espionage: Corporate theft and the new Patriot Act. Hastings Sci. & Tech. LJ, 9, 89.
Rothrock, R. (2018). Digital resilience: Is your company ready for the next cyber threat? Amacom.
Schinagl, S., Schoon, K., & Paans, R. (2015a). A Framework for Designing a Security Operations Centre (SOC). 2015 48th Hawaii International Conference on System Sciences, 2253–2262. https://doi.org/10.1109/HICSS.2015.270
Schinagl, S., Schoon, K., & Paans, R. (2015b). A Framework for Designing a Security Operations Centre (SOC). 2015 48th Hawaii International Conference on System Sciences, 2253–2262. https://doi.org/10.1109/HICSS.2015.270
Scholz, C., Schauer, S., & Latzenhofer, M. (2022). The emergence of new critical infrastructures. Is the COVID-19 pandemic shifting our perspective on what critical infrastructures are? International Journal of Disaster Risk Reduction, 83, 103419.
Shull, C., Dietz, J. E., & Kenley, C. R. (2016). CRITICAL INFRASTRUCTURE PROTECTION AND RECOVERY: A Common Recovery Model for Grid Systems in Blackout Scenarios. INSIGHT, 19(4), 47–49. https://doi.org/10.1002/inst.12127
Smith, D. (2003). Business continuity and crisis management. Management Quarterly, 44(1), 27–33.
Tompkins, J. A. (2007). 4 Steps to Business Resilience. Industrial Management, 49(4), 14–18.
Trierweiler, M. K., & Krumay, B. (2023). Managing Cybersecurity and Other Fraud Risks in Small and Medium Enterprises – A Framework to Build a Fraud Management Program in Times of Digitalization. Wirtschaftsinformatik 2023 Proceedings. https://aisel.aisnet.org/wi2023/28
Vielberth, M., Böhm, F., Fichtinger, I., & Pernul, G. (2020). Security Operations Center: A Systematic Study and Open Challenges. IEEE Access, 8, 227756–227779. https://doi.org/10.1109/ACCESS.2020.3045514

Wall, T., & Rodrick, J. (2021). The Demand for Cybersecurity and SOC Analysts. In T. Wall & J. Rodrick, Jump-start Your SOC Analyst Career (pp. 1–10). Apress. https://doi.org/10.1007/978-1-4842-6904-6_1

Warda, K. (2000). A Fundamental and Essential look into Managed Security Services. 14.

Whitehead, D. E., Owens, K., Gammel, D., & Smith, J. (2017). Ukraine cyber-induced power outage: Analysis and practical mitigation strategies. 2017 70th Annual Conference for Protective Relay Engineers (CPRE), 1–8. https://doi.org/10.1109/CPRE.2017.8090056

Wright, R. T., Jensen, M. L., Thatcher, J. B., Dinger, M., & Marett, K. (2014). Influence Techniques in Phishing Attacks: An Examination of Vulnerability and Resistance. Information Systems Research, 25(2), 385-400,438-439,441. ABI/INFORM Global.

Yin, R. K. (2013). Case Study Research: Design and Methods (5th ed.). Sage Publications, Inc.

20. The Short, Medium, and Long-Term Effects of ICT Use on Corruption: A Panel Data Model

Anol Bhattacherjee
University of South Florida
abhatt@usf.edu

Avijit Sengupta
University of Queensland
avijit.sengupta@uq.edu.au

Sumantra Sarkar
State University of New York, Binghamton
ssarkar@binghamton.edu

Abstract

While prior studies provide mixed evidence on whether ICT can help fight corruption, this study disaggregates the effect of ICT (specifically ICT penetration and regulation) on corruption over the short, medium, and long term to examine how this effect unfolds over time. Using a 11-year panel of secondary data (2010-2021) constructed from multiple sources (Transparency International, Global IT Report, World Economic Forum, and World Bank), we employ random-effects panel data models to examine how long it takes (in years) to realize the anti-corruption effects of ICT penetration and regulation, and how long those effects last. Our analysis shows that ICT penetration has an immediate significant effect on corruption; this effect is small between years 0 and 2 and strengthens between years 3 and 8. In contrast, ICT regulation has no significant immediate effect on corruption, a small but significant effect from year 1 through 7, and a stronger effect in year 8. Implications for these findings for corruption research and policy are discussed.

Keywords: Corruption, Information and Communication Technology, Lagged Effects, Panel analysis.

1. Introduction

The United Nations estimates that corruption costs the global economy more than 5% of global GDP or about US$3.6 trillion per year (Johnson, 2018). The United Nations Development Program defines corruption as the misuse of entrusted power for private gain (UNDP, 2011). Corruption may take many forms, such as bribery, embezzlement, money laundering, tax evasion, extortion, fraud, misappropriation, influence peddling, cronyism, and many others. It often leads to lower tax revenue, lower economic growth, lower public trust, weaker public institutions, less prosperity, and more accidents and environmental disasters, and increases business costs by as much as 20% (UNODC, 2005). Any reduction in corruption can therefore potentially improve economic efficiency, drive economic growth, build stronger institutions, free up resources for public programs, and increase quality of life for citizens.

Despite its widespread occurrence, the investigation and prevention of corruption-related crimes are made difficult by the lack of observable or usable evidence. Unlike traditional

crimes such as robbery, arson, or homicide where there are often physical and forensic evidence, evidence is often sparse in corruption investigations, since corrupt officials avoid detection by using cash transactions and by avoiding electronic communications (Gottschalk et al., 2011). Corruption investigations typically entails following unauthorized financial trails, whistleblower accounts, and wiretapped conversations, which are inadmissible in courts of law in many countries. Moreover, individuals involved in corrupt acts are often wealthy, well-connected, and influential members of society, who leverage their political connections to escape prosecution, especially in countries with weak legal systems (OECD, 2010).

Since corruption is caused by the opacity of financial and other information, it stands to reason that information and communication technologies (ICT), that enables the transparency of financial transactions to the citizenry and press, can be a key tool in combating corruption. ICT, in this context refers to a wide range of hardware, software, and networking tools, including surveillance tools, searchable databases, and data mining programs, among many others. For example, digitization of government processes and services may help search for electronic documents, create online trails of money flows for tracking or auditing suspicious transactions, enable information sharing across government agencies, cut out corrupt individuals from the loop, provide third parties access to information that was previously privy only to a select few, and bringing in much needed transparency and openness in government and business functioning (Srivastava et al., 2016; Bhattacherjee & Shrivastava, 2018).

Despite the above expectations, the empirical evidence on the relationship between ICT and corruption has remained mixed. Some studies report that ICT reduces corruption at the country level (e.g., Bertot et al., 2012; Kim et al., 2009; Lio et al., 2011), while others found no such effect (e.g., Charoensukmongkol & Moqbel, 2012; Garcia-Murillo, 2013; Heeks, 1998; Wescott, 2001). There are at least three potential explanations for such inconsistent effects. First, it may be that ICT may not have an immediate effect on corruption, but that this effect may be realized over time. If corruption studies consider the temporal effects of ICT using longitudinal or panel data, they may find evidence of corruption that may not exist in cross-sectional data. Second, the mere presence of ICT may be inadequate to fight corruption if that ICT is not used widely by the population. For instance, when people see corruption, they should be able to report it to appropriate agencies, which is not possible without high ICT penetration. Third, ICT use, in itself, may be insufficient to fight corruption; countries may also need a well-developed set of ICT laws and regulations to discourage and prosecute corruption. For example, requiring that government agencies and businesses maintain electronic records for a certain period of time, that they report and audit suspicious transactions, and that courts of law accept electronic messages and digital signatures as admissible evidence may act as deterrents to corruption.

In this paper, we explore the above explanations on the effect (or lack thereof) of ICT on corruption by examining the following two research questions:

1. What is the effect of ICT penetration and ICT regulation on global corruption?
2. How long does it take to realize the expected effects of ICT penetration and regulation on corruption and how long do those effects last?

We investigate these research questions using 11 years of archival data on corruption, ICT penetration, ICT regulation, and related constructs from 2010 to 2020 (inclusive of both years) by combining data from multiple sources: Transparency International, World Economic Forum, and the World Bank. We employ random-effects panel data models to explore the short-term, medium-term, and long-term effects of ICT penetration and regulation on country-level corruption at different lags from 0 years to 8 years. Our analysis indicates that ICT penetration reduces corruption for at least up to eight years; this effect is low in magnitude between 0 and 2 years, and higher between 3 and 8 years. ICT regulation has a non-significant effect on corruption in year 0, a low effect between 1 and 7 years, and a higher effect on year 8. Implications of these findings for research and practice are discussed.

2. Related Research

Prior studies have examined the relationship between ICT and corruption at the project and country levels of analysis in government and business settings. At the project level, Kim et al. (2009) examined electronic government data in South Korea over the 1999–2007 period and found that using e-government services increased transparency in the public sector and reduced corruption. However, in this study, ICT use was complemented with increasing public awareness of corruption, encouraging corruption detecting and reporting, firing corrupt public officials, and committing resources to curb corruption. Bhuiyan (2011) observed that e-government expansion in Bangladesh helped reduce corruption, and Muñoz-Cañavate and Hípola (2011) reported that deploying e-government services in construction projects combined with increased policing of public officials and imprisonment of several politicians lowered widespread corruption in Spanish municipalities.

Among country-level studies, Andersen (2009) evaluated data from 149 countries over the 1996–2006 decade to report that increased e-government services helped reduce corruption significantly. Similarly, using data from 70 countries over the 1998–2005 period, Lio et al. (2011) showed that Internet adoption in a country is negatively related to its level of corruption. Bertot et al. (2012) noted that corruption was higher in countries where face-to-face or phone-based interactions were preferred over e-government, and that ICT, such as e-government services, increased public transparency and reduced corruption. However, Charoensukmongkol and Moqbel (2012) observed a U-shaped relationship between ICT investment and corruption in that increasing ICT investment reduced corruption up to a certain extent, beyond which more ICT investments created new avenues for corruption. Likewise, Garcia-Murillo (2013) contended that ICT often have no impact on corruption because they are not accompanied by substantive process or role changes in corrupt systems designed to benefit the people in power.

In the corporate sector, Foucault (1977) suggested that if employees perceive their actions to be visible to others, they will tend to follow rules and procedures and be less corrupt. Elmes et al. (2005) observed that enterprise resource planning (ERP) systems and other ICT applications improve data visibility and thereby increase discipline among business employees, reduce inconsistencies in their work, and detect flow of funds that may be linked to corrupt activities. However, Heeks (1998) found that ICT does not significantly reduce corruption and can sometimes create new opportunities for corruption if employees with advanced technology

skills engage in corrupt activities by taking advantage of system vulnerabilities and unskilled employees' lack of ICT knowledge (Wescott, 2001).

From the above studies, the link between ICT and corruption appears to be inconclusive at both project and country levels of analysis and in both government and private sectors. This link appears to be the strongest in panel studies with multiple years of data and weakest or non-significant in cross-sectional studies. There is also some evidence that ICT by itself may not be sufficient to curb corruption, unless it is backed by systemic process improvements and strong legal enforcements. Furthermore, we did not find a single study that examined how long it may take to realize the expected effects of ICT penetration and regulation on corruption and how long those effects may last (our second research question). These gaps in the literature motivated our research study, our use of panel data, and our choice of ICT penetration and ICT regulation as the key predictors of interest in our corruption study.

3. Theory & Hypotheses

Traditionally, there has been three approaches at tackling corruption (Shim & Eom 2009): (1) administrative or process reforms, featuring formal rules of conduct, accountability, and responsibility (e.g., in hiring and promotion practices), quality improvement practices (often involving ICT for information intake, processing, and tracking), and sometimes a watchdog to monitor the actions of relevant parties, (2) law enforcement, to complement administrative reforms by ensuring that an appropriate legal system is in place to punish corrupt individuals, and (3) social change, to empower citizens by allowing them to participate in institutional reform by identifying and shaming corrupt individuals as a long-term deterrent to corruption. ICT directly enables administrative reforms and social change by bringing in visibility, transparency, and accountability in previously opaque government processes. While ICT also helps law enforcement source the information needed to prosecute corrupt individuals, the technology may be ineffective without the presence of ICT laws that can allow law enforcement to use the power of ICT for their investigation and prosecution efforts.

3.1 ICT and Transparency

In recent years, many countries have used ICT to initiate administrative reforms, for example, for online tax filing, online business licensing, and distributing financial benefits. These ICT interventions seek to streamline and automate government processes, restrict discretion of officials and intermediaries, and reduce red tape and bureaucracy in government functioning. A less discussed aspect of these reforms is their impact on government transparency and corruption, by increasing visibility of the actions of the state and its agents to its citizenry (Davies & Fumega, 2014). ICT has enabled transparency by greatly reducing the cost of collecting, distributing, and accessing government information, which in turn, is empowering citizens with the ability to track their governments' activities, demand change, and voice against corruption (Roberts, 2006). ICT seems to address corruption in several ways (Chene, 2016):

- By streamlining and automating government processes, ICT limits the discretion of office bearers and reduces their opportunities to extract bribes.
- By removing the role of intermediaries who may demand bribes to "grease the wheels," ICT makes government processes move smoother.
- By reducing red-tape and bureaucracy, which are often entry points for corruption.

- By allowing citizens, press, and other interested third parties to track their governments' actions or inactions, voice dissatisfaction, identify problems, and report corruption.
- By allowing citizens to organize, report and mobilize against corruption using a growing repertoire of online collective action tools and platforms.

The theoretical connection between ICT and corruption can be understood using the lens of information asymmetry (Akerlof, 1970). This perspective suggests that information asymmetry exists when two parties engage in a transaction and one of those parties has greater information about the transaction than the other party. In such situations, the party with more information may try to take advantage of the information asymmetry to extract more economic rent from the transaction, such as selling an inferior quality product at the contracted price. Taking advantage of undisclosed information in an economic contract is also called the "adverse selection" problem.

Prior to the Internet and ICT, government officials held a strong information advantage over the citizenry in the provisioning of public services. Opportunistic officials often took advantage of the opacity of their work to demand bribes and favors. By enabling the collection, storage, and distribution of information to all relevant partied, ICT has eroded that information asymmetry, limiting opportunities for public officials to engage in corrupt behavior, giving citizens more visibility and transparency into government activities, and thereby curbing corruption.

There are many examples of how ICT decreased corruption by improving transparency in government operations (cf. Chene, 2016). For example, the Bhoomi project in Karnataka, India digitized paper-based land ownership records and created a software to control changes to land registry. By replacing discretionary decisions made by public officials with auditable software processes and transparent documentation, this system significantly reduced corruption in the recording and management of land records. In another instance, Brazil created a transparency portal in 2004, which allowed its citizens to view the financial execution of all programs and actions of Brazil's federal government, including transfer of funds from the federal government to state municipalities, disbursements of financial benefit to citizens, and federal government spending on procurement or contracts for projects and services. This portal also improved transparency and reduced corruption in Brazil. Likewise, India created an "I Paid a Bribe" website and invited its citizens to report corruption in various government agencies from tax to police, judiciary, and municipal services via mobile phone or the web to an anti-corruption agency. The system also reports the outcome of those investigations to citizens and displays corruption trends. To date, this website has documented 38,168 instances of bribery worth a total of INR 30 billion and led to the arrest and prosecution of many senior government officers.

However, the anti-corruption effect of ICT may not be instantaneous but may take several years to materialize. The reason for this expectation is that it takes time for public officials to get used to the new transparent environment and understand the import of their actions in this new environment and for the citizenry to gain familiarity with a growing variety of ICT tools to check on government work, report corruption, and mobilize support against corruption. In light of this expectation, we hypothesize:

H1. ICT penetration will have a positive effect on corruption, but this effect may be realized over the medium to long term, rather than over the short-term.

3.2 ICT and Law Enforcement

The corruption literature also suggests that the availability and use of technology is rarely a comprehensive solution to complex problems like corruption and that strong ICT laws and regulations are also needed to prosecute criminals and discourage future offenders (Bhattacherjee & Shrivastava, 2018). Investigative and law enforcement agencies must be empowered to use ICT-based information and tools like mobile data records, spatial mapping software, and prediction modeling to track suspicious activities and transactions and bring corrupt individuals to justice.

ICT regulations can allow investigative agencies and journalists to use surveillance tools such as wiretapping, spy cameras, and sensors to capture corruption data without the offender's knowledge, and courts of law to accept such evidence that are otherwise considered illegal and an invasion to privacy, to build a compelling case against corrupt criminals (Hartle et al., 2011). Investigative journalists have used spycams to record government officials accepting bribes and even ordinary citizens use smartphones to record instances of government malfeasance. Furthermore, data-based policing enabled by ICT can complement the knowledge and experience of law enforcement agencies and help them improve their capacity to detect and prosecute corruption.

To understand why and how ICT regulation may reduce corruption, general deterrence theory (GDT) may serve as a useful guide. GDT is a theory of criminal justice that examines why rational people commit criminal acts and what can be done to prevent such acts (Gibbs 1975). This theory assumes that criminals (particularly, white-collar criminals) are not pathological or disturbed people who need help or psychological counselling, but are rather rational individuals who actions are guided by a cost-benefit calculus. If the expected benefit from committing a crime is higher than its expected cost, they will likely commit the crime. Conversely, if the expected cost is higher, they are less likely to engage in criminal activities.

GDT postulates that criminal acts can be deterred by increasing the expected cost of committing a crime, for example, by increasing the perceived celerity (swiftness), certainty, and severity of punishment related to the crime (Gibbs 1975). Hence, techniques that increase the certainty of punishment (e.g., strong anti-corruption laws and enforcement of such laws by courts of law), celerity (e.g., fast track or special courts for corruption cases), and severity (e.g., mandatory jail time for corruption, minimum sentencing guidelines, and “three strikes law”) can help deter corruption.

Criminals who have been prosecuted for corruption using ICT regulations include American financier Bernie Madoff who masterminded the largest Ponzi scheme in world history (for which he received a 150-year prison term), former Worldcom CEO Bernard Ebbers for illegal corporate accounting (sentenced to a 25-year term), and hedge fund owner Raj Rajaratman for insider trading (sentenced to a 11-year term). Although countries with strong ICT laws can be expected to have lower corruption, it is unknown to what extent this effect is actually realized. If it takes several years for ICT laws to be institutionalized in a country’s criminal justice and

court systems, it is possible that ICT laws may have a multi-year lagged effect on corruption. Based on this expectation, we propose:

> H2. ICT regulation will have a positive effect on corruption, but this effect may be realized over the medium to long term, rather than over the short-term.

4. Research Methods

The proposed hypotheses were testing using data compiled from multiple third-party sources over an eleven-year period from 2010 to 2020. Given the infeasibility of collecting primary data on corruption at the country-level and the prior availability of such data from reputable third-party sources, we chose secondary data for this study.

Corruption data at the country level of analysis was sourced from Transparency International - a non-profit, anti-corruption group that publishes annually a corruption metric called Corruption Perception Index (CPI). CPI is based on a detailed evaluation of corruption in the government, private sector, and individual level, and is widely considered to be the "gold standard" metric of corruption. This metric varies from 0 (high corruption) to 10 (low corruption), i.e., higher values indicate less corruption.

Data on ICT penetration and ICT regulation, the two independent variables in this study, were obtained from the Global Information Technology Report and the World Economic Forum (WEF) respectively. ICT penetration was measured using the total ICT usage sub-index (ICT), which aggregates ICT use among individuals, businesses, and government on a scale of 0 (low penetration) to 10 (high penetration). Individual use includes factors related to citizens' use of ICT, such as the number of mobile phone subscriptions and number of people using the Internet. Business use considers measures related to firms' use of ICT, such as technology absorption, business Internet use, and staff training. Government use includes public initiatives and push for the use of technology, such as prioritization of ICT and provision of online services for citizens, and so forth.

ICT regulation data was derived from WEF's survey of business executives and experts from over 142 countries. The survey is administered annually in 20 languages in multiple formats: face-to-face interview, telephone interview, and mail-in questionnaire. Respondents are asked to evaluate the maturity of ICT laws in their countries, such as those related to digital signatures, electronic commerce, and consumer protection on a scale of 1 (low) to 7 (high). The data has been demonstrated to have high inter-rater reliability and construct validity.

In addition to the above variables, we controlled for the potential confounding effects of countries' education level, economic level, political environment, and media freedom that are presumed to influence corruption. These variables were measured as literacy rate, gross domestic product (GDP) per capita, political stability, and press freedom index respectively. All four variables were sourced from World Bank databases. Country-level literacy rate is a function of the percentage of population aged 15 and higher who can read and write, with adequate understanding, a short, simple statement about their everyday life. GDP per capita is a standard metric of economic prosperity, which, in our data set, is adjusted for purchasing power parity. Political stability is measured using a combination of metrics such as the extent

of armed conflict, violent demonstrations, and social unrest, on a scale of −2.5 (low stability) to 2.5 (high stability). Lastly, press freedom index is calculated based on two components: a quantitative tally of abuses against media and journalists in connection with their work and a qualitative analysis of press freedom in each country based on surveys of journalists and media personnel. This index is scored from 0 (worst possible score) to 100 (best score).

5. Data Analysis and Results

We ran random-effects panel regression models with CPI as the dependent variable, ICT penetration, ICT regulation, literacy rate, GDP per capita, political stability, and press freedom index as predictors, and country and year as random effects, for nine yearly lags between CPI and predictors, as shown in Equation 1. These lags were motivated by our hypotheses, which distinguished between short-term (0-2 years), medium-term (3-5 years), and long-term (6 or more years) effects. Although we had 11 years of data, it was not possible to estimate lags of 9 years or more due to limited sample size. To estimate the marginal effects of ICT penetration and regulation, we compared the above models with a base model that did not include ICT penetration and regulation. The results of this analysis are presented in Tables 1 and 2.

$$y_{ijt_n} = \beta_0 + \beta_1 ICTPenetration_{ijt-n} + \beta_2 ICTRegulation_{ijt-n} + \beta_3 LiteracyRate_{ijt-n} + \beta_4 GDPPerCapita_{ijt-n} + \beta_5 PoliticalStability_{ijt-n} + \beta_6 PressFreedomIndex_{ijt-n} + U_{ij} + Z_{it} + \varepsilon_{ijt} \quad (1)$$

where n (values 0 to 8) is the time lag in years between corruption and its predictors, U_{ij} represents the time-specific random effect, and Z_{it} is the nation-specific random effect.

Predictors	Betas (Standard Errors)				
	Base Model	Lag 0 Model	Lag 1 Model	Lag 2 Model	Lag 3 Model
Country Random Effects	Yes	Yes	Yes	Yes	Yes
Year Random Effects	Yes	Yes	Yes	Yes	Yes
ICT Penetration	-	0.067*** (0.023)	0.067*** (0.025)	0.054** (0.022)	0.120*** (0.027)
ICT Regulation	-	0.032 (0.028)	0.067** (0.026)	0.047* (0.024)	0.071*** (0.026)
Literacy Rate	0.152*** (0.036)	0.063 (0.041)	0.059 (0.039)	0.067* (0.038)	0.030 (0.038)
GDP Per Capita	0.0003*** (0.00005)	0.0003*** (0.00005)	0.0002*** (0.00005)	0.0002*** (0.00005)	0.0001*** (0.00005)
Political Stability	3.365*** (0.775)	3.803*** (0.838)	5.196*** (0.692)	3.822*** (0.599)	4.240*** (0.633)
Press Freedom Index	0.185*** (0.046)	0.167*** (0.048)	0.002 (0.013)	-0.002 (0.010)	0.008 (0.011)
Constant	12.922*** (4.680)	16.252*** (4.996)	25.611*** (3.635)	26.420*** (3.413)	25.760*** (3.521)
Observations	352	290	342	390	361
Log Likelihood	-1,063.369	-879.072	-1,035.593	-1,175.375	-1,102.335
Akaike Information Criterion	2,142.739	1,778.145	2,091.187	2,370.750	2,224.670
Bayesian Information Criterion	2,173.648	1,814.843	2,129.535	2,410.412	2,263.558
*p<0.10; **p<0.05, ***p<0.01 Note: High CPI denotes less corruption; hence positive betas imply negative effects on corruption.					

Table 1: Parameter Estimates of Random Effects Panel Models with 0-3 Year Lags

The lower Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values of the lag 0 model compared to the base model (with no ICT penetration or regulation)

confirmed that adding ICT penetration or regulation to the corruption model increases model goodness of fit and adds explanatory power to the model. In the contemporaneous (zero lag) model, ICT penetration had a significant but small negative effect on corruption (ß=0.067, p<0.01), when controlled for literacy rate, GDP per capita, political stability, and press freedom index, and the effect of ICT regulation was non-significant (ß=0.032, p>0.10). Note that positive beta coefficients imply negative effects on corruption, because the high values of the corruption metric CPI denote less corruption.

Both ICT penetration and regulation had significant but small negative effects on corruption in lag 1 and lag 2 models (ß=0.0054 to 0.067 for ICT penetration and 0.047 to 0.067 for ICT regulation). This effect increased in magnitude for ICT penetration between lag 3 and 8 (ß=0.120 to 0.160), with the highest value in lag 5, before declining gradually to the lag 8 model (ß=0.124). In case of ICT regulation, however, the effect on corruption remained small until lag 7 (ß=0.069 to 0.083, between years 1 and 7), before increasing in magnitude in the lag 8 model (ß=0.128). All of these effects, except one, were significant at p<0.05.

Predictors	Betas (Standard Errors)				
	Lag 4 Model	Lag 5 Model	Lag 6 Model	Lag 7 Model	Lag 8 Model
Country Random Effects	Yes	Yes	Yes	Yes	Yes
Year Random Effects	Yes	Yes	Yes	Yes	Yes
ICT Penetration	0.104*** (0.030)	0.160*** (0.036)	0.130*** (0.036)	0.136*** (0.037)	0.124*** (0.047)
ICT Regulation	0.073** (0.029)	0.071** (0.032)	0.083** (0.032)	0.069** (0.033)	0.128*** (0.038)
Literacy Rate	0.049 (0.044)	0.023 (0.046)	0.060 (0.046)	0.045 (0.046)	-0.003 (0.053)
GDP Per Capita	0.0002*** (0.0001)	0.0001** (0.0001)	0.0002*** (0.0001)	0.0001** (0.0001)	0.0002*** (0.0001)
Political Stability	4.557*** (0.706)	4.101*** (0.748)	3.619*** (0.717)	3.761*** (0.741)	4.649*** (0.879)
Press Freedom Index	-0.005 (0.012)	0.002 (0.013)	0.014 (0.012)	0.017 (0.011)	0.019 (0.013)
Constant	26.105*** (3.945)	26.475*** (4.135)	22.577*** (4.147)	25.033*** (4.188)	25.824*** (4.928)
Observations	333	271	241	214	178
Log Likelihood	-1,046.473	-848.763	-740.974	-665.594	-572.873
Akaike Information Criterion	2,112.947	1,717.527	1,501.948	1,351.187	1,165.745
Bayesian Information Criterion	2,151.028	1,753.548	1,536.796	1,384.847	1,197.563
*p<0.10; **p<0.05, ***p<0.01 Note: High CPI denotes less corruption; hence positive betas imply negative effects on corruption.					

Table 2: Parameter Estimates of Random Effects Panel Models with 4-8 Year Lags

The results of the panel analysis confirm our hypotheses that ICT penetration and regulation have negative lagged effects on corruption at the country level. However, while we expected these effects over the medium to long-term, our analysis also found evidence for short-term effects, although these effects were small in magnitude. In fact, ICT penetration had contemporaneous effects (in lag 0 model), while significant effects of ICT regulation started in the lag 1 model. The strongest effects for ICT penetration were medium-to-long term, while that for ICT regulation was long-term only. Implications for these findings are discussed in the next section.

6. Discussion and Conclusions

6.1 Implications for Research

This study may be one of the earliest to disaggregate the effects of ICT penetration and regulation on corruption into short, medium, and long-term. We present evidence regarding the significance and magnitude of these effects at different lags in years. Although we did not expect ICT penetration or regulation to have immediate effects, our analysis shows that ICT penetration does indeed have an immediate albeit small effect if public officials and the citizenry can take advantage of ICT, while the effect of ICT regulation becomes significant after one year of regulations. Nonetheless, the effect of ICT penetration is strongest after a lag of 3 years until 8 years (and possibly beyond), while a strong effect of ICT regulation is realized only after a lag of 8 years.

Our analysis demonstrates the importance of panel data analysis for analyzing corruption and similar economic data. While cross-sectional models only provide a contemporaneous account of the effect of predictor variables, which may provide an inaccurate estimate of ICT's impact on corruption, panel data models can help uncover the temporal nature of these effects, which may help policy makers formulate appropriate short, medium, and long-term strategies to fight corruption.

ICT includes a wide range of technologies, from generic technologies like servers and networks to specific applications such as surveillance systems and e-government systems. Not all of these applications may help equally in combating corruption. Future research may investigate at a more granular level which systems and applications are more effective at curbing corruption than others. Such research may help developing countries prioritize scarce financial resources and use them more effectively against corruption. Future research may also examine whether the effect of ICT on corruption is more pronounced in business, citizen, or government sectors.

6.2 Implications for Practice

For policy makers, this study demonstrates that ICT penetration and regulation can indeed be effective tools in the war on corruption. Governments should therefore invest in appropriate ICT, inform their public officials and citizenry about these tools, educate them on how to use these tools, and in addition, build a set of ICT laws to provide law enforcement, investigative agencies, and the judiciary the necessary "teeth" needed to fight corruption.

Second, contrary to expectations, our study found that ICT penetration and regulation have short term effects on corruption, although these effects are small in magnitude. This finding may be of interest to developing countries with scarce resources, who may be forced to demonstrate immediate return of investments to sustain their anti-corruption initiatives, against the will of vested interests.

6.3 Concluding Remarks

While prior studies provide mixed evidence on whether ICT can help fight corruption, this study examines how the effects of ICT (specifically ICT penetration and regulation) on corruption unfolds over time. Using panel data analysis, we disaggregate ICT's effects over the short, medium, and long term, and show how long it takes (in years) to realize the anti-

corruption benefits of ICT penetration and regulation and how long those effects last. We hope that our study will stimulate further interest on panel data analysis of ICT and corruption, such as the search for moderators, and will help influence policy choices for countries that are struggling with high levels of corruption.

References

Anderson, T. B. (2009). "E-Government as an Anti-Corruption Strategy." *Information Economics and Policy*, *21*(3), 201-210.

Akerlof, G. A. (1970). "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics*, *84*(3), 488–500.

Bertot, J., P. Jaeger, and J. Grimes (2012). "Promoting Transparency and Accountability Through ICTs, Social Media, and Collaborative E-Government." *Transforming Government: People, Process and Policy*, *6*(1), 78–91.

Bhattacherjee, A. and U. Shrivastava (2018). "The Effects of ICT Use and ICT Laws on Corruption: A General Deterrence Theory Perspective." *Government Information Quarterly*, *35*(4), 703-712.

Bhuiyan. S. H. (2011). "Social Capital and Community Development: An Analysis of Two Cases from India and Bangladesh." *Journal of Asian and African Studies*, *46*(6), 533-545.

Charoensukmongkol, P. and M. Moqbel (2012). "Does Investment in ICT Curb or Create More Corruption? A Cross-Country Analysis." *Public Organization Review*, *14*(1), 51–63.

Chene, M. (2016). Literature Review: The Use of ICTs in the Fight Against Corruption. Transparency International, Retrieved 22 January 2024, from https://www.u4.no/publications/literature-review-the-use-of-icts-in-the-fight-against-corruption.

Davies, T. and S. Fumega (2014). *Mixed Incentives: Adopting ICT Innovations for Transparency, Accountability, and Anti-Corruption*. Michelsen Institute, Retrieved 22 January 2024, from https://www.cmi.no/publications/5172-mixed-incentives.

Elmes, M., D. Strong, and O. Volkoff (2005). "Panoptic Empowerment and Reflective Conformity in Enterprise Systems-Enabled Organizations." *Information and Organization*, *15*(1), 1–37.

Foucault, M. (1977). *Discipline and Punish*. New York: Pantheon Books.

Garcia-Murillo, M. (2013). "Does a Government Web Presence Reduce Perceptions of Corruption?" *Information Technology for Development*, *19*(2), 151–175.

Gibbs, J. (1975). *Crime, Punishment, and Deterrence*. New York: Elsevier.

Gottschalk, P., C. Filstad, R. Glomseth, and H. Solli-Sæther (2011). "Information Management for Investigation and Prevention of White-Collar Crime." *International Journal of Information Management*, *31*(3), 226–233.

Hartle, F., M. Parker, and C. Wydra (2011). "The Digital Case File: The Future of Fighting Crime with Big Data." *Issues in Information Systems*, *15*(1), 257–266.

Heeks, R. (1998). *Information Technology and Public Sector Corruption*. Manchester: Institute for Development Policy and Management.

Johnson, S. (2018, December). *Corruption is Costing the Global Economy $3.6 Trillion Dollars Every Year*. World Economic Forum. Retrieved 22 January 2024, from https://www.weforum.org/agenda/2018/12/the-global-economy-loses-3-6-trillion-to-corruption-each-year-says-u-n.

Kim, S., H. Kim, and H. Lee (2009). "An Institutional Analysis of an E-Government System for Anti-Corruption: The Case of OPEN." *Government Information Quarterly*, *26*(1), 42–50.

Lio, M., M. Liu, and Y. Ou (2011). "Can the Internet Reduce Corruption? A Cross-Country Study based on Dynamic Panel Data Models." *Government Information Quarterly*, *28*(1), 47–53.

Muñoz-Cañavate, A., and P. Hípola (2011). Electronic Administration in Spain: From Its Beginnings to the Present." *Government Information Quarterly*, *28*(1), 74–90.

Roberts, A. (2006). *Blacked Out: Government Secrecy in the Information Age*. New York: Cambridge University Press.

Shim, D. C. and T. H. Eom. (2009). "Anticorruption Effects of Information Communication and Technology and Social Capital." *International Review of Administrative Sciences*, *75*(1), 99-116.

Srivastava, S. C., T. S. H. Teo, and S. Devaraj (2016). "You Can't Bribe a Computer: Dealing with the Societal Challenge of Corruption Through ICT." *MIS Quarterly*, *40*(2), 511-526

OECD (2010). *Collusion and Corruption in Public Procurement*. Organization for Economic Cooperation and Development. Retrieved 22 January 2024, from http://www.oecd.org/competition/cartels/46235399.pdf.
UNDP (2011). *Staying on Track: Tackling Corruption Risks in Climate Change*. United Nations Development Programme. Retrieved 22 January 2024, from http://www.undp.org/content/dam/aplaws/publication/en/publications/democratic-governance/dg-publications-for-website/staying-on-track–tacklingcorruption-risks-in-climate-change/Staying_on_Track_corruption_risk_in_CC. pdf.
UNODC (2005). *Corruption Facts*. United Nations Office on Drugs and Crime. Retrieved 22 January 2024, from http://www.unodc.org/pdf/facts_E.pdf.
Wescott, C. (2001). "E-government in the Asia-Pacific Region." *Asian Journal of Political Science*, *9*(2), 1–24.

22. Unrevealing the digital thread: Exploring students' LMS digital behavior and its impact on academic performance in Kuwait higher education

Ibtisam L. Almutairi
Public Authority for Applied Education and Training
E_almutairi@hotmail.com

Brad McKenna
University of East Anglia
b.mckenna@uea.ac.uk

Adrian Benfell
University of East Anglia
a.benfell@uea.ac.uk

Abstract

This study aimed to investigate the influence of students' digital behavior in the Learning Management System (LMS) on their academic performance. Educational Data Mining (EDM) algorithms, specifically clustering analysis, will be used to analyze student log data, specifically within the context of Kuwait University (KU). By utilizing EDM algorithms, various aspects of students' actual digital behavior will be analyzed, including forum posts and views, frequency of logins, files downloaded, attempts and finalization at exams, and quizzes. Then multiple linear regression will be applied to examine the influence of students' digital behavior in the LMS on their academic performance represented by their grades in LMS log data. The findings of this research could help to better understand students' digital behavior through LMS, which can assist in formulating strategies to enhance student engagement and optimize the learning experience. In addition, these findings can inform the design and implementation of LMS at KU, ensuring that it is more closely aligned with the preferences and expectations of students. Since this alignment comes at a cost, it would be wise to invest in it only if it ultimately contributes to enhancing student academic performance which is the question that will be answered in this study.

Key words: Kuwait University (KU), Learning Management System (LMS), Educational Data Mining (EDM), LMS log data, academic performance.

1. Introduction

In an era marked by the integration of advanced Information and Communication Technologies (ICTs) and the widespread adoption of e-learning systems such as LMS in higher education, it is important to note that the mere availability of these systems does not guarantee improved academic or teaching performance (Al-Fraihat et al., 2020; El-Sayad et al., 2021; Rajabalee & Santally, 2021). This paradox is particularly evident in the context of Kuwait (Ghinea et al., 2013). Despite significant investments in e-learning, the desired benefits have not materialized, and performance has often fallen short of expectations as evident by relatively Low evaluation

metrics of KU's performance (QS Quacquarelli Symonds, 2022; RUR Rankings Russian Federation Agency, 2021; Times Higher Education, 2022).

LMS is a server or cloud-based systems that stores and manages information related to users, programs, and content, thereby meeting the needs of all stakeholders within educational institutions or organizations (Veluvali & Surisetti, 2022). Moodle, one open-source LMS, that is widely recognized for its adaptability and extensive capabilities, making it a popular choice among educators and students (Al-Fraihat et al., 2020b). KU, the primary public institution of higher education in Kuwait, has embraced Moodle as the primary system to foster an interactive learning environment with the aim of enhancing educational outcomes (Kuwait University, 2018). The university integrates Moodle as part of a blended e-learning approach rather than as a replacement for traditional instructional methods. While students typically engage in full-time on-campus attendance, the university utilizes e-learning systems, such as Moodle, to maximize their benefits (KU E-Learning Centre, 2018).

2. Literature Review

Students engage in a range of digital activities while utilizing the LMS (Huang et al., 2020). The LMS records the students' clicks, with each clickable function containing a click counter that tracks the students' digital behavior. EDM is a significant area of study focused on analyzing data obtained from the LMS system. EDM algorithms facilitate the examination of various aspects of students' actual digital behavior, such as their time spent on the LMS, frequency of logins, and types of activities performed. This study aims to address existing knowledge gaps concerning the effects of different indicators on student academic performance within blended learning settings (Bessadok et al., 2021), with a specific focus on at KU which would offer practical insights and contribute to this area. In the field of Information Systems (IS)/ e-learning literature, this study aims to fill the gap in understanding methods for visualizing and structuring students' online behavior by analyzing action logs from LMS rather than depending solely on surveys or interviews (Bessadok et al., 2021; Kara & Yildirim, 2022; Leem, 2023). Furthermore, it extends beyond this by examining the impact of students' online behavior on their academic performance, thereby paving the way for potential future research endeavors.

The substantial volume of data produced by LMS logs has led numerous researchers to utilize various analysis methods to address research questions. Yildirim & Gülbahar (2022) employed Moodle engagement analytics and learners' characteristics to forecast superior final performance by utilizing the Decision Tree algorithm, a machine learning technique that constructs a tree-like model of decisions and their potential outcomes to identify the most significant predictors of high final performance. The study findings support that learner behavior is a significant predictor of their final performance (Yildirim & Gülbahar, 2022). The literature has raised concerns about the Decision Tree algorithm due to its limitations in capturing complex interactions and nonlinear relationships and its classification as a black-box model presents difficulties in interpreting the decision-making process and the rules generated by the algorithm (Huang et al., 2020).

Bessadok et al. (2021) employed EDM, specifically K-means clustering, to uncover the digital activities of students recorded within the Blackboard LMS, including uploaded files, viewed

courses, completed quizzes, and finalized homework assignments. Subsequently, the authors utilized univariate analysis of variance (ANOVA) to evaluate the impact of students' digital activities on their academic performance. The research findings indicated that the analysis of data derived from an LMS platform is essential for enhancing student achievement in e-learning courses (Bessadok et al., 2021). Riestra-González et al. (2021) used similar analysis methods to predict student performance using LMS log data. The research revealed that students who demonstrated promptness in completing quizzes outperformed their peers, whereas those who procrastinated and accessed course materials later exhibited poorer performance (Riestra-González et al., 2021). Clustering analysis has demonstrated its efficacy in discerning patterns, correlations, and outliers within a complex and multifaceted dataset, such as LMS log data which can be instrumental in facilitating informed decision-making and identifying potential issues or challenges within the LMS system (Ramadan et al., 2020). The following section will discuss the literature take of conceptualizing students' LMS data log activities/actions to certain research variables in order to develop research hypothesis.

2.1. Content Interaction

This study specifically focuses on student-content interaction, which can be defined as the manner in which students interact with instructional materials and planned activities within the learning systems (Muir et al., 2022). The literature refers to students' actions registered in the LMS related to all files uploaded/download on the course, and course views as student content interaction (Bessadok et al., 2021; Muir et al., 2022; Vlachopoulos & Makri, 2019). The course view refers to the process of signing into the user interface in a LMS that displays the tasks assigned to students (Bessadok et al., 2021). Buckley et al. (2021) indicated that the number of views a course receives serves as a reliable indicator of a student's interest and understanding of a subject. Furthermore, Vlachopoulos & Makri (2019) suggests that students who choose to download their course materials more frequently demonstrate greater independence in their academic pursuits. These interactions may also signify students' intention to review the information, which can aid in memorization and comprehension. Vlachopoulos & Makri (2019) and Leem (2023) discovered that the interaction between students and the content in e-learning platforms has a direct and positive influence on their performance. Additionally, Muir et al. (2022) underscores the significance of learner-content interaction in promoting student engagement and improving their learning outcomes. As a result, it can be hypothesized that:
H1: Content interaction has a positive effect on student's academic performance.

2.2. Student Engagement

The LMS plays a role in facilitating student engagement, with a specific focus on the significance of student behavioral engagement in relation to LMS usage (El-Sayad et al., 2021). This study will focus on behavioral engagement as it is the engagement type that is documented in the LMS log data, defined as the level of involvement and actions demonstrated by students in their learning and academic pursuits, as measured through active participation in discussions, consistent efforts, and contributions to class discussions (Salas-Pilco et al., 2022). In an e-learning context, it is more feasible to monitor and assess students' engagement behavior, such as their participation in online discussions, Forum posts, Forum views (Lu & Cutumisu, 2022; Muir et al., 2022; Salas-Pilco et al., 2022). Student engagement with LMS through structured and facilitated discussions, such as those in forums or discussion boards, can enhance the quality of relationships and cultivate an environment conducive to voluntary engagement

which offers opportunities a deeper thinking (Vlachopoulos & Makri, 2019). Huang et al. (2020) utilized data on students' online discussion activities and their findings indicated that students who were more engaged and participated more actively in online discussions tended to achieve higher academic performance compared to those who were less active or did not participate at all (Huang et al., 2020). Additionally, Vlachopoulos & Makri (2019) asserted that students who actively engage in online courses are more likely to attain improved learning outcomes. Therefore, it can be argued that:

H2: Student behavioural engagement has a positive effect on student's academic performance.

2.3. Assessment

Assessment is the process of finalizing or attempting to finalize LMS graded activities, such as assignments, exams, or quizzes, by students (Bessadok et al., 2021). Assessment has become a crucial element of higher education and is integral to academic achievement (Lu & Cutumisu, 2022). Online assessment offers learners the opportunity for self- and feedback, enabling them to adjust and regulate their own learning (Lu & Cutumisu, 2022). Additionally, it can provide deeper insights into comprehending the subject matter (Yildirim & Gülbahar, 2022). Assessment activities demonstrate a high level of accuracy in predicting students' academic performance. Furthermore, comparing the predictive capabilities of different subsets of features, the assessment dataset showing a superior accuracy rate (Bessadok et al., 2021). Yildirim & Gülbahar (2022) research indicated that assessments in the form of pre-tests or midterms activities had a positive influence on final grades. Furthermore, Lu & Cutumisu (2022) and Zainuddin et al. (2019) confirmed that online assessments through LMS were effective in evaluating students' learning performance. Consequently, it can be hypothesized that:
H3: Assessment has a positive effect on student's academic performance.
Hypothesis are presented in the expected model in Figure 1.

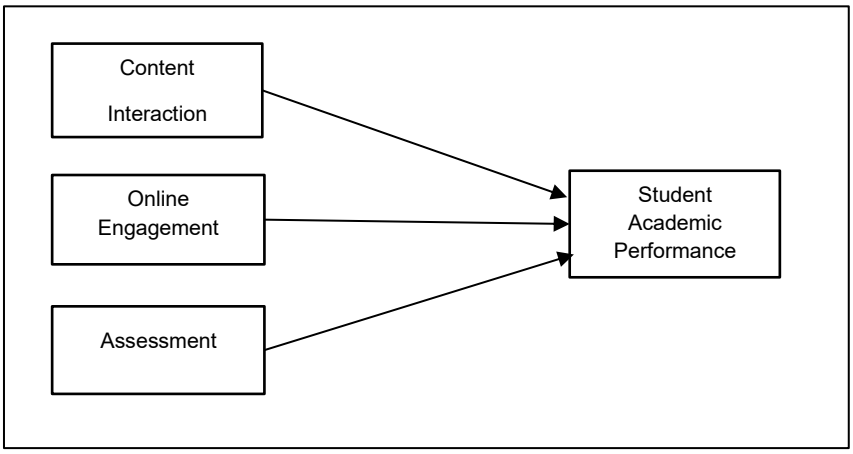


Figure 3 The expected research model.

3. Methodology

3.1. Data Set

The data analysed and processed in this study was obtained from the KU e-learning Centre, specifically the data recorded in the LMS across the various colleges within the university for the academic year 2022-2023. Consisting of 17 colleges, this study focuses specifically on the undergraduate programs offered by its 15 colleges with a population of 35,910 students (Kuwait University, 2018). Each student registered in the system has a record with several attributes, and the variety of these attributes depends on how the system is activated and utilized by the course instructor. To address the research question, a combination of activities and course grades is necessary. Consequently, only student records that meet these criteria are retained for subsequent analysis in this study. Specifically, 8,204 records from the first semester and 7,158 records from the second semester were obtained, resulting in a total of 15,362 student records.

During data preparation phase, the initial log files were processed to clean and prepare them for further analysis. This is essential because Moodle's extracted data sets often contain missing values, noisy data, and irrelevant or redundant information. To address this, the cleaning procedures occurred after the raw log files were initially imported into an Excel worksheet. Identifying data was selectively removed, and the dataset was anonymized by eliminating any personal information.

This study focused on three categories for analysis: content interaction, behavioural engagement, and assessment/evaluation as discussed in the literature review section. The student records kept represent these categories and reflect the digital activities of the students within the LMS, in line with the relevant literature discussed earlier. Detailed set of activities/action types used in the study for analysing students' digital behaviour presented in Table1.

Description	Activity type
Contents interaction	Course view, files downloaded on the course (Bessadok et al., 2021).
Behavioral engagement	Forum posts, Forum views (Lu & Cutumisu, 2022).
Assessment	Exam/ quizzes attempt, finalize exam/ quizzes, Assignment attempt, Finalize assignment (Bessadok et al., 2021; Lu & Cutumisu, 2022; Yildirim & Gülbahar, 2022) .
Student academic performance	Course grade.

Table 2 Student Activities Description

3.2. Analysis Methods

EDM and Learning Analytics are the two main fields of study that focus on analysing the data obtained from an LMS platform (Aldowah et al., 2019). Although there are some differences between the two concepts' purposes and scope, learning analytics is frequently related to EDM (Lemay et al., 2021). While EDM is about utilizing methods for the analysis of learning data to discover previously unknown trends, for example, cluster analysis, learning analytics is a data-driven decision-making strategy that deals with the explanation and contextualization of

that data mostly to address predetermined questions for learning improvement (Tomasevic et al., 2020). Due to the exploratory nature of this study, where we examine the kind of digital activities students engage in within the LMS and whether these behaviours influence on academic performance, EDM appears relevant.

In order to analyse the digital behaviour of distinguished students and categorize them into clusters, the study utilized EDM clustering analysis, specifically K-means, a non-hierarchical method that aims to partition a given dataset into a predetermined number of clusters (k) by iteratively assigning data points to the nearest cluster centroid (Ramadan et al., 2020). One challenge associated with this type of analysis is the determination of the number of clusters by the researcher (Tomasevic et al., 2020). Various methods can be employed to determine the number of clusters, one of which involves using Hierarchical Cluster Analysis (Ramadan et al., 2020). In this study, IBM SPSS 29 was used to apply Hierarchical Cluster Analysis using Ward's method to create evenly sized clusters. The dendrogram was visually inspected to identify a reasonable number of clusters for the final clustering solution, resulting in the identification of two distinguished clusters (two unique digital behaviours).

Subsequently, K-means algorithms were applied using IBM SPSS 29, involving the performance of several clusters with different predetermined numbers until the optimal number of clusters was obtained, which also found to be two, confirming the previous analysis. This approach enabled the identification of distinct digital behaviours for each of the two clusters.
In order to examine the significance of the differences in means among the independent variables of the clusters, this research will extend the analysis by employing Multiple Linear Regression through IBM SPSS 29. This approach is especially suitable for forecasting student achievement using LMS log data (Kara & Yildirim, 2022). The Multiple Linear Regression will be utilized to assess the research hypothesis and ascertain whether the student digital behaviour determined by clustering, can predict student academic performance.

4. Conclusion

In conclusion, this study aims to uncover the hidden relationships between students' digital behaviour in the LMS and their academic performance at Kuwait University. By utilizing EDM algorithms and multiple regression, researchers plan to identify key patterns and predictors of success within the KU LMS environment. Such insights have the potential to inform personalized interventions. Also, understanding student preferences and engagement patterns within the LMS can guide the development of adaptive learning platforms and personalized instruction. Furthermore, the findings of this study would facilitate data-driven decision-making in resource allocation, curriculum development, redesigning learning systems, and prioritizing features.

References

Aldowah, H., Al-Samarraie, H., & Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, *37*, 13–49. https://doi.org/10.1016/J.TELE.2019.01.007

Al-Fraihat, D., Joy, M., & Sinclair, J. (2020). Evaluating E-learning systems success: An empirical study. *Computers in Human Behavior*, *120*, 67–68. https://doi.org/10.1016/j.chb.2019.08.004

Bessadok, A., Abouzinadah, E., & Rabie, O. (2021). Exploring students digital activities and performances through their activities logged in learning management system using educational data mining approach. *Interactive Technology and Smart Education*. https://doi.org/10.1108/ITSE-08-2021-0148/FULL/HTML

Buckley, K. , Fairman, K. , Pogge, E., & Raney, E. (2021). Novel use of LMS data to predict online learning success in a pharmacy capstone course. *American Journal of Pharmaceutical Education.* https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&as_ylo=2019&inst=15450697844440806189&q=Novel+Use+of+LMS+Data+to+Predict+Online+Learning+Success+in+A+Pharmacy+Capstone+Course.&btnG=

El-Sayad, G., Hasliza, N., Md Saad, N., & Thurasamy, R. (2021). How higher education students in Egypt perceived online learning engagement and satisfaction during the COVID-19 pandemic. *Journal of Computers in Education*, *8*(4), 527–550. https://doi.org/10.1007/s40692-021-00191-y

Ghinea, G., Alkharang, M. M., & Ghinea, G. (2013). E-learning in Higher Educational Institutions in Kuwait: Experiences and Challenges. *International Journal of Advanced Computer Science and Applications*, *4*(4). https://doi.org/10.14569/IJACSA.2013.040401

Huang, A. Y. Q., Lu, O. H. T., Huang, J. C. H., Yin, C. J., & Yang, S. J. H. (2020). Predicting students' academic performance by using educational big data and learning analytics: evaluation of classification methods and learning logs. *Interactive Learning Environments*, *28*(2), 206–230. https://doi.org/10.1080/10494820.2019.1636086

Kara, M., & Yildirim, Z. (2022). Faculty Performance Improvement in Distance Education: Causes of the Performance Deficiencies (Part I). *Performance Improvement Quarterly*, *34*(4), 573–601. https://doi.org/10.1002/PIQ.21367

KU E-Learning centre. (2018). *Kuwait eLearning strategy*. http://ku.edu.kw/DLC/index.htm

Kuwait University. (2018). *About Kuwait University*. https://kuweb.ku.edu.kw/ku/index.htm

Leem, B. (2023). Impact of interactivity on learning outcome in online learning settings: Ordinal logit model. *Journal of Engineering Business Management*, *15*, 1–10. https://doi.org/10.1177/18479790231203107

Lemay, D. J., Baek, C., & Doleck, T. (2021). Comparison of learning analytics and educational data mining: A topic modeling approach. *Computers and Education: Artificial Intelligence*, *2*, 100016. https://doi.org/10.1016/J.CAEAI.2021.100016

Lu, C., & Cutumisu, M. (2022). Online engagement and performance on formative assessments mediate the relationship between attendance and course performance. *International Journal of Educational Technology in Higher Education*, *19*(1). https://doi.org/10.1186/S41239-021-00307-5

Muir, T., Wang, I., Trimble, A., & Mainsbridge, C. (2022). Using interactive online pedagogical approaches to promote student engagement. *Education Sciences*, *12*(6), 415. https://www.mdpi.com/2227-7102/12/6/415

QS Quacquarelli Symonds. (2022). *QS World University Rankings*. https://www.qs.com/

Rajabalee, Y. B., & Santally, M. I. (2021). Learner satisfaction, engagement and performances in an online module: Implications for institutional e-learning policy. *Education and Information Technologies*, *26*(3), 2623–2656. https://doi.org/10.1007/S10639-020-10375-1

Ramadan, R. A., Alhaisoni, M. M., & Khedr, A. Y. (2020). Multiobjective clustering algorithm for complex data in learning management systems. *Complex Adaptive Systems Modeling*, *8*(1). https://doi.org/10.1186/S40294-020-00071-9

Riestra-González, M., Del Puerto Paule-Ruíz, M., Ortin, F., & Del Puerto Paule-Ruiz, M. (2021). Massive LMS log data analysis for the early prediction of course-agnostic student performance. *Computers & Education*, *163*, 104108–104128. https://doi.org/10.1016/j.compedu.2020.104108

RUR Rankings Russian Federation Agency. (2021). *Round University Ranking RUR*. Thomson Reuters. https://roundranking.com/

Salas-Pilco, S. Z., Yang, Y., & Zhang, Z. (2022). Student engagement in online learning in Latin American higher education during the COVID-19 pandemic: A systematic review. *British Journal of Educational Technology*, *53*(3), 593–619. https://doi.org/10.1111/BJET.13190

Times Higher Education. (2022). *Times Higher Education's global portfolio of university rankings*. https://www.timeshighereducation.com/

Tomasevic, N., Gvozdenovic, N., & Vranes, S. (2020). An overview and comparison of supervised data mining techniques for student exam performance prediction. *Computers & Education*, *143*, 103676. https://doi.org/10.1016/J.COMPEDU.2019.103676

Vlachopoulos, D., & Makri, A. (2019). Online communication and interaction in distance higher education: A framework study of good practice. *International Review of Education*, *65*(4), 605–632. https://doi.org/10.1007/S11159-019-09792-3

Yildirim, D., & Gülbahar, Y. (2022). Implementation of Learning Analytics Indicators for Increasing Learners' Final Performance. *Technology, Knowledge and Learning*, *27*(2), 479–504. https://doi.org/10.1007/S10758-021-09583-6

Zainuddin, Z., Shujahat, M., Chu, S. K. W., Haruna, H., & Farida, R. (2019). The effects of gamified flipped instruction on learner performance and need satisfaction: A study in a low-tech setting. *Information and Learning Science*, *120*(11–12), 789–802. https://doi.org/10.1108/ILS-07-2019-0067/FULL/HTML

23. Deep learning methods for hepatic tumor detection

Manal Makram
Department of Data Science FGSSR
Cairo University
manal.hana64@gmail.com

Mohammad Elhemeily
Radiology Department
Ain Shams University
m.elhemeily@med.asu.edu.eg

Ammar Mohammed
Faculty of CS, MSA University & FGSSR,
Cairo University
ammohammed@msa.edu.eg

Abstract

LIVER tumors pose a substantial global health threat, and in Egypt, they are a significant public health concern. A report in 2020 from the World Health Organization indicated that hepatic tumors contribute to 4.57% ranking Egypt second globally in hepatic tumor mortality and the first in liver disease-related deaths, which constitute 11.20% of total mortality in the nation. It's encouraging to note that within less than a decade, Egypt has drastically reduced its hepatitis C prevalence from 10% to just 0.38%, achieving one of the lowest rates globally by 2023. Recognizing the urgency of accurate diagnoses, the development of computer-assisted imaging techniques using deep learning has become pivotal in recent years. This paper makes noteworthy contributions to the field of artificial intelligence (AI) in the context of medical record digitization. Leveraging AI, especially in medical record systems, has become increasingly essential. The proposed liver dataset, obtained from Ain Shams University Specialized Hospital (ASUSH) in Egypt and annotated by expert radiologists, comprises 280 patients and encompasses approximately 14,096 computed tomography (CT) images sourced from the Picture Archiving and Communication System (PACS). The experiments conducted in two phases are significant. In the initial phase, expert radiologists classify the dataset into binary classification (normal vs. abnormal) cases. Subsequently, in the second phase, the radiologists further categorize abnormal cases into multi-class classifications, these categories include non-tumorous abscesses, malignant metastases, malignant hepatocellular carcinoma treated with radiofrequency ablation, malignant HCC, liver cirrhosis, benign hemangioma, and focal fatty sparing. Notably, this dataset stands out as the first multi-class liver cancer dataset in Egypt, making a novel and impactful contribution to the field. The thesis also benchmarks the dataset using various deep learning models, including VGG16, VGG19, CNN, ResNet50, ResNet101, EfficientNetB1, EfficientNetB2, Xception, Vision Transformer (ViT), Inception-V3, and InceptionRestNetV2. The experimental results are remarkable, with all models achieving an impressive 98% accuracy in predicting liver cancer variants. The evaluation includes several performance metrics, highlighting the robustness and effectiveness of the proposed models in liver cancer diagnosis.

Keywords: Deep Learning methods, Hepatocellular Carcinoma, Vision Transformer, Transfer Learning, Computed Tomography (CT) scans.

1. Introduction

The human body contains 79 organs, the liver is the second-largest internal organ as well as the largest gland within the human body, the liver plays a pivotal role in numerous physiological functions, it carries out more than 500 vital functions. The liver is located on the right side of the abdomen and weighs about three pounds. The liver is primarily composed of hepatocytes, bile canaliculi, and hepatic sinusoids. Cancer is a major global health concern because it can affect various parts of the human anatomy, including the liver. The biological behavior of liver tumors can classify them as benign or malignant. Benign tumors are non-cancerous growths that do not spread to other areas of the body (Sinthuja, Hatti, and Thavamani, 2022; Marrero et al., 2005; Tan & Le, 2019), whereas malignant tumors are cancerous and have the capacity to metastasize to distant organs (Chaffer, Christine, & Robert, 2011; Sween, Samar & Binu, 2018; Sharma, N, 2023; Kader, Andronikou, & Welman, 2001; Nemane et al., 2022). In light of this, early detection of liver tumors is essential to successful treatment and better patient outcomes, such as reducing time and getting quick treatment. Numerous diagnostic procedures, including imaging tests like computed tomography (CT) scans, MRI, US, and PET-CT scans, are required to detect liver tumors. The Picture Archiving and Communication System (PACS) relies on the triple-phase liver CT protocol can be used to evaluate focal liver lesions, endocrine tumors, and hyper vascular liver metastases (Murphy, Dabirifar & Bell, 2023). The triple-phase liver is an initial non-contrast image included before the intravenous contrast medium is given, commonly requested for patients who received an injection of drug-eluting beads. Additionally, the four-phase liver CT protocol is a useful examination for liver tumors. This protocol includes a non-contrast liver phase, followed by a dedicated late arterial phase, portal venous phase, and delayed phase acquisition. The subsequent phases offer detailed imaging of the liver's arterial and venous supply, providing essential insights for identifying various types of liver lesions or tumors. The delayed phase acquisition further enhances details about the lesions, particularly in terms of their enhancement patterns over time (Shaked et al., 2011). In Egypt, liver cancer has been ranked as the first form of cancer among males, while it ranks second after breast cancer among females (Ibrahim et al., 2014). This paper is crucial in addressing the difficulties and opportunities in liver tumor detection and diagnosis, cutting-edge technology, and innovative research, all of which underscore the importance of this thesis in addressing the challenges and opportunities in liver tumor detection and diagnosis. At Ain Shams University, the exploration of cutting-edge ideas, visions, scientific insights, and technological applications in the field of liver diseases is pivotal for the millions of patients facing these health challenges. The objective spans diagnosis, treatment, surgery, and organ transplantation, showcases advancements that produce tangible results, offer valuable recommendations, and contribute to the global endeavor in diagnosing and treating liver diseases. This pursuit underscores the university's commitment to advancing knowledge and making meaningful contributions to the broader scientific and medical community. AI technology has made significant progress in recent years, leading to the development of more efficient solutions for a variety of tasks. This has resulted in exponential growth in the demand for AI-powered tools and services in the medical field. AI is more capable of solving numerous issues than radiography and biopsy. Artificial intelligence technology can assist medical professionals in providing healthcare services (Olczak et al., 2021).

2. Related Work

Deep Learning techniques have been effectively implemented in a variety of application domains (Gabruseva, Poplavskiy & Alnin, 2020; Kora & Mohammed, 2023; Rana & Bhushan, 2023; Amin, Hefny & Mohammed, 2021; Makram, Elhemeily & Mohammed, 2023; Mansour et al., 2022; Jiang et al, 2023; Abdelhay et al., 2023). The efficient application of deep learning has produced notable developments in the medical and healthcare fields, enabling accurate diagnosis (Shafiq & Gu, 2022; Alabrak et al., 2023; Makram, Ali & Mohammed, 2023; Mazurowski et al., 2019; Munir et al., 2019). Many papers as in (Jiang et al., 2023) concerned their research the medical images by different types of scans X-ray, ultrasound (US), computed tomography (CT), magnetic resonance imaging (MRI), positron emission computed tomography (PET) (Zhou et al., 2019; Nishida et al., 2022; Wang et al. 2020), and histopathological images. Their paper discussed advanced neural networks that have emerged recently, such as transfer learning, ensemble learning (EL), graph neural networks, and ViT. To prevent overfitting, contemporary research underscores the efficacy of batch normalization, dropout, weight initialization, and data augmentation techniques. This focus is part of a broader research effort exploring the application of deep learning algorithms, specifically (CNNs), for liver segmentation and hepatic lesion classification in medical imaging. For instance, in the study published in (Hamm et al. 2019), outperforming two board-certified radiologists, CNN was used to classify common hepatic lesions in MRI images with accuracy, sensitivity, and specificity of 92%, 92%, and 98%, respectively. The research demonstrates that hepatic lesion classification can be enhanced by deep learning algorithms. Similarly, the authors (Zhen et al. 2020) developed a CNN model to classify liver tumors using enhanced and unenhanced MR images and clinical data. Their proposed system was compared to the performance of three expert radiologists, the AUC was 0.946, with a 95 % confidence interval ranging from 0.914 to 0.979. This compared closely to the radiologists' AUC of 0.951, with a confidence interval of 0.919 to 0.982, and a P-value of 0.664, indicating statistical significance in the comparison. Likewise, the authors of the study presented in (Wang et al., 2020) proposed an approach using deep residual neural networks with a transfer learning strategy for classifying the pathological types of lung tumors from CT images. The proposed approach achieved an accuracy of 85.71% in identifying the pathological types of lung tumors. Similar work has been suggested by (Tang et al., 2020), where the authors aim to evaluate the use of CNN for liver segmentation in CT images of patients receiving selective internal radiation therapy (SIRT) patients and reduce inter-observer variability. The model performed well for CT images with normal liver shapes and low tumor burden but required minor clinical adjustments for 87% of the segmentation in images with unusual liver positions or shapes. In the study by (Young et al., 2002), a system based on deep learning for joint segmentation of the liver and hepatic tumors in CT images was presented. The results showed promise for joint liver and tumor segmentation using deep learning methods, highlighting the method's promising capabilities in medical imaging analysis. Similar work on segmentation is proposed in (Ronneberger, Fischer and Brox, 2015), where the authors also utilized the U-Net architecture to effectively segment neuronal structures in electron microscopy images, outperforming prior state-of-the-art techniques in this regard. The advantages include a symmetric network structure, improved feature fusion through deconvolution and convolution, and the use of skip connections to concatenate features. According to the authors in, (Biswas et al. 2018) deep learning systems using CNN performed better than traditional machine learning systems at fatty liver disease risk stratification and detection. Similarly, the authors (Wang et al., 2019) identified different focal

liver diseases using a sparse auto-encoder and a SoftMax layer, showcasing the versatility and efficacy of deep learning techniques in medical diagnosis. Another contribution was a deep convolution neural network (CNN) model with transfer learning for liver steatosis put forth by the authors (Byra et al., 2018). Deposition of fat in the liver which causes death (Biswas et al., 2018) assessment in B-mode ultrasound images. They used a pre-trained deep CNN on the ImageNet dataset to extract high-level features, which were then used by a support vector machine algorithm to classify images. They used the features and Lasso regression method to evaluate the degree of steatosis. Additionally, the authors (Constantinescu et al., 2021) recommended using different CNN architectures to classify fatty liver disease in images, and they suggest further studies to establish the clinical implementation of these algorithms. In the study, (Yuan et al., 2021) propose Vision Transformers (ViT) as an innovative alternative to traditional CNNs for computer vision tasks. They emphasize ViT's reliance on the Transformer architecture and a unique patch-based approach for image processing, enabling effective handling of complex natural images. This marks a significant shift in the methodology for tasks like object detection and image classification, positioning ViT as a promising new paradigm in computer vision.

3. Evaluation Metrics

Various metrics are utilized to evaluate the performance of deep-learning liver tumor classifiers, with the confusion matrix offering a comprehensive view of their effectiveness. The accuracy metric defined in (Equation 1), is described as the model's ability to correctly predict the class label of new or unseen data. Additionally, Sensitivity (Equation 2) and specificity (Equation 3) further assess the classifier's capability to distinguish between actual positive cases and negative cases, respectively (Purwar & Singh, 2015). On the other hand, precision as defined in (Equation 4), measures the number of correctly classified positive tumor cases to the total number of examples classified as positive. The precision-recall harmonic mean is encapsulated in Equation (5). In the context of diagnosing Hepatocellular Carcinoma (HCC), it becomes imperative to calculate or extract the true negative (TN), false negative (FN), false positive (FP), true positive (TP), specificity, and sensitivity values directly or indirectly.

$$\text{Accuracy (Acc)} = \frac{TP+TN}{TP+FN+TN+FP} \quad \text{Equation 1}$$

$$\text{Sensitivity} = \text{Recall} = \frac{TP}{TP+FN} = \frac{TP}{Diseased} \quad \text{Equation 2}$$

$$\text{Specificity} = \frac{TN}{FP+TN} = \frac{TN}{NoDiseased} \quad \text{Equation 3}$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad \text{Equation 4}$$

$$\text{F1-Score} = 2 * \frac{Precision*Recall}{Precision+Recall} \quad \text{Equation 5}$$

Receiver Operating Characteristics (ROC) graphs serve as a prevalent technique for visualizing and assessing the effectiveness of classifiers, particularly used in medical decision-making. These graphs the relationship between the false positive rate (FPR) and the true positive rate (TPR) across different classification thresholds for a binary classifier. The resulting curve offers a means to compare the performance of different classifiers. The decision to select an

optimal classifier is often guided by the area under the ROC curve (AUC), where higher AUC values indicate better overall performance (Fawcett, 2006).

4. Proposed Approach

In this paper, deep-learning models were employed to identify liver tumors in CT images. The proposed approach consists of a specified pipeline and provides details regarding the dataset used for evaluating this pipeline.

4.1 The proposed pipeline

The study approach pipeline deploys deep learning models of the medical liver CT image classification involves several stages. Every stage in the pipeline takes the output from the stage before it as an input and produces outputs as results that are utilized by the next stage. The proposed approach pipeline has the following stages as shown in figure 1. The steps of the suggested approach pipeline are as follows.

4.1.1 Data Acquisition

The data acquisition stage involves collecting, handling, and preparing CT liver images from the hospital PACs system and converting images from DICOM to TIFF format. Collaboration with radiologists, careful examination, and classification of images, along with the incorporation of laboratory tests and transfer learning, contribute to building a robust system for the detection of liver tumors. Ethical considerations and privacy concerns are prioritized throughout the data acquisition process. The main emphasis is on creating a well-rounded dataset that combines visual and textual information, expert knowledge, and cutting-edge deep learning techniques to enhance the accuracy and effectiveness of liver tumor detection.

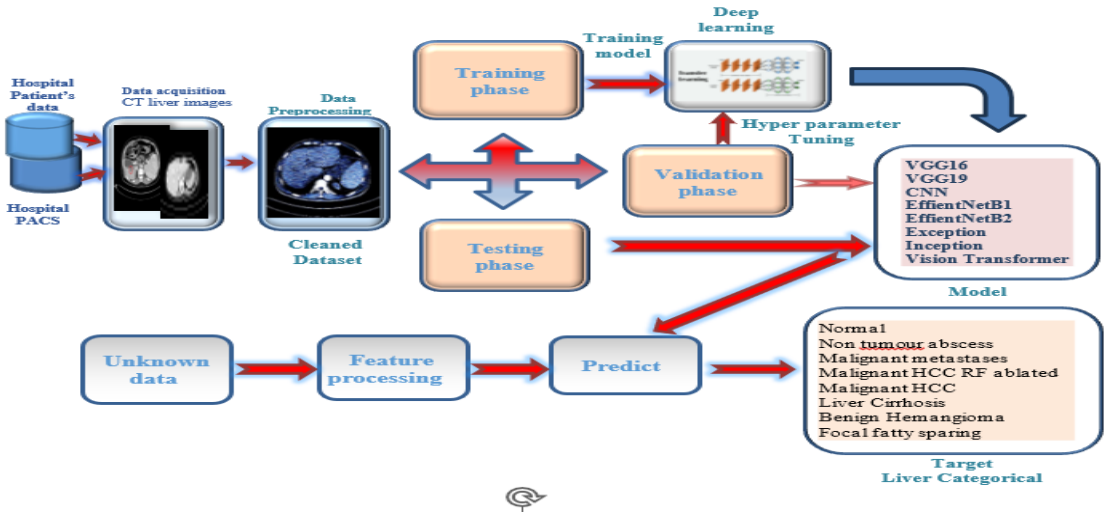
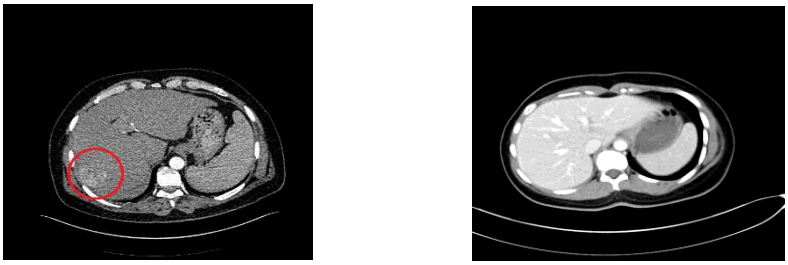


Figure 4: Pipeline aproach

4.1.2 Data Preprocessing

in the data pre-processing stage, several crucial steps are undertaken to enhance the quality and uniformity of the dataset for effective deep learning model training.



(a) CT scan showing HCC liver tumor (b) CT scan of a non-diseased liver

Figure 05: shows CT scans from the ASUSH medical dataset

Figure 2 illustrates two different conditions of the liver. The red circle on the diseased liver image is used to highlight the area of the HCC tumor as in fig. (2.a). Such images are valuable in medical research and practice for diagnosing, studying, and understanding liver diseases, including liver cancer.

4.1.3 Role of Visualization

Visualization aids in understanding complex datasets, particularly CT images, crucial for diagnosing liver health, especially in cases of Hepatocellular Carcinoma (HCC), by distinguishing between diseased and non-diseased liver images (Cheng et al., 2021; Zhou et al., 2019; Nemane et al., 2022; Mishra & Mohapatro, 2019; El-Zayadi et al., 2005).

4.1.4 Model Creation

In the model creation stage, various deep learning models for liver tumor detection, including VGG, ResNets, EfficientNets, ViT, and CNN, are deployed using transfer learning on ASUSH CT medical images. This approach aims to facilitate the diagnosis and treatment of patients with liver tumors. Moving to model selection and training, the dataset is divided into training, validation, and testing sets in an 80:20 ratio. Multiple models are trained using the training set to maximize accuracy. The Grid Search CV method is utilized to optimize hyperparameters, enhancing the performance of the models (Olczak et al., 2021).

4.1.5 Model Evaluation

In the Model Evaluation phase, a systematic approach is followed to assess the performance and generalization ability of the trained model. The evaluation process culminates in two distinctive views. The first view involves classifying the dataset into a binary categorization of normal or abnormal. In the second view, the liver dataset is further categorized into specific conditions, including Normal, Non-tumorous abscess, Malignant metastases, Malignant HCC ablated, Malignant Cirrhosis, Liver Hemangioma, Benign Hemangioma, and Focal fatty sparing. This comprehensive classification provides detailed insights into the model's diagnostic capabilities across various liver conditions. Implemented entirely using the Python programming language and leveraging the TensorFlow, Keras, and PyTorch libraries, the training and testing process demonstrates the flexibility and efficiency of these tools in developing a robust model for liver tumor detection. ColabPro Plus, a cloud-based computing platform, was employed to harness the computational power required for training transfer learning models.

4.2 Dataset Description

This paper, authored by the undersigned, utilizes a medical database, specifically an electronic health record system, to identify patients treated by liver oncology specialists and radiologists. The medical records used in the study were sourced from the ASUSH hospital PACs databases and servers. The dataset used in the study is from real historical data collected between 2018 and 2023. The study employed a robust methodology, systematically expanding the sample size to encompass 280 patients sourced from ASUSH. The downloading of scans necessitated considerable time and effort, particularly in filtering the images to isolate those featuring the liver or other organs such as the lungs and heart. Each patient's scans could comprise thousands, contingent on the interval between scan sessions. The dataset under study contains more than 14,096 CT images. It encompasses both normal and abnormal male and female subjects, with ages ranging from 18 to 90 years old. The liver tumor dataset includes various types of lesions. This approach likely enhanced the statistical reliability of the results. The expert radiologists classified the dataset into two views. The first view is binary classification (normal vs. abnormal) cases. Subsequently, in the second view, for multi-class classification. The researcher primarily focuses on and addresses the following aspects. When it comes to detecting hepatic tumors, the attention of oncologists and radiologists is directed toward the type of liver CT images, specifically those of pelvises. These images are captured in three different phases using a contrast dye, and the research also involves an array of laboratory tests. The dataset was enriched by including lists of patients under the care of participating doctors and the International Classification of Diseases and Related Health Problems (ICD10) for 2023/2024. The dataset includes patients who have undergone specific types of liver surgeries. In addition to medical images, various types of data, such as structured, unstructured, and semi-structured data, were collected and pre-processed using Oracle (PL/SQL), forms, and reports developer. The dataset used in this study is stored in a data center that stores various types of data, such as laboratory results, radiology reports, PAC images, and patients' demographics, especially age and gender. We used the patients' medical images in DICOM and TIFF formats and had an expert radiologist from ASUSH review and classify the images and doctors' reports to diagnose the patient's condition. Any discrepancies were corrected to ensure an accurate diagnosis. The procedures used are part of the regular daily medical practice. We employed transfer-learning methods with Keras to aid in the detection of liver tumors (Baykal et al., 2020; Wang et al., 2019).

5. Experimental Results

5.1 Views and Hyperparameters

The work's author, with the expert radiologists, classified the dataset into two views. the patients were divided into two main categories. First, Normal: healthy liver with no focal lesions of any nature. Second, Abnormal: focal hepatic lesion is detected regardless of its nature. The expert radiologists further divided the dataset into multi-class frameworks in the second view. Plus, a cloud-based computing platform was used. Utilizing Google Colab with a GPU significantly boosts the performance and speed of training deep learning models. The experiments' training and testing phases are conducted using the Python programming language, with support from TensorFlow and Keras libraries, both crucial in the model training and evaluation process, especially for transfer learning models. Overall, the experimental environment description provided in the section is comprehensive, and the selection of libraries and hardware demonstrates careful consideration of resources and tools necessary to achieve the thesis goals.

The choice of VGGs, CNN, exception, inception, EffientNets, ViT, and ResNets as the base models for transfer learning is well-justified, as they are well-known architectures with strong performance on image recognition tasks. This paper demonstrates a comprehensive analysis of hyperparameters, covering a range of options to optimize model performance. The findings, particularly in the context of liver disease diagnosis, have the potential to improve accuracy and consistency, aiding radiologists in making more informed decisions.

Table 1(a) illustrates the hyperparameters that cover a wide range, demonstrating a systematic exploration of options to fine-tune the model. The various combinations allow for finding the optimal set of hyperparameters that balance model complexity and generalization. The findings could have a significant impact on the field of liver disease diagnosis using CT images, reducing potential variability in report results based on the experience and proficiency of radiologists in liver disease diagnosis. Table1(b) explains the hyper parameters used for fine-tuning ViT model. Table 1(c) provides a comprehensive list of hyperparameters and the settings for the training model, and it includes various augmentation techniques, such as rotation, width and height shift ranges, rescaling, shearing, brightness adjustments, and flips both horizontally and vertically. These augmentations are crucial for training robust models by introducing variability, which helps in generalizing better to new data.

Hyper Parameters			Value
Image Size (pixels)			224x224
Batch size			32
Early Stopping			Yes
Learning Rate			0.000004
Split data			80:20
weights			Imagenet
Hyper parameters combination	regularization	L1	0.01, 0.001, 0.0001
		L2	0.01, 0.001, 0.0001
	Dropout		0.5, 0.8
	optimizers		RMSprop, Adam, Adagrad

(a) General hyper parameters

MODEL	Vision transformer
Early Stopping	Yes
Regularization Techniques	Data Augmentation, weight Decay, and Early Stopping
Evaluation Metrics	precision, recall, ROC/AUC and F1-score
Rotation range	20
Drop out	0.5
Weights	Imagenet
Learning Rate	3×10^{-5}
Loss Function	Cross Entropy Loss
Optimizer	Adam
Ignore Mismatched Sizes	True

(b) Vision transformer

Augmentation Hyper parameters	Value
Rotation range	15
Width shift range	0.05
Height shift range	0.05
Rescale	1/255
Shear range	0.05
Brightness range	0.1,1.5
Horizontal flip	True
Vertical flip	True

(c) Data augmentation techniques

Table 1: Hyper parameters used for fine-tuning ASUSH dataset

5.2 Experimental Outcomes

The experiments are executed in two rounds.

5.2.1 PART ONE (View 1): Binary classification

The dataset is for binary classification (Normal vs Abnormal). The dataset comprises 14,095 CT images, with 7,473 CT images labeled as normal patients and 6,623 as abnormal. The deep-learning-based approach for detecting hepatic lesions in CT images resulted in improved accuracy compared to other algorithms. Performance metrics, including accuracy, precision,

recall, F1-Score, and ROC-AUC are presented in Table 4. The dataset is split into 11,276 CT images for training, and 2819 CT images for validation and testing. Eleven deep learning models are compared, and the overall experimental results showcase their effectiveness in detecting liver tumors. The performance metrics displayed in Table 2 highlight the models' high accuracy in classifying CT liver images into normal or abnormal categories. These models consistently exhibit strong performance in both training and testing. ViT model stands out by achieving perfect sensitivity/recall, precision, and F-1 score, possibly attributed to its capability to capture global dependencies in the images. While most models demonstrate high specificity, slight variations exist. Maintaining high specificity is crucial in medical diagnostics to minimize false positives. The overall consistency in performance metrics across diverse models emphasizes their robustness in the task of liver tumor detection from CT images.

Table of results view one (Normal or Abnormal CT images)											
MODEL	**VGG16**	**VGG19**	**CNN**	**ResNet-50**	**ResNet-101**	**EfficientNetB1**	**EfficientNetB2**	**Vision transformer**	**Xception**	**Inception-V3**	InceptionResNetV2
Training Accuracy	0.9970	0.9966	0.9975	0.9947	0.9975	0.9982	0.9961	0.9999	0.9929	0.9929	0.9929
Testing Accuracy	0.9929	0.9965	0.9975	0.9947	0.9975	0.9982	0.9961	0.9979	0.9929	0.9929	0.9929
ROC/AUC	0.9994	0.9999	0.9996	0.9985	0.9975	0.9993	0.9996	0.9999	0.9963	0.9944	0.9171
Sensitivity	0.9929	0.9965	0.9975	0.9947	0.9975	0.9982	0.9961	1.00	0.9929	0.9929	0.9929
Specificity	0.9993	0.9993	0.9987	0.9987	0.9993	0.9973	0.9987	1.00	0.9973	0.9973	0.9973
Precision	0.9930	0.9965	0.9965	0.9947	0.9975	0.9982	0.9961	1.00	0.9929	0.9929	0.9929
F-1 Score	0.9929	0.9965	0.9965	0.9947	1.00	0.9982	0.9961	1.00	0.9929	0.9929	0.9929

Table 2: Experimental Results for view one

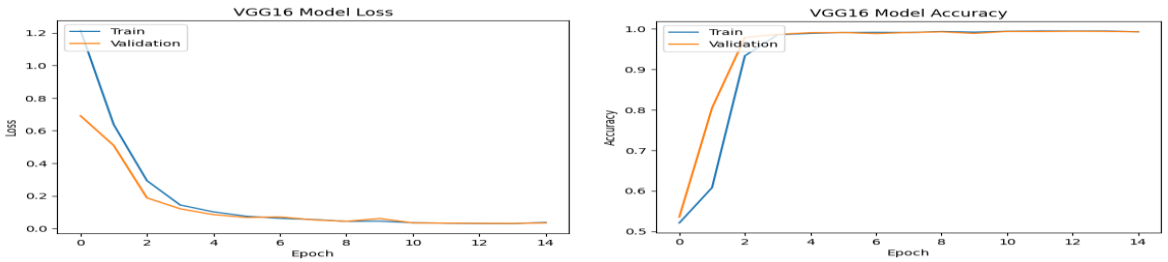


Figure 3: CNN model

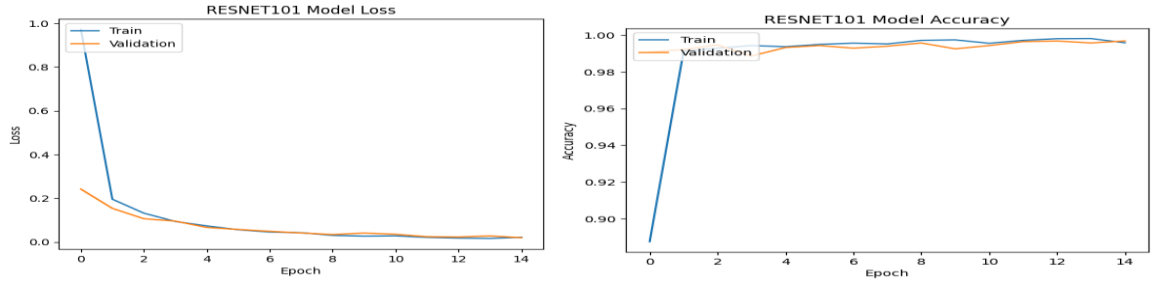


Figure 4: ResNet101 model

The comparison of accuracies, as shown in Figures 3, and 4 provides insights into the performance of different models on the task. The models exhibit high training and testing accuracies, indicating effective learning and generalization. The comparison of confusion matrices is crucial for understanding the model's performance in different aspects.

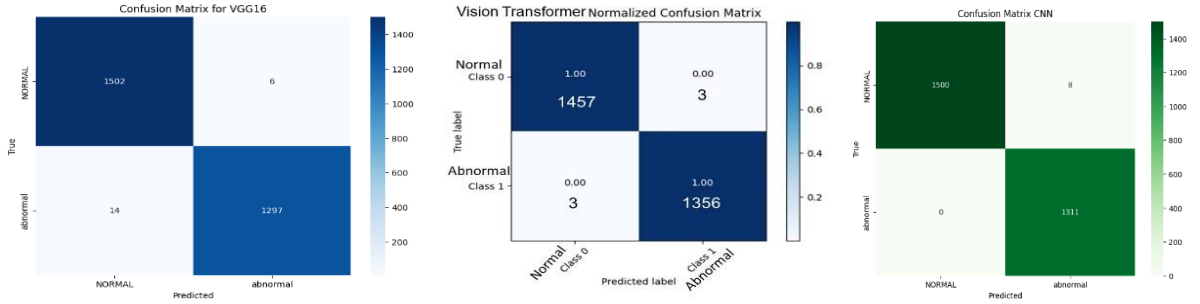


Figure 5: Confusion matrices

As shown in Figure 5, most models have high TP and TN rates indicating strong performance. FP rates are low for most models, suggesting good precision.

5.2.2 PART TWO (View Two): Multi-class dataset

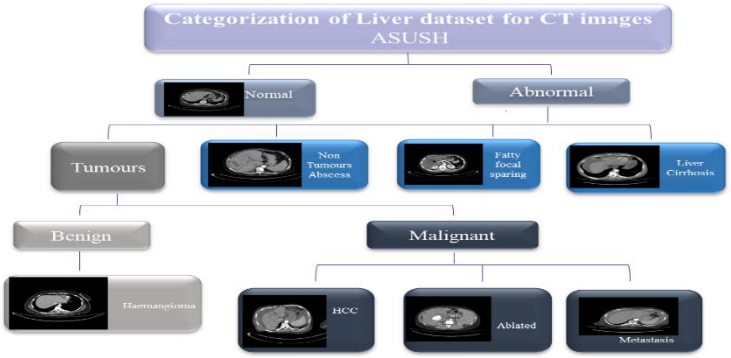


Figure 6: Description of ASUSH dataset categories

Class	Numbers
Normal	1270
Hemangioma	729
Focal Fatty Sparing	438
Cirrhosis	1283
HCC	1888
Ablated	1605
Metastasis	752
Abscess	273
Total	8237

Table 3: ASUSH dataset categorization

This view involves categorizing the ASUSH dataset of liver CT images into 8 categories, which can greatly enhance the utilization of real-world data. They are Normal or Abnormal. Then

further work. The expert radiologists categorize the abnormal patients into Non-Tumor Abscesses, Malignant Metastases, Malignant HCC–RF-Ablated, Malignant HCC, Liver Cirrhosis, Benign Hemangioma, and Focal Fatty Sparing categories. The dataset contains 11,330 CT images. Then prepare data by using data cleaning. Subsequently, the normal class was down-sampled to 1270 CT images as illustrated in Table 3 and Figure 6. Various transfer learning techniques are trained on this dataset. This is especially for improving the diagnosis and treatment of liver diseases. This practical approach allows for the categorization of the hospital dataset, which can greatly enhance the utilization of real-world data.

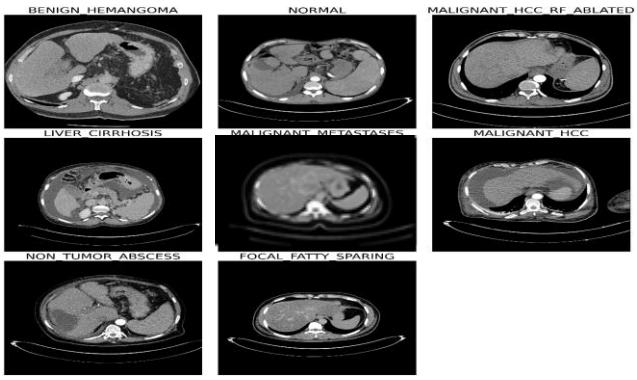


Figure 7: Sample of ASUSH dataset

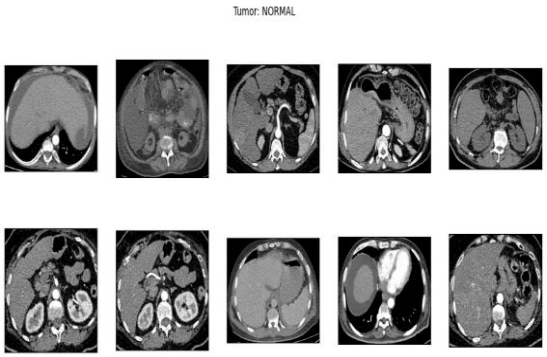


Figure 8: Sample of liver CT cropping Normal images

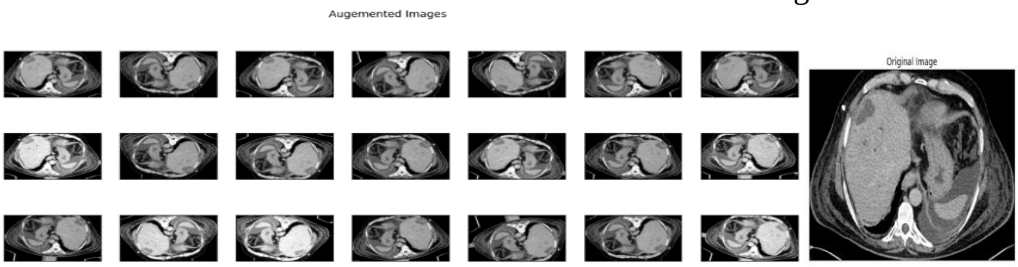
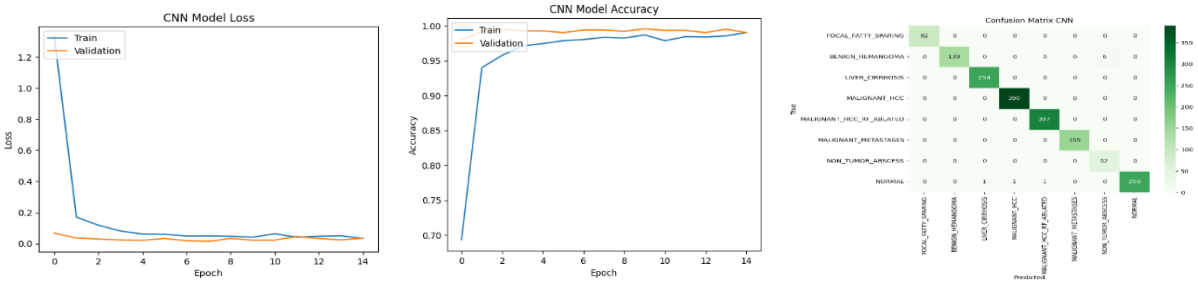
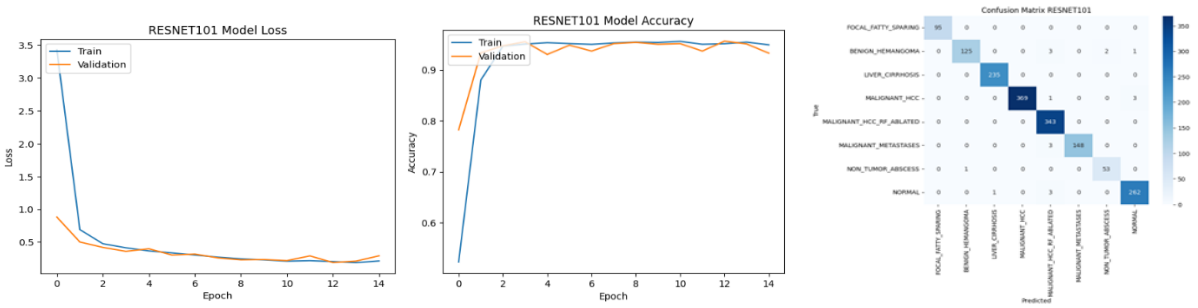


Figure 9: The process of augmenting images



(a) Results of CNN model



(b) Results of ResNet101 model

Figure 10: model's results

Table of results views two (Normal or Abnormal) into 8 subclasses of CT images						
MODEL	Training Accuracy	Testing Accuracy	Precision	Recall	F1-Score	ROC/AUC
VGG16	0.998027	0.992718	0.992783	0.992718	0.992683	0.98
VGG19	0.998634	0.995146	0.995372	0.995146	0.995174	1.00
Vision transformer	0.995752	0.9964	1.00	0.99	0.99	1.00
Xception	0.995143	0.988471	0.988972	0.988471	0.988479	0.99
CNN	0.994840	0.996966	0.997002	0.996966	0.996964	1.00
ResNet50	0.993474	0.988471	0.989106	0.988471	0.988485	1.00
ResNet101	0.957505	0.959951	0.965154	0.959951	0.959982	1.00
EfficientNetB1	0.994081	0.992718	0.992852	0.992718	0.992716	1.00
EfficientNetB2	0.994536	0.989078	0.989174	0.989078	0.989009	1.00
InceptionV3	0.994385	0.988471	0.988743	0.988471	0.988473	1.00
InceptionResNetV2	0.826681	0.779733	0.831807	0.779733	0.768353	0.71

Table 4: Experimental results for viewing two

Figures 7, 8, and 9 visualize and explore the contents of the ASUSH dataset. By plotting one image from each subclass, it gives you a quick glimpse into the different categories of data. The ASUSH CT grayscale images hold significance as they effectively reduce the computational burden of the algorithm. Figure 10 indicates the learning process of each model, showing how well they are learning to generalize from the training data to unseen validation data. The accuracy of each model is quite high, suggesting effective learning capabilities for the tasks on which they are trained. Table 4 shows that models like VGG19, ViT, and CNN is particularly effective for complex classification tasks within medical imaging. They can distinguish between multiple subclasses with high accuracy and reliability. However, the performance of InceptionResNetV2 highlights the variability in model suitability depending on the specificity and complexity of the classification task.

6. Conclusion and Future Work

This paper addressed the critical challenge of liver cancer, a rapidly growing global threat and a leading cause of mortality in Egypt. The paper made several key contributions to support accurate diagnosis and improve patient outcomes. Recognizing the oncologists and radiologists face in detecting and treating treatment of liver tumors, deep learning has recently been explored as a tool to aid in these efforts. As a result, patients can get accurate treatment at the right time with suitable costs. The dataset was meticulously collected from Ain Shams

Specialized Hospital, focusing primarily on CT liver images. The research was bifurcated into two distinct experimental approaches: binary classification and multi-class classification. In the binary classification, patients were categorized as either normal or abnormal. Further, the multi-class is categorized into eight categories: normal, non-tumor abscess, malignant metastases, malignant HCC RF-ablated, malignant HCC, liver cirrhosis, benign hemangioma, and focal fatty sparing. However, it's worth noting that, to our latest knowledge, this dataset is the first multi-class liver cancer dataset in Egypt, making a significant contribution to the field. A variety of sophisticated and novel deep-learning models were employed in this study. The result was a remarkable achievement of over 99% accuracy, precision, recall, F1-Score, and ROC/AUC in classification, and categorization tasks. The reported accuracy of 99% is based on testing data, not training data. There's no evidence that there's overfitting since we tested the model on unseen data and there's no gap between training and testing accuracy results. The implications of these findings are significant: with enhanced detection accuracy, patients can receive timely and precise treatments. This research not only demonstrates the viability of deep-learning models in medical imaging analysis but also paves the way for future innovations in the field.

References

Abdelaziz Ismael, S. A., Mohammed, A., & Hefny, H. (2020). An enhanced deep learning approach for brain cancer MRI images classification using residual networks. Artificial Intelligence in Medicine, 102, 101779 https://www.sciencedirect.com/science/article/pii/S0933365719306177

Abdelhay, M., Mohammed, A., & Hefny, H. A. (2023). Deep learning for Arabic healthcare. Social Network Analysis and Mining, 13(1), 71 https://doi.org/10.1007/s13278-023-00892-1

Alabrak, M. M. A., Megahed, M., Alkhouly, A. A., Mohammed, A., Elfandy, H., Tahoun, N., & Ismail, H. A. R. (2023). Artificial intelligence role in subclassifying cytology of thyroid follicular neoplasm. Asian Pacific Journal of Cancer Prevention: APJCP, 24(4), 1379. https://doi.org/10.31557/APJCP.2023.24.4.1379

Amin, M., Hefny, H., & Mohammed, A. (2021). Sign language gloss translation using deep learning models. International Journal of Advanced Computer Science and Applications, 12(11). https://dx.doi.org/10.14569/IJACSA.2021.0121178

Baykal, E., Dogan, H., Ercin, M. E., Ersoz, S., & Ekinci, M. (2020). Transfer learning with pre-trained deep convolutional neural networks for serous cell classification. Multimedia Tools and Applications,79, 15593–15611. https://doi.org/10.1007/s11042-020-09017-0

Biswas, M., Kuppili, V., Edla, D. R., Suri, H. S., Saba, L., Marinho, R. T., Sanches, J. M., & Suri, J. S. (2018). Symtosis: A liver ultrasound tissue characterization and risk stratification in an optimized deep learning paradigm. Computer Methods and Programs in Biomedicine, 155, 165–177.

Byra, M., Styczynski, G., Szmigielski, C., Kalinowski, P., Michalowski, L., Paluszkiewicz, R., Ziarkiewicz-Wróblewska, B., Zieniewicz, K., Sobieraj, P., & Nowicki, A. (2018). Transfer learning with deep convolutional neural networks for liver steatosis assessment in ultrasound images. International Journal of Computer Assisted Radiology and Surgery,13,1895-1903.

Chaffer, C. L., & Weinberg, R. A. (2011). A perspective on cancer cell metastasis. Science, 331(6024), 1559–1564 https://doi.org/10.1126/science.1203543.

Cheng, P. M., Montagnon, E., Yamashita, R., Pan, I., Cadrin-Chenevert, A., Perdigón Romero, F., Chartrand, G., Kadoury, S., & Tang, A. (2021). Deep learning: an update for radiologists. Radiographics, 41(5), 1427–1445.

Constantinescu, E. C., Udris, A.-L., Udris, C., Iacob, A. V., Gruionu, L. G., Gruionu, G., Sandulescu, L., & Săftoiu, A. (2021). Transfer learning with pre-trained deep convolutional neural networks for the automatic assessment of liver steatosis in ultrasound images. Medical Ultrasonography, 23(2), 135–139.

El-Zayadi, A.-R., Badran, H. M., Barakat, E. M., Attia, M. E.-D., Shawky, S., Mohamed, M. K., Selim, O., & Saeid, A. (2005). Hepatocellular carcinoma in Egypt: A single center study over a decade. World Journal of Gastroenterology: WJG, 11(33), 5193.

Fawcett, T. (2006). Introduction to receiver operator curves. Pattern Recognit. Lett, 27, 861– 874.

Gabruseva, T., Poplavskiy, D., & Alnin, A. (2020). Deep learning for automatic pneumonia detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 350-351).

Hamm, C. A., Wang, C. J., Savic, L. J., Ferrante, M., Schobert, I., Schlachter, T., Lin, M., Duncan, J. S., Weinreb, J. C., Chapiro, J., & et al. (2019). Deep learning for liver tumour diagnosis part I: Development of a convolutional neural network classifier for multi- phasic MRI.European Radiology, 29, 3338–3347.

Ibrahim, A. S., Khaled, H. M., Mikhail, N. N., Baraka, H., Kamel, H., & et al. (2014). Cancer incidence in Egypt: Results of the national population-based cancer registry program. Journal of Cancer Epidemiology, 2014. https://doi.org/10.1155/2014/437971

Jiang, X., Hu, Z., Wang, S., & Zhang, Y. (2023). Deep learning for medical image-based cancer diagnosis. Cancers, 15(14), 3608. https://doi.org/10.3390/ncers15143608

Kader, E., Andronikou, S., & Welman, C. (2001). Computed tomography appearances of focal and diffuse hepatic disease in children. South African Journal of Radiology,5(1), 23– 28.

Kora, R., & Mohammed, A. (2023). An enhanced approach for sentiment analysis based on meta-ensemble deep learning. Social Network Analysis and Mining, 13(1), 38. https://doi.org/10.1007/s13278-023-00888-x

Makram, M., Ali, N., & Mohammed, A. (2022). Machine learning approach for diagnosis of heart diseases. In 2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC) (pp. 69-74). IEEE.

Makram, M., Elhemeily, M., & Mohammed, A. (2023). Deep learning approach for liver tumor diagnosis. In 2023 Intelligent Methods, Systems, and Applications (IMSA) (pp. 210-216). IEEE.

Marrero, J. A., Fontana, R. J., Fu, S., Conjeevaram, H. S., Su, G. L., & Lok, A. S. (2005). Alcohol, tobacco and obesity are synergistic risk factors for hepatocellular carcinoma. Journal of Hepatology, 42(2), 218–224.

Mansour, A. E., Mohammed, A., Elsayed, H. A. E. A., & Elramly, S. (2022). Spatial-net for human-object interaction detection. *IEEE Access*, *10*, 88920-88931.

Mazurowski, M. A., Buda, M., Saha, A., & Bashir, M. R. (2019). Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI. Journal of Magnetic Resonance Imaging, 49(4), 939-954.

Mishra, A., & Mohapatro, M. (2019). An IoT framework for bio-medical sensor data acquisition and machine learning for early detection. International Journal of Advanced Technology and Engineering Exploration, 6(54), 112–125.

Munir, K., Elahi, H., Ayub, A., Frezza, F., & Rizzi, A. (2019). Cancer diagnosis using deep learning: a bibliographic review. Cancers, 11(9), 1235.

Murphy, A., Dabirifar, S., & Bell, D. (n.d.). CT triple-phase liver (protocol). Radiopaedia.org. Retrieved October 26, 2023, from https://doi.org/10.53347/rID-94730

Nemane, R., Thakare, A., Pillai, S., Shiturkar, N., & Nair, A.(2022). Comparative analysis of intelligent learning techniques for diagnosis of liver tumour from CT images. In Advances in Data Computing, Communication and Security: Proceedings of I3CS2021 (pp. 27–37). Springer.

Nishida, N., Yamakawa, M., Shiina, T., Mekada, Y., Nishida, M., Sakamoto, N., Nishimura, T., Iijima, H., Hirai, T., & Takahashi, K. et al. (2022). Artificial intelligence (AI) models for the ultrasonographic diagnosis of liver tumours and comparison of diagnostic accuracies between AI and human experts. Journal of Gastroenterology, 57(4), 309–321.

Olczak, J., Pavlopoulos, J., Prijs, J., Ijpma, F. F. A., Doornberg, J. N., Lundström, C., Hedlund, J., & Gordon, M. (2021). Presenting artificial intelligence, deep learning, and machine learning studies to clinicians and healthcare stakeholders: An introductory reference with a guideline and a clinical AI research (CAIR) checklist proposal. Acta Orthopaedica, 92(5), 513–525. https://doi.org/10.1080/17453674.2021.1918389

Purwar, A., & Singh, S. K. (2015). Hybrid prediction model with missing value imputation for medical data. Expert Systems with Applications, 42(13), 5621–5631. https://doi.org/10.1016/j.eswa.2015.02.059

Rana, M., & Bhushan, M. (2023). Machine learning and deep learning approach for medical image analysis: diagnosis to detection. Multimed Tools Appl, 82, 26731–26769. https://doi.org/10.1007/s11042-022-14305-w

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III (pp. 234–241). Springer.

Shafiq, M., & Gu, Z. (2022). Deep residual learning for image recognition: a survey. Applied Sciences, 12(18), 8972.

Shaked, O., Siegelman, E. S., Olthoff, K., & Reddy, K. R. (2011). Biologic and clinical features of benign solid and cystic lesions of the liver. Clin Gastroenterol Hepatol, 9(7), 547-62. e1-4. doi: 10.1016/j.cgh.2011.03.007. Epub 2011 Mar 11. PMID: 21397723

Sharma, A., & Nagalli S. Chronic Liver Disease (2023). In StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK554597/

Sinthuja, U., Hatti, V., & Thavamani, S. (2022). Analysis and prediction of liver disease for the patients in India using various machine Learning algorithms. in Advances in Data Computing, Communication and Security: Proceedings of I3CS2021.Springer, 022, pp. 433–442.

Sween, S. Samar, C. and Binu, S, "Triple-phase mdct of liver: Scan protocol modification to obtain optimal vascular and lesional contrast," *Indian Journal of Radiology and Imaging*, vol. 28, no. 03, pp. 315–319, 2018. Sween, S., Samar, C., & Binu, S. (2018). Triple-phase mdct of liver: Scan protocol modification to obtain optimal vascular and lesional contrast. *Indian Journal of Radiology and Imaging, 28*(3), 315–319.

Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In International Conference on Machine Learning. PMLR (pp. 6105–6114).

Tang, X., Jafargholi Rangraz, E., Coudyzer, W., Bertels, J., Robben, D., Schramm, G., Deckers, W., Maleux, G., Baete, K., Verslype, C., et al. (2020). Whole liver segmentation based on deep learning and manual adjustment for clinical use in SIRT. *European Journal of Nuclear Medicine and Molecular Imaging, 47*, 2742–2752.

Wang, W., Yang, Y., Wang, X., Wang, W., & Li, J. (2019). Development of convolutional neural network and its application in image classification: A survey. *Optical Engineering, 58*(4), 040901.

Wang, S., Dong, L., Wang, X., & Wang, X. (2020). Classification of pathological types of lung cancer from CT images by deep residual neural networks with transfer learning strategy. *Open Medicine, 15*(1), 190–197.

Young, M. et al. (2002). Technical writer's handbook. University Science Books.

Yuan, L., Chen, Y., Wang, T., Yu, W., Shi, Y., Jiang, Z.-H., Tay, F. E., Feng, J., & Yan, S. (2021). Tokens-to-token ViT: Training vision transformers from scratch on ImageNet. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 558–567).

Zhou, L.-Q., Wang, J.-Y., Yu, S.-Y., Wu, G.-G., Wei, Q., Deng, Y.-B., Wu, X.-L., Cui, X.- W., & Dietrich, C. F. (2019). Artificial intelligence in medical imaging of the liver. World Journal of Gastroenterology, 25(6), 672.

24. Examining The Influence of Explainable Artificial Intelligence on Stakeholder Decision-Making Processes in an Audit Firm

Sinthuri Konar
University of Cape Town
knrsin001@myuct.ac.za

Grant Oosterwyk
University of Cape Town
grant.oosterwyk@uct.ac.za

Pitso Tsibolane
University of Cape Town
pitso.tsibolane@uct.ac.za

Abstract

Decision-making, a complex process central to organizational success, is the focus of increasing attention within the field of Information Systems (IS) research, particularly in the context of Explainable Artificial Intelligence (XAI). Grounded in General Systems Theory (GST), this study examines the influence of XAI on decision-making processes within a South African audit firm, where the ability to understand and explain AI-driven decisions is imperative for meeting regulatory requirements and gaining stakeholders' trust. The research employed a single case study methodology with semi-structured interviews as the primary means of data collection. Thematic analysis was used to facilitate the analysis process by grouping the findings into main themes. The study revealed how various socio-technical aspects influence the utilization of XAI in decision-making processes and are crucial components in generating business value, information quality, and insights. The study also presents a conceptual model that illustrates the relationship between XAI and stakeholder decision-making.

***Keywords**:* Explainable Artificial Intelligence (XAI), Artificial Intelligence (AI), audit firm, decision-making processes, Stakeholders.

1. Introduction

The growth in Artificial Intelligence (AI) capabilities, driven by advanced techniques (Mikalef & Gupta, 2021), has positioned AI as a significant value driver for companies. The evolution of advanced AI technologies necessitates strategic integration by organizations across diverse sectors to effectively improve decision-making processes operationally and strategically (Haque et al., 2023). This is particularly relevant in South Africa, where organizations are actively expanding their research efforts in the field of AI (Taljaard & Gerber, 2022). As the landscape of AI adoption evolves, it becomes increasingly essential for companies to harness AI to achieve outcomes that not only optimize operational efficiencies but also contribute to strategic decision-making across various domains. Decisions made by autonomous computational algorithms often have severe consequences for both individuals and organizations as they become more invisible, thereby making it more difficult for decision-makers to understand and explain AI models (Das and Rad, 2020). The lack of explainability and transparency of these models can become problematic in real-world decision contexts

because the results of such analyses are never 100% accurate despite constant technological improvements (Wheeler and Carroll, 2023). Therefore, existing models used in various AI systems do not reveal enough information about the processes and methods that lead to a conclusion (Ali et al., 2023). Explainable Artificial Intelligence (XAI) has emerged, intending to present explanations purveyed to human understanding, trust, and transparency (Haque, Mikalef, and Islam, 2023). XAI refers to the process of human users understanding how an AI system decides, predicts, and executes its operations to make complex AI models more transparent by providing users with understandable explanations (Meske et al., 2022). Explainability in XAI refers to the ability of the AI system to give the users a deeper methodological explanation of predictions derived from an AI model (Langer et al., 2021). The design of XAI serves as a bridge to provide stakeholders with an understanding of AI's outcomes, results, and even behaviour, thereby allowing the stakeholders to use what AI has to offer to produce more informed and conscious decisions (Wang and Yin, 2022). Research is needed to include the exploration of XAI among various professions and users of systems (Haque, Islam and Mikalef, 2023). Research on XAI for decision-making tends to either be generic or predominantly focused on medical use cases (Sahoh & Choksuriwong, 2023; Scholar et al., 2021). AI presents various challenges for South African audit firms that need to be resolved; these difficulties include ethical and understandability issues with AI amongst employees and clients, which gives rise to moral concerns regarding privacy, security, bias, fairness, and accountability of the outputs of AI models (Taljaard & Gerber, 2022). Hence, there is a need to examine the influence of XAI from the perspectives of stakeholders in the context of auditing firms. Consequently, the research objective examines how XAI influences stakeholders' decision-making processes in an audit firm. To address this research objective, we have collected empirical data using semi-structured interviews in this study. The study aims to examine the influence of utilising XAI on decision-making processes among stakeholders within an organization. Moreover, the research sought to capture the technology (i.e., XAI) and social (i.e., decision-making) elements drawing from the GST, modelled as an IS artefact with information emanating from the interaction between the two elements (Chatterjee et al., 2021; Sarker et al., 2019). We ask the following research question: *"How does XAI influence the stakeholder decision-making processes in an audit firm?"* The paper is structured as follows. Firstly, we introduce a narrative literature survey of XAI and decision-making processes. Secondly, we present the IS model representing a GST-informed conception of XAI and decision-making artefact. Finally, we present the findings and implications for future research.

2. Literature Review

This section utilizes a narrative literature review, which takes an adaptable approach to mapping and expanding on the existing state of knowledge of XAI and decision-making processes. This is accomplished by identifying potential gaps in prior XAI-related research. A narrative review is ideal for presenting findings from published journals and studies as it allows the researcher to summarize and analyze the evidence regarding the state of knowledge currently available on the relevant topic (Templier and Paré, 2018). This narrative review reports the repute of the sources used when searching for and selecting publications to be analyzed. Firstly, a search was conducted using the Web of Science and Scopus databases. The IS Senior Scholar basket of journals was used to identify relevant papers. The search process began with seminal articles on the topic of interest and limitations within the inclusion and exclusion were performed (Günther et al., 2017). Figure 1 depicts the complete screening and

selection process, in which journals were searched on Title and Topic with the key search terms relating to the research question, and during the timespan from 2019 to 2023. Specifically, the authors decided only to use Scopus and Web of Science databases to search for affiliated IS journals, and the leading four IS conferences and HICCS were further selected from the AISeL (IS electronic library). After reading the titles and abstracts, 93 papers were identified. However, many of these papers were discovered to lack empirical evidence or papers that were not within the scope of the research theme. Hence, these papers were excluded. To end, the remaining articles were examined based on their abstracts and main objectives in relation to whether they offered substantive literature surrounding XAI and stakeholders in decision-making processes. Full texts of the remaining 51 papers were thoroughly reviewed, selecting only 42 papers that were deemed sufficient to be further analysed.

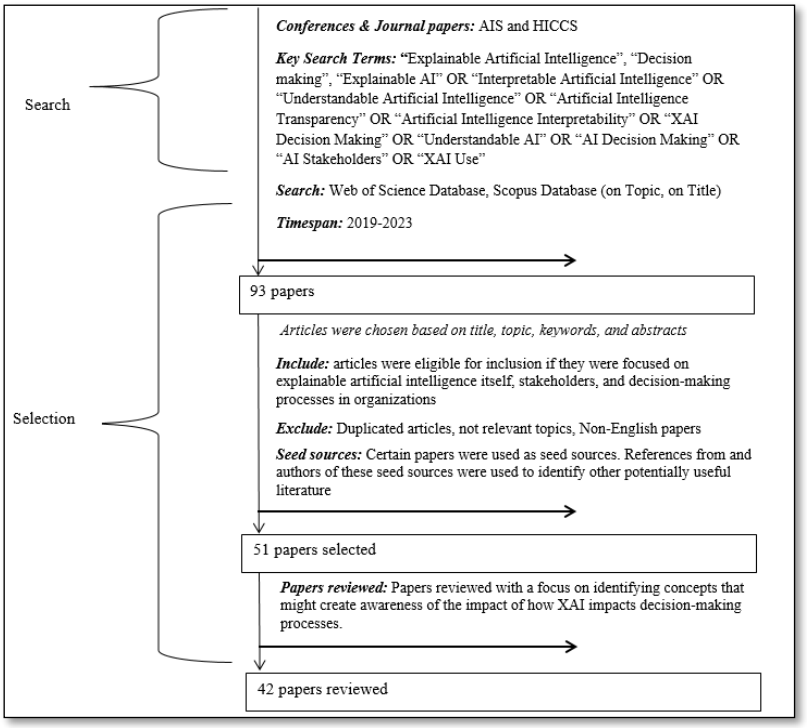


Figure 1: Search and selection process

2.1 Explainable Artificial Intelligence

Understanding AI systems becomes challenging when they are too complex or inherently opaque (Chaddad et al., 2023). Explainable Artificial Intelligence (XAI) has emerged to address this challenge by providing explanations for trust, human understanding, and transparency (Eisbach, Langer and Hertel, 2023). XAI aims to uncover the reasoning behind AI predictions and present this information in an understandable manner through user interfaces (Scholar, Chakrobartty and El-Gayar, 2021). By making AI outputs interpretable to end users and stakeholders, XAI increases transparency without sacrificing performance (Clement et al., 2023). Operational efficiency of AI systems is enhanced when they offer meaningful information alongside outputs to affected individuals (Meske and Bunde, 2020), facilitating quicker and more accurate decision-making (Eisbach, Langer and Hertel, 2023).

However, challenges such as trust, ethics, skills, and biases in decision-making processes hinder XAI utilization (Ding et al., 2022). Tailoring explanations to stakeholders' characteristics and interests enhances system utility, particularly in high-stake decision-making scenarios (Haque et al., 2023; Dwivedi et al., 2023). Decision trees are common XAI visualization techniques, offering insights into how features influence AI decisions (Gerlings, Shollo and Constantiou, 2020). XAI strategies, including model-specific and model-agnostic approaches, provide local explanations for specific decisions, either integrating interpretability into learning models or using them to explain outputs (Meske et al., 2022; Gerlach et al., 2022).

2.2 Stakeholder Decision-making

As defined by Shibayama et al. (2020), stakeholders exert influence on organizational goals and are influenced by them. These groups, identified by Meske et al. (2022), directly impact project implementation. Understanding stakeholders is crucial due to their diversity and competing objectives (Carbone et al., 2022). Decision-making, selecting the best alternatives from options, involves complex mathematical reasoning (Sharma, Mithas and Kankanhalli, 2014), typically grounded in the phase theorem (Mintzberg, Raisinghani and Théorêt, 1976). This theory suggests that decisions follow a systematic process of gathering information, assessing options, and committing to alternatives (Blohm, 2019), involving iterative data-processing flows. Decision-making models, such as those discussed by Turpin and Marais (2004), transform needs and goals into desired outcomes, guiding complex decisions and enhancing organizational performance (Devi, Nayak and Patnaik, 2020). Stakeholder involvement in decision-making is increasingly vital (Le Pira et al., 2017), contributing to rational and intuitive processes. The hierarchical nature of stakeholder engagement involves multiple levels (Carbone et al., 2022), fostering shared decision-making and consensus through stakeholder interaction (Edelenbos and Klijn, 2006).

2.3 Challenges and Benefits Associated with XAI Use in Decision Making

Trust in AI models is influenced by factors like accuracy, reliability, consistency, and fairness of results (Kaur et al., 2023). Explainable AI (XAI) enhances trust by offering clear explanations for users to comprehend, confirm, and verify results (Kaur et al., 2023). However, unclear or incomplete explanations can undermine trust and affect decision-making (Ding et al., 2022 Bias may be increased or introduced through the explanations produced by XAI, as they may be subjective or due to the influence of human cognitive biases. The inexplicability of AI decision results may increase the likelihood of biases, as human decision-makers may struggle to recognize errors in the data or the model that affect the AI system's decisions, thereby causing limitations or discrepancies (Albahri et al., 2023). Furthermore, ethical and legal concerns arise regarding privacy, accountability, and fairness in decision-making processes (Saeed and Omlin, 2023). Accountability may be compromised as human decision-makers struggle to monitor AI behaviour or justify decisions based on AI output (Ding et al., 2022), leading to uncertainty or conflict over responsibility (Antoniadi et al., 2021). Despite challenges, XAI helps stakeholders understand AI systems, enhancing trust and decision-making by providing clear justifications (Barredo Arrieta et al., 2020; Meske and Bunde, 2020). However, the complexity of explanations may overwhelm users, affecting their skills (Carvalho & Silva, 2021). Customizable explanations accommodate diverse user preferences, enhancing decision-making efficiency (Das and Rad, 2020; Praveen and Joshi, 2023). Thus, XAI

facilitates quicker, more accurate decision-making and improves business efficiencies (Pesecan and Stoicu-Tivadar, 2023).

3. Theoretical Background

The IS artefact model emphasizes the interplay between technology and social subsystems, which is crucial for system behaviour. Applying general systems theory (GST), the symbiotic relationship between these subsystems characterizes the interaction between XAI and decision-making as an IS Artifact. To illustrate the socio-technical elements influencing decision-making, a paradigm depicted in Figure 2 is utilized (Chatterjee and Bhattacharjee, 2020). Permeable subsystem boundaries facilitate resource exchange, fostering synergism (Sarker et al., 2019). Decision-making emerges as a superordinate system shaped by social and technical subsystem interactions (Sarker et al., 2019). Stakeholders, such as AI managers, represent the social subsystem, while XAI methods constitute the technical subsystem, enhancing stakeholder understanding of AI outputs. The feedback loops maintain system stability and establish transformation (Chatterjee et al., 2021; Sarker et al., 2019).

The efficacy of the decision is not entirely dependent on the outputs of XAI but also on the decision maker's knowledge, which is "promoted by the permeable boundaries of the subsystems" Chatterjee et al., 2020). A South African audit firm is a case study chosen for its suitability in examining XAI's influence on stakeholder decision-making within the audit context, employing semi-structured interviews for data collection.

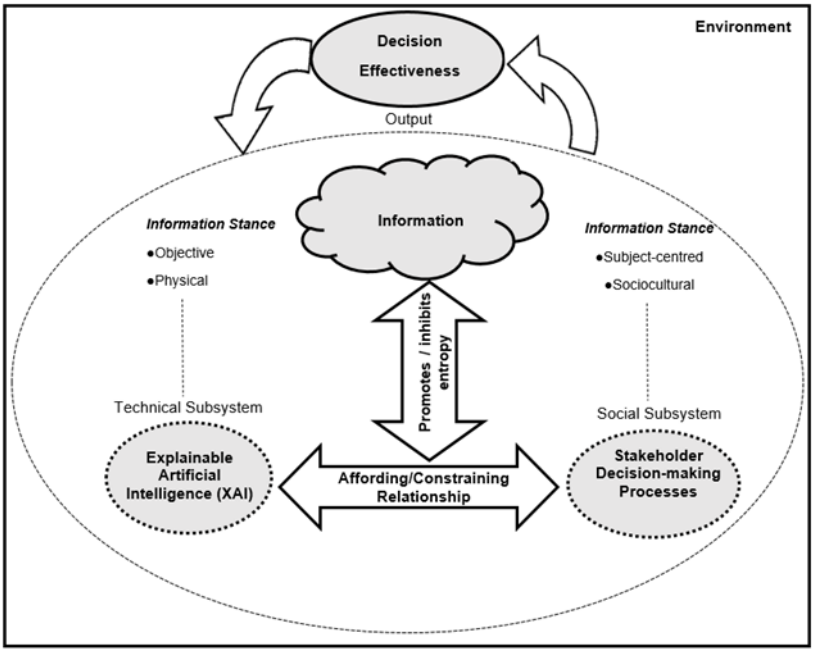


Figure 2: A GST-informed conception of XAI and decision-making artefact (Source: Chatterjee et al., 2021; Sarker et al., 2019)

4. Research Methodology

Firstly, South Africa's dynamic economic landscape, characterized by multiple industries and regulatory environments, presents a rich and multifaceted setting for studying complex interactions between XAI and stakeholder decision-making. Secondly, the regulatory framework in South Africa, including the Companies Act and International Financial Reporting Standards (IFRS), imposes stringent requirements on audit practices, making it imperative for audit firms to operate in mixed compliance regimes. Qualitative semi-structured interviews fit well with the interpretive approach and enabled the researcher to acquire an understanding of the participants' perspectives on the research subject (Braun and Clarke, 2006). An interpretivism philosophy is adopted for this study to analyse the phenomenon through the subjective interpretation of interviewing individuals to gain insights from the data collected (Pervin and Mokhtar, 2022). Before conducting the interviews, we collected explicit consent from the participants regarding the purpose of the interview and how the data will be processed. We did not use any personal data while analyzing the interview data. The sample size consisted of 15 participants who agreed to be interviewed with the aim of answering the study's research question. The participants were selected as IT professional experts who are well-versed in AI tools or who are leading AI-driven initiatives within the audit firm. The participants' work experience spans 3 to 17 years in their respective fields, with professions ranging from intermediate to senior roles, such as consultants to senior managers. Ultimately, this provided a sample set from a diverse and knowledgeable target population, which ensured diverse views were obtained; the participants were a mix of different job roles across the organisation. The complete list of the participants and their respective roles are presented in Table 1.

Participant ID	**Job Title**	**Years of experience**
INT1	Manager in Digital Cognitive Advantage	8
INT2	Senior Consultant in AI and Risk Advisory	8
INT3	Manager in Data Science and engineering	6
INT4	Manager in AI engineering	6
INT5	Manager in Data Science and Machine Learning	4
INT6	Manager in Data Analytics	4
INT7	Senior Consultant in Digital engineering	6
INT8	Manager in Business intelligence	8
INT9	Senior Consultant in AI & Data analytics	3
INT10	Manager in AI and Data	10
INT11	Head of the Intelligent Automation Capability	17
INT12	Senior Data Migration Consultant	5
INT13	Associate Director in AI and Data analytics	12
INT14	Customer Analytics Leader	7
INT15	Data Science Consultant	4

Table 1: Interview Participants

All interviews were conducted virtually using Microsoft Teams. After collecting sufficient data, the interview sessions were transcribed using Microsoft Teams and saved to a secure cloud storage location. Qualitative data collected from the interviews was exported to Microsoft Word for cleaning and uploaded to NVivo software for analysis. This study adopted the thematic analysis approach to facilitate the data analysis process (Thompson, 2022). First-level codes (direct quotes) were derived from interview data, and arranged into categories and themes, guided by their shared characteristics and associations. Building on the legacy of the sociotechnical perspective given the nature of XAI, the themes were grouped between the social and technical components of the phenomenon to understand the mutual interaction and influence of information quality on decision-making processes in the context of XAI. These preliminary themes were further refined to ensure they precisely captured the data they originated from and were well-defined, unique, and pertinent to the research goals. Lastly, these themes were characterized and given names to offer a succinct and clear explanation of their content. A link to the interview questions has been provided as part of the appendix section.

5. Findings

To address the primary research question, this study followed the socio-technical tradition by investigating how XAI (i.e., the technical perspective) influences the decision-making processes (i.e., the social perspective) to produce information (such as insights) amongst stakeholders (Chatterjee et al., 2021).

5.1 Social Aspects

Trust is a central prevalent social theme emerging from the empirical data, which is the challenge to the utilization of XAI for decision-making. Participant 11, the head of the intelligent automation capability, described an exemplary case of such a challenge whereby he explained: "*I think with every new technology being implemented or adopted, there's always going to be some initial feeling issues, which obviously will not get people to trust it. But over a period of time, I do believe that we will be able to kind of trust and accept this technology*" [INT11]. Furthermore, Participant 13, an associate director in AI & Data analytics, for example, described: "*there's quite a lot of pushback from the general people in the organizations because they feel like AI will take over their jobs*" [INT13]. Participants, including AI and data managers in senior roles, expressed concerns about being subject to cognitive **bias** in their decision-making processes stemming from bias and legal implications of the explanations produced and characterized by the AI models. Participant 10, a manager in AI & Data highlights the criticality of this theme as explained: "*The questions around trust, explainability, bias, fairness, robustness, security and so forth are all super important considerations when we decide what we do like at the start of the project, such as who is affected by the solutions, and we need to think about what variables do we want to use*" [INT10]. Participant 11, the head of the intelligent automation capability shared positive sentiments related to the social benefit of the influence of XAI on their **confidence** levels in upskilling and learning. He explained: "*Personally, I would try and understand well how did AI derive at that output, and why could I not have thought about it in that manner. So, like I said, it's an opportunity for me to look at things differently and obviously learn at the same time*" [INT11]. A social construct of **business support** emerged as the prominent enabler for the use of XAI in providing **business insights** within decision-making processes. Participant 5, a manager in data science and machine learning, explained: "*I think so far it can help in little*

things like improving certain aspects of the way we work and maybe allowing us to get some insights from data" [INT5]. The issue of **ethical concerns** and considerations were raised by how stakeholders need to ensure that they maintain their professional scepticism and independence in order to avoid over-reliance or under-reliance on the technology due to the lack of human judgment associated with XAI. Participant 11, expressed his concerns: "*In terms of AI, I guess the market is moving quite rapidly with this technology because of the value it can unlock with the sort of con that comes in is the trust OK, so people do not have that trust right now with this tool, and because AI does not have any emotions or judgments when it makes any sort of decisions as well in comparison to humans, so that element is not there, which then poses a bit of an ethical sort of question as well*" [INT11]. When an organization optimizes its firm-level resources and adopts XAI, economic value emerges in terms of providing **transparency**. In contexts with highly competent IT professionals, trust-building between the client and stakeholder should be encouraged as a measure to promote transparency and trust, for example, Participant 10, a manager in AI and data, noted: "*you try and explain to the clients how the model is working, how it's arriving at, what it is arriving at and you know try to make it as transparent as possible because otherwise, we've seen like the lot of issues from an adoption perspective, right*" [INT10]. Understanding the explanations as to why AI systems make certain choices and what are the sources of **uncertainty** or error in AI systems emerged as one of the significant influences of XAI on stakeholder decision-making processes; this is a social construct. Participant 14, a customer analytics leader explained: "*It's critical for people to understand you know how more or less these types of solutions work and why something is being outputted as what it is right. XAI reduces the uncertainty that decision-makers have to deal with when they have to make really important decisions*" [INT14].

5.2 Technical Aspects

A significant social barrier to utilizing XAI in decision-making was **skills**. The empirical data suggested how the lack of skills among stakeholders may impede utilizing the XAI's full capabilities for decision-making. Participant 1, a manager in the digital cognitive advantage offering, notes that "*people obviously required a deep understanding of the AI tool itself and the mathematics behind the model itself, because if you don't know that you can't use XAI effectively when making a decision*" [INT1]. The sentiment that stakeholders still shy away from taking responsibility for decisions taken with XAI due to fear and not having the skills to understand other participants further affirms AI. Participant 13 referred to such limitations as "*I think people are afraid of it because they don't understand just how it works and because of when they don't understand how it works, they naturally are scared to use it*" [INT13]. The empirical data demonstrated the **technical complexities** of utilizing XAI as a prevalent technical construct that emerges when using XAI. Participant 1, a manager in digital cognitive advantage, described how most people find it difficult to comprehend the explanations produced and derived from complex AI models. He explained that "*as you move further down the spectrum of problem complexity and linearity, you would need to start using more complex, more versatile, and flexible models, and typically with increased complexity becomes less explainability. It'll tell you something, but it's very hard to actually figure out how it got there, especially in the case of deep neural networks*" [INT1]. Understanding the explanations for the AI was noted by Participant 10 as a barrier, he noted that "*one of the challenges that we've seen with AI is the understanding of the people that not everybody's on the same page or understands what exactly is happening or what is expected even off of this solution, whatever*

we are creating for them" [INT10]. The empirical data revealed that **efficiency** as a technical construct is an enabler for using XAI in decision-making. Participant 12, a senior data migration consultant, stressed the importance of how XAI assists in performing tasks more accurately and in a **time-efficient** manner, "*in terms of the output that I have to produce so I'm pretty much dependent on it to provide accurate things like the data, I really depend on that that it's quite vital and it makes my life easier*" [INT12]. In parallel, the participants shared positive sentiments related to the influence of XAI. The derived **business value** was noted by Participant 6, a manager in data analytics, because of less time required to execute tasks. He notes that "*its 100 percent helps with efficiency, XAI helps us to do tasks in a shorter period of time, which enables people to then now focus on other things such as where you do need a lot more, like human sensibility or like trying to ration things out*" [INT6]. Whilst XAI is incorporated into decision-making, the decision outcome is considered the stakeholders' responsibility. Consequently, the need for an adequate understanding of the explainability of AI is crucial through various interpretability methods. Participant 1 explained: "*We'll try and run, you know, multiple linear regressions because that's easy for the general person to understand, and at the end of that, we can get to the same prediction and the value of that is to say this is how the models deconstructed. These are the important parameters that were used in making that consideration*" [INT1]. To make complex models more comprehensible and explainable to both technical and non-technical stakeholders, Participant 14, added that he employs XAI methods to make complex AI analyses and predictions fully transparent and the results of AI more interpretable to both technical and non-technical users. Model-specific XAI and model-agnostic XAI are two major types of XAI strategies for improving explainability. Individual predictions are generated by complex models which can be explained using **interpretability methods**, such as Local Interpretable Model-agnostic Explanations (LIME) (Meske & Bunde, 2020). These strategies provide local explanations, indicating that they explain the reasoning behind the model's judgments for a specific decision (Gerlach et al., 2022). Participant 14 noted: "*there's a predefined approach that you can apply to abstract explainability and one of them is you know what comes to mind is called LIME. LIME as an approach brings forward the variables that were used to create that prediction. It says if you are going to buy this product and these are the 10 reasons why you going to buy the product, which goes a long way in, you know, in helping people with no technical sort of background understand how decisions are being made, and that are the variables that that, that, that are going into it."* [INT14]. Options for XAI-embedded decision-making processes are proposed which highlights how AI and humans interchangeably make decisions within an organization. Gathering a rich **quality of information** from a variety of stakeholders is characterized by different backgrounds, perspectives, and levels of **technical expertise**, thus resulting in a decision-making process characterized by gut feel, rationality, and intuition. Participant 15, a data science consultant, explained: "*Sometimes you make decisions based on like your gut or based on the experience, because we can't really rely on what the machine tells you, you also have to bring in your experience*" [INT15]. Additionally, Participant 10, a manager in AI and Data, described: "*You need to consider the fact that as much as it's a very powerful machine learning algorithm that's giving you the outputs, it has got its own limitations, so you need to factor in those limitations and then consider the expert inputs on that. You need to complement this with the expertise and the experience that people have and they, together with AI, can arrive at better decisions*" [INT10].

6. Proposed Model

A conceptual model was developed and produced according to the insights collected from the thematic analysis to understand better how stakeholders currently perceive the utilisation of XAI for decision-making. Concepts such as the benefits and challenges of XAI, the tools and techniques of XAI, and the influence of the utilization of XAI on decision-making processes were in the form of NVivo codes. Subsequently, these concepts were merged into relevant second-order concepts. Applying the GST lens to these different concepts helps understand the socio-technical aspects of the influence of XAI on decision-making processes. Due to their similarities, several concepts were merged. Figure 2 serves as a graphical representation of a conceptual model that demonstrates the integration of the GST lens, and how the researcher progressed from raw data to the proposed themes as guided by the method.

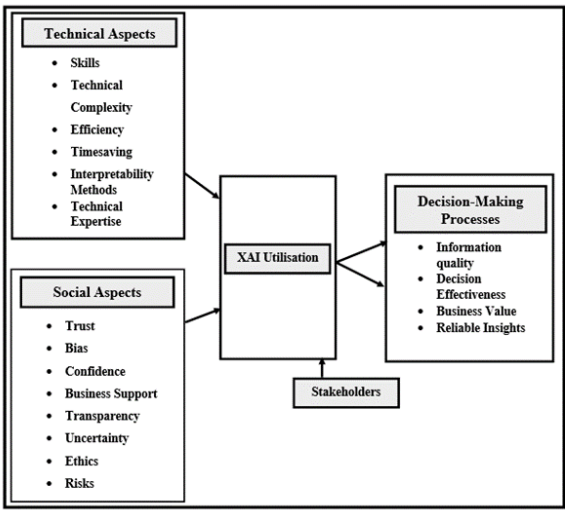


Figure 2: Conceptual Model for the Socio-Technical Influence of XAI on Stakeholder Decision-making

5. Conclusion

The study aimed to examine the influence of the utilization of XAI on stakeholder decision-making processes within a South African audit firm. This research identified the following socio-technical factors influencing using XAI in decision-making processes: trust, technical complexities, efficiency, uncertainty, skills, bias, customer value, transparency, confidence, business support, and ethical concerns. This study confirmed that utilising XAI can improve stakeholders' decision-making performance by providing value through insights, information quality, and efficiency. Furthermore, the results showed how crucial it is for stakeholders to utilize XAI to assist with tasks that need to be completed faster and more accurately. To further foster the utilization of XAI, this research highlighted the roles of technical and social aspects in the increased utilization of XAI, as well as examined the benefits and challenges of the use of XAI for decision-making processes. This study offered a socio-technical perspective in an organizational context, which has remained neglected so far (Sarker et al. 2019). These findings are also of great practical relevance as they point to potential organizational strategies for establishing decisions that are effectively supported by XAI, translating into improved decision

outcomes. Regarding the theoretical contribution, the proposed conceptual model aims to provide guidelines to stakeholders to perceive and derive the underlying benefits and challenges of utilising XAI to establish decisions amongst stakeholders. While an effort was made to explain the term XAI to the participants, some appeared not fully understanding the term well. A lack of knowledge of XAI was evident in how some of the interview questions were answered. The scope of this study was limited to a single audit firm in South Africa. This suggests that the findings of this case study may be case-specific and limited to the contexts of the organization investigated in this study. In future, we recommend investigating the potential challenges and benefits associated with using XAI, particularly in the context of decision-making across various industries and not limited to audit firms.

References

Albahri, A.S. et al. (2023) 'A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion', Information Fusion, 96, pp. 156–191.

Ali, S. et al. (2023) 'Explainable Artificial Intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence', Information Fusion, 99.

Antoniadi, A.M. et al. (2021) 'Current challenges and future opportunities for xai in machine learning-based clinical decision support systems: A systematic review', Applied Sciences (Switzerland), 11(11).

Barredo Arrieta, A. et al. (2020) 'Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI', Information Fusion, 58, pp. 82–115.

Blohm, I. (2019) Patterns of Data-Driven Decision-Making: How Decision-Makers Leverage Crowdsourced Data Competence Center Crowdsourcing View project Leveraraging Data-Driven Organizations View project.

Braun, V. and Clarke, V. (2006) 'Using thematic analysis in psychology', Qualitative Research in Psychology, 3(2), pp. 77–101.

de Bruijn, H., Warnier, M. and Janssen, M. (2022) 'The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making', Government Information Quarterly, 39(2).

Carbone, S. et al. (2022) 'Stakeholder involvement in care transition planning for older adults and the factors guiding their decision-making: A scoping review', BMJ Open, 12(6).

Chaddad, A. et al. (2023) 'Survey of Explainable AI Techniques in Healthcare', Sensors. MDPI.

Chatterjee, S. et al. (2021) 'A possible conceptualization of the information systems (IS) artifact: A general systems theory perspective1', Information Systems Journal, 31(4).

Chatterjee, S. and Bhattacharjee, K.K. (2020) 'Adoption of artificial intelligence in higher education: a quantitative analysis using structural equation modelling', Education and Information Technologies, 25(5), pp. 3443–3463.

Clement, T. et al. (2023) 'XAIR: A Systematic Metareview of Explainable AI (XAI) Aligned to the Software Development Process', Machine Learning and Knowledge Extraction, 5(1), pp. 78–108..

Das, A. and Rad, P. (2020) 'Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey'.

Devi, S., Nayak, M.M. and Patnaik, S. (2020) 'Decision-making models and tools: A critical study', International Journal of Management and Decision Making, 19(2), pp. 176–206.

Ding, W. et al. (2022) 'Explainability of artificial intelligence methods, applications and challenges: A comprehensive survey', Information Sciences, 615, pp. 238–292.

Dwivedi, R. et al. (2023) 'Explainable AI (XAI): Core Ideas, Techniques, and Solutions', ACM Computing Surveys, 55(9), pp. 1–33.

Edelenbos, J. and Klijn, E.-H. (2006) 'Managing stakeholder involvement in decision making: A comparative analysis of six interactive processes in the Netherlands', Journal of Public Administration Research and Theory, 16(3), pp. 417 – 446.

Eisbach, S., Langer, M. and Hertel, G. (2023) 'Optimizing human-AI collaboration: Effects of motivation and accuracy information in AI-supported decision-making', Computers in Human Behavior: Artificial Humans, 1(2), p. 100015.

Gerlach, J. et al. (2022) 'Decision support for efficient XAI services - A morphological analysis, business model archetypes, and a decision tree', Electronic Markets, 32(4), pp. 2139–2158.

Gerlings, J., Shollo, A. and Constantiou, I. (2020) 'Reviewing the need for explainable artificial intelligence (XAI)', in Proceedings of the Annual Hawaii International Conference on System Sciences.
Günther, W.A. et al. (2017) 'Debating big data: A literature review on realizing value from big data', Journal of Strategic Information Systems, 26(3).
Haque, A.B., Islam, A.K.M.N. and Mikalef, P. (2023) 'Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research', Technological Forecasting and Social Change, 186.
Haque, B., Mikalef, P. and Islam, A.K.M.N. (2023) To Explain Or Not To Explain: An Empirical Investigation Of Ai-Based Recommendations On Social Media Platforms.
Kaur, D. et al. (2023) 'Trustworthy Artificial Intelligence: A Review', ACM Computing Surveys, 55(2).
Langer, M. et al. (2021) 'What do we want from Explainable Artificial Intelligence (XAI)? – A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research', Artificial Intelligence, 296.
Meske, C. et al. (2022) 'Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities', Information Systems Management, 39(1).
Meske, C. and Bunde, E. (2020) 'Transparency and trust in human-AI-interaction: The role of model-agnostic explanations in computer vision-based decision support', in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, pp. 54–69.
Mikalef, P. and Gupta, M. (2021) 'Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance', Information and Management, 58(3).
Mintzberg, H., Raisinghani, D. and Théorêt, A. (1976) 'The Structure of "Unstructured" Decision Processes', Administrative Science Quarterly, pp. 246–275.
Nowell, L.S. et al. (2017) 'Thematic Analysis: Striving to Meet the Trustworthiness Criteria', International Journal of Qualitative Methods, 16(1).
Pervin, N. and Mokhtar, M. (2022) 'The Interpretivist Research Paradigm: A Subjective Notion of a Social Context', International Journal of Academic Research in Progressive Education and Development, 11(2).
Pesecan, C.-M. and Stoicu-Tivadar, L. (2023) 'Increasing Trust in AI Using Explainable Artificial Intelligence for Histopathology - An Overview', Studies in Health Technology and Informatics, 305, pp. 14–17.
Le Pira, M. et al. (2017) 'Towards a decision-support procedure to foster stakeholder involvement and acceptability of urban freight transport policies', European Transport Research Review, 9(4).
Praveen, S. and Joshi, K. (2023) 'Explainable Artificial Intelligence in Health Care: How XAI Improves User Trust in High-Risk Decisions', Studies in Computational Intelligence, 1072, pp. 89–99.
Saeed, W. and Omlin, C. (2023) 'Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities', Knowledge-Based Systems, 263.
Sarker, S. et al. (2019) 'The sociotechnical axis of cohesion for the IS discipline: Its historical legacy and its continued relevance', MIS Quarterly: Management Information Systems, 43(3).
Scholar, B., Chakrobartty, S. and El-Gayar, O. (2021) Explainable Artificial Intelligence in the Medical Domain: A Explainable Artificial Intelligence in the Medical Domain: A Systematic Review Systematic Review, Systematic Review.
Sharma, R., Mithas, S. and Kankanhalli, A. (2014) 'Transforming decision-making processes: A research agenda for understanding the impact of business analytics on organisations', European Journal of Information Systems. Palgrave Macmillan Ltd., pp. 433–441.
Shibayama, T. et al. (2020) 'Stakeholder involvement in decision-making process: A test assessment towards transition to autonomous vehicles', in Transportation Research Procedia, pp. 2550–2568.
Taljaard, T. and Gerber, A. (2022) 'The Preparation of South African Companies for the Impact of Artificial Intelligence', in Artificial Intelligence Research. Springer Nature Switzerland, pp. 348–367.
Templier, M. and Paré, G. (2018) 'Transparency in literature reviews: an assessment of reporting practices across review types and genres in top IS journals', European Journal of Information Systems, 27(5), pp. 503 – 550.
Thompson, J. (2022) 'A Guide to Abductive Thematic Analysis', Qualitative Report, 27(5).
Turpin, S.M. and Marais, M.A. (2004) Decision-making: Theory and practice. Available at: http://www.orssa.org.za.
Wang, X. and Yin, M. (2022) 'Effects of Explanations in AI-Assisted Decision Making: Principles and Comparisons', ACM Transactions on Interactive Intelligent Systems, 12(4), pp. 1–36.

Wheeler, R. and Carroll, F. (2023) 'An Explainable AI Solution: Exploring Extended Reality as a Way to Make Artificial Intelligence More Transparent and Trustworthy', in Springer Proceedings in Complexity, pp. 255–276.

Appendix 1

A copy of the link of the interview questionnaire can be found at Interview Questionnaire.

25. Optimization of Call Center Agent Resources Using Various Machine Learning Methods: A Systematic Review

Can Yüzüak
Boğaziçi University
can.yuzuak@std.bogazici.edu.tr

Furkan Nurdağ
ComData
furkan.nurdag@comdatagroup.com

Mesut Tartuk
ComData
mesut.tartuk@comdatagroup.com

\

Abstract

Prediction problems have been highly studied in the literature especially in service systems. Recently machine learning based solutions have been added to find optimal solutions to such problems. In this study, the aim is to compile the existing recent literature about prediction problems in various fields. A systematic literature review process has been followed, using ScienceDirect, Scopus and Web of Science. After filtering the outcomes of the databases 24 articles are identified to be fully analyzed. The results indicate that formulating the prediction as a classifier problem, which enables usage of XGBoost, Bayesian Classifier, regression etc. provides need for labelled data, and provide computational ease, whereas can be less accurate, compared to deep learning solutions. Deep learning solutions are, however, "black-box" as in in any field, literature cannot fully identify each iteration that has been recorded. Finally, literature suggests combining into hybrid solutions, as in using non-ML and ML solutions together, for example heuristic rulesets and deep learning to obtain a better fit.

Keywords: agent performance optimization, model comparison, call prediction, capacity management, resource optimization, machine learning

1. Introduction

In numerous fields, especially with the ones that have high variability regarding its demand, successfully supplying all demand is a complicated process. There are internationally defined supply metrics, such as Quality of Experience (QoE) and Quality of Service (QoS) for technological fields. In order to perform adequately in terms of supply satisfaction, workload management and workload prediction is proved beneficial. Having infrastructures that are planned and resource allocated professionally as described, so that these infrastructures can supply the workloads, moreover with the power of being scalable in manner that decreases operational overhead costs and offers the most competitive services for end users (Fernández-Cerero et al., 2022). Insufficient allocation of both personnel and capital resources leads to Service Level Agreement (SLA) breaches, diminished quality of service (QoS), and customer discontent. Any organisation doing so may face penalties and compensation demands. Moreover, other than financial damages, the negative user experience, such as prolonged response times, reputation loss and overall negative branding may also hurt the organisation's

well-being. On the other hand, over-allocating resources results in wastage of energy, resources, and elevated costs for agent salaries and fees, technological over expenses, increased maintenance costs and overall unnecessary overhead costs (Matoussi & Hamrouni, 2022). Striking the right balance is crucial to avoid both service quality degradation and unnecessary resource expenditures (Matoussi & Hamrouni, 2022). To address this challenge, many scholars have suggested AI/ML based solutions, as well as more traditional statistical/heuristical solutions. Given that the accuracy of energy demand predictions for buildings or energy consumers profoundly influences the decision-making process, it becomes crucial to minimize errors in these forecasts (Ajayi & Heymann, 2021). Another instance of the similar practice, would be in solar radiation systems. Global solar radiation prediction is used for energy systems and environmental purposes all together (Ghimire et al., 2022). For the specific instance of call centers, there are numerous use cases in regards of AI usage in optimization purposes. Valle et al. (2012) evaluate individuals that are in their onboarding processes by employing a naive Bayes classifier to differentiate between those who meet the minimum performance standards required by the company to retain their position, based on demographic and operational characteristics. Another very similar point of view in the call center context, which again focuses the agent quality itself, (de Oliveira et al., 2019) There are also examples of studies including scheduling and assigning agents to departments (Phikulngoen et al., n.d.) and inbound call volume prediction (Mohammed, 2017).

The role of AI can be quite variant, in sense of computational ease, prediction and scaling can all be prioritised with tools within ML scope. In the realm of machine learning (ML), distinct paradigms include supervised, unsupervised, semi-supervised, and Reinforcement Learning (RL). Supervised learning relies on labelled data, unsupervised learning does not require labels, semi-supervised learning utilises a combination of labelled and unlabeled data for training, and RL aims to learn a strategy maximising a predefined quantity. Within ML, deep learning (DL) methodologies, particularly Neural Networks (NNs), have gained widespread popularity.

2. Research Methodology

In this study, the purpose is to systematically organize the literature work for the research questions of i)“How can the optimization of call center agent resources be achieved through the integration of call prediction models and agent scheduling using various machine learning methods”, and ii) “what are the comparative effectiveness and efficiency outcomes of these methods with correct data and feature preprocessing in terms of improving overall call center performance?” The research questions are aiming to improve upon the identified literature gap of recent studies that examines call volume forecasting and prediction’s correlation with agent cost and performance. This article is an initial systematic literature review, which will be the base study of a thesis that will be published within summer 2024. This article is curated in January 2024.

Systematic review, which follows PRISMA guidelines (Moher et al., 2009) involves searching for relevant articles to address specific research questions. The literature gathering and filtering method employed follows a comprehensive top-down approach. Initially, a broad set of phrases is used to collect a wide range of articles. Subsequently, a filtering process is applied to exclude any articles that appear irrelevant to the research, ensuring a focused and pertinent selection.

For the specified research question, the used keywords are provided in Table 1. Multiple objectives, such as “resource optimization”, “capacity management”, “call prediction” and “agent performance optimization” since the research question takes into account multiple perspectives in order to achieve optimal output of call center agent scheduling. These keywords are chosen to capture best practices and methods that are being utilised in various sectors so that these best practises can be addressed into call center context.

agent performance optimization	capacity management
model comparison	resource optimization
call prediction	machine learning

Table 1: Keywords used with the “AND” boolean operator

These keywords are run through three databases: Scopus, Science Direct and Web of Science. Scopus and Web of Science databases are chosen in regards of their extensive and internationally accepted usage (Zancanaro et al., 2015).Therefore, they entirely cover application areas of multimedia learning principles. This process overall provided 991 articles, then with the addition of 6 articles from the preliminary search there was a cumulative of 997 articles that were considered. Due to the exclusion criteria, which are provided in Table 2, 904 of the articles are excluded and eliminated from the scope of this research. The screening phase was conducted with 93 articles. Through abstract elimination, 69 articles is found out of scope in regards of the research question, thus full text assessment was conducted with 24 articles. Final 24 articles are read and analysed thoroughly, and are being referenced in this article. These articles are examined and the information received is separated upon the titles in this article to be able to convey the details optimally. Detailed process of literature search and review is displayed in Figure 1.

3. Result Discussion

3.1: Feature Selection, Feature Transformation & Data Pre-processing

3.1.1 Data Pre-processing & Feature Selection

Machine learning (ML) empowers systems to autonomously learn and enhance their performance through experience, eliminating the need for explicit programming. This learning process occurs as ML algorithms enable computers to acquire the knowledge necessary for a particular task by analyzing a substantial amount of relevant data. In traditional ML, data needs to be represented in terms of a fixed number of features. Consequently, a "feature extraction" step is essential, involving pre-processing the data to extract specific attributes that encapsulate the most characteristic information. In essence, feature extraction is a critical pre-processing step that facilitates the effective training of ML systems by highlighting and representing key characteristics of the data (Avci et al., 2021).

In the information layer, data-handling methods encompass vital processes, such as data pre-processing, anomaly detection, and data analysis. These methods play a crucial role in shaping appropriate data models. The primary objective of data pre-processing is to generate a time series devoid of anomalies, facilitating effective data analysis and forecasting (Richter et al., 2022).

When dealing with a new time series lacking prior knowledge of its characteristics, it is advisable to undergo various pre-processing steps to gain a comprehensive understanding of its general features. The next step involves the application of different data transformation techniques before selecting a suitable forecasting model. This systematic approach ensures that the time series is prepared optimally for analysis and forecasting, considering its unique characteristics and potential anomalies (Richter et al., 2022).

However, since every study entails very different data sources throughout their processes, data pre-processing and feature selection can differentiate greatly. Data conversion (data to timestamp etc.), autocorrelation analysis, standardization and normalizations are very common practises to scale out datasets (Matoussi & Hamrouni, 2022). While no methods can guarantee 100% accuracy, even small enhancements in data pre-processing and forecasting models can lead to a substantial increase in forecasting accuracy and a reduction in costs. This highlights the importance of continuous efforts to refine and optimize data pre-processing techniques and forecasting models, as incremental improvements can have a meaningful impact on the overall effectiveness and efficiency of the forecasting process. Investing in refining these aspects of the workflow can contribute to more reliable predictions and cost savings over time (Carballal et al., 2021).

3.1.2 Data Partitioning

It is a common practise to partition the data as test and validation sets. In the process of model development, the training set is used to calibrate the models, allowing them to learn from the provided data. Meanwhile, the validation set is distinct and does not play a role in the actual training process; instead, it is utilized to fine-tune the models during the development phase. The primary purpose of the validation set is to assist in optimizing various parameters and enhancing the model's performance. After the models have been trained on the combined training and validation sets, the test set comes into play. The test set is exclusively reserved for evaluating the performance of the model. It provides an independent dataset that the model has not encountered during training or validation, allowing for a robust assessment of how well the model generalizes to new, unseen data. This separation into training, validation, and test sets helps ensure the reliability and effectiveness of machine learning models in real-world applications (Ghimire et al., 2022). Usage of stratified sampling is also very common, aim is to avoid data leakage and to obtain an unbiased data analysis for training and test sets (Lim et al., 2022).

3.2 Building Prediction Models

After finalizing data pre-processing, feature selection and deciding upon the data partitions, model building is the next step to the finalizing prediction solution to call center/service system optimization. The goal is to successfully predict incoming workload, and the consumption rate of the workload predicted. (Matijašević et al., 2022). For instance, Farahzadi and Kioumarsi (2023) claim that classical optimization algorithms tend to provide more accurate estimation results compared to numerical and statistical methods. Hybrid approaches combine the strengths of classical optimization methods with ML techniques, offering improved performance and versatility. By integrating optimization strategies into ML models, these

hybridized methods aim to enhance the robustness, efficiency, and overall effectiveness of the learning process, providing a more comprehensive solution to complex problems.

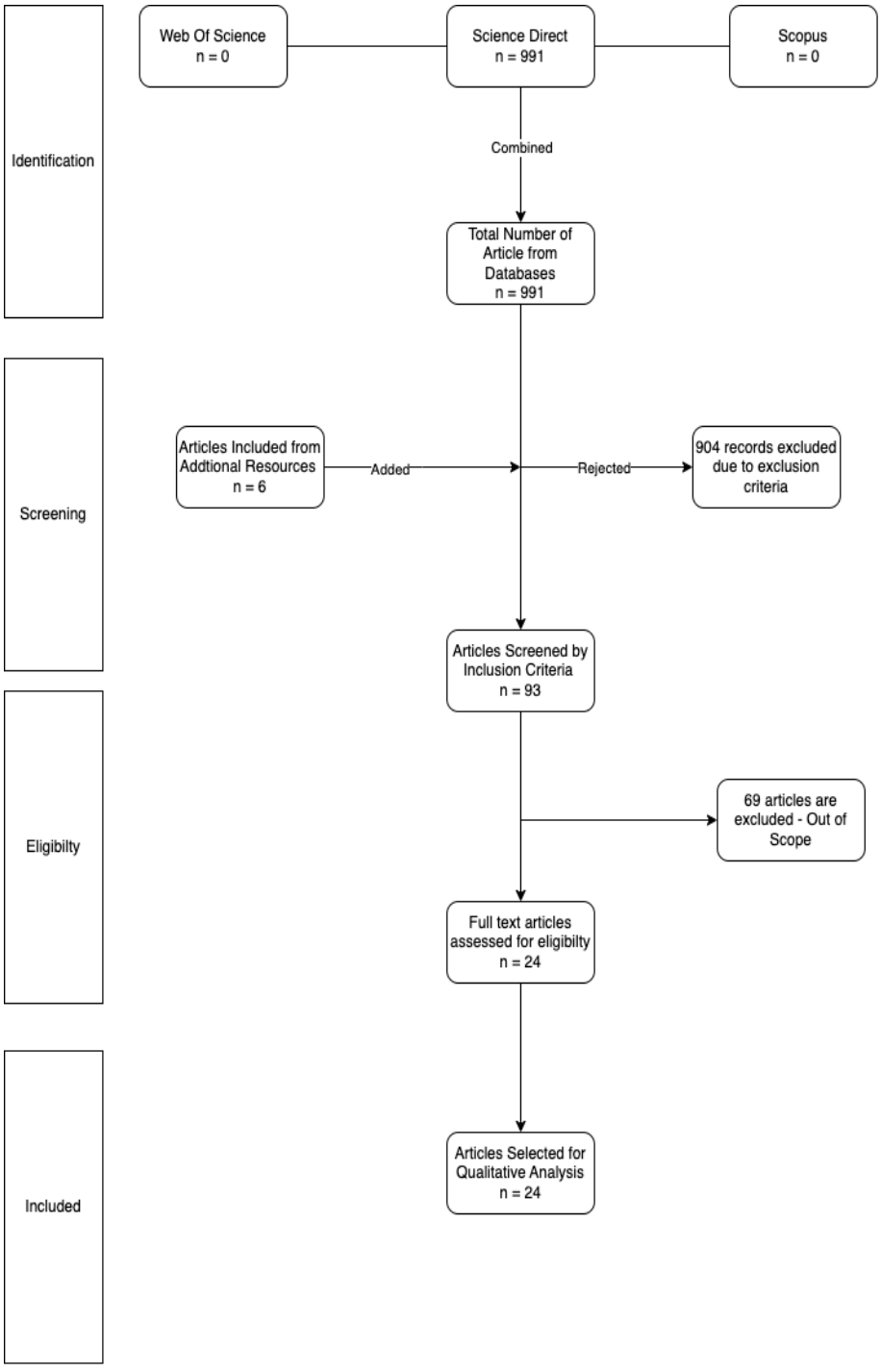


Figure 1: Flow Diagram of Article Selection Process

Inclusion Criteria	Exclusion Criteria
Articles should have been published in journals indexed by WoS, Scopus, and ScienceDirect databases	Conference proceedings, dissertations/thesis, and book chapters are not selected
Articles are peer-reviewed	Article type should be review or research articles
Full-text available	Articles that are not published in English
	Articles that are published within the time range of 2018 to present

Table 2: Inclusion and exclusion criteria

3.2.1 Supervised Models

Regression serves as a supervised machine learning method employed to forecast a continuous outcome variable by considering one or more predictor variables. The central objective of regression analysis is to establish a model that illustrates the connection between independent variables (features) and the dependent variable (target) to facilitate predictions. Within the realms of data science and statistics, regression analysis proves to be a versatile tool, offering valuable insights into variable relationships and facilitating precise predictions for continuous outcomes. There are many cases that regression is used in predicting computational potential load in cloud computing systems (Kamila et al., 2022). Kamila et al., (2022) also suggest that using computing data, load balancer performance can be predicted by python regressor functions. In the similar field, however, in a macro scope, Fernández-Cerero et al., (2022) use regression models, specifically gradient regression boosting, which is tasked with estimating the scheduling performance of each resource management model within a catalogue for a specific data-center operating scenario and timeframe. Based on this performance estimation, the optimal resource management model is implemented for the corresponding period.

Tree Based algorithms, with their representative hero being Random Forest (RF) is being used as the classifier/predictor in many areas of expertise. Tree-based classifiers belong to a category of machine learning algorithms that rely on a hierarchical tree-like structure to make decisions. Decision Trees, a foundational model within the tree-based approach, segment the dataset into subsets based on the values of input features, creating a tree structure where each node signifies a decision. (Lim et al., 2022)

Random Forest, an ensemble learning technique, extends the concept of Decision Trees. It generates numerous decision trees during the training phase and produces the mode of classes for classification tasks or the average prediction for regression tasks. The "random" aspect in Random Forest arises from utilizing a random subset of features for each split in the trees, introducing diversity among the individual trees. This diversification is instrumental in averting overfitting, resulting in a more resilient and accurate model (Lim et al., 2022). Keon et al., (2018) demonstrated that regular decision tree models have less accuracy to successfully

predict total cumulative call duration for the specified call center, then random forests, which is a special version of decision tree models.

On the other hand, Bayesian classifiers determine the most probable class for each sample based on the vector values of its variables. This model has found applications in various domains, including drug discovery, where it is utilized to predict potential drug targets (Carballal et al., 2021). In call center use case, usually performance variables that are being used as input and the labels, all of which tend to be continuous, underwent a discretization process for the Bayesian classifier. This discretization was carried out using a minimum entropy heuristic (Valle et al., 2012). By doing so, approaching to the call center performance prediction in a different perspective, Valle et al. (2012) used Naive Bayes, a specific Bayesian approach to classify agents. Agent classification can be used in parallel with the already discussed incoming workload predictions, to create a complex hybrid approach.

Support vector machine (SVM) is within supervised learning. Its core objective is to determine an optimal hyperplane that maximizes the margin between distinct classes within the input data. The significance of Support Vectors, or data points closest to the hyperplane, is crucial in establishing this optimal boundary, enhancing the algorithm's resilience. Its adaptability and proficiency in handling intricate data relationships contribute to its widespread adoption in machine learning applications. The primary objectives are to detect anomalies and pinpoint the root causes of bottlenecks, particularly based on Software-Defined Networking (SDN). The SVM algorithm is trained using performance metrics data, and a baseline model is constructed to represent normal application behavior (El-Shamy et al., 2021).

Finally, XGBoost (XGB), a very specific model is a solution worth mentioning. In a study that is also considering an edge scenario of “absenteeism” de Oliveira et al., (2019) uses XGB which basically uses gradient boosting framework, in order to build a strong predictive model by combining the predictions of multiple weak models, typically decision trees, in order to classify call center agents next t periods absenteeism.

3.2.2 Neural Networks & Deep Learning

As mentioned earlier, deep learning is one of relatively later introduced solutions amongst the other listed. Firstly, in order to understand the most fundamental version of the neural networks, Artificial Neural Networks (ANN), it is essential to start with the artificial neuron, a fundamental component of the network. After calculating the net value, a trigger function is applied to determine the output of the processing element (Carballal et al., 2021).

Based on the concept of an artificial neuron, multiple neurons can be interconnected to form a network, where the output of one neuron becomes the input for another. In the context of ANNs, input nodes are crucial for obtaining information from the external environment, constituting the input layer of the network (Theodoropoulos et al., 2023). The network also requires output nodes, situated in the hidden layer, to convey the final result of the ANN. The remaining nodes, known as hidden nodes, facilitate the transmission of information between neurons and are grouped in one or several hidden layers (Carballal et al., 2021).

Evaluating the output of an ANN involves assessing all the neurons within the network, and during training, this process is iterative. Consequently, ANNs typically have a limited number of neurons, leading to a complex web of connections where the function of each part is not easily interpretable. This characteristic renders an ANN as a "black box," signifying that the knowledge is embedded in the intricate connections without explicit insight into the individual actions of each component (Carballal et al., 2021). Carballal (2021) in detail focuses upon medical usages of ANN practices. The focus is very parallel to what the research question of this study is, it being predicting the function and outcome of a drug to a certain cell group.

Similarity-Based Methods (SBM) constitute a family of techniques designed to assess feature importance based on their ability to preserve data similarity, and they naturally find realizations in neural network models. This family encompasses neural network models like Radial Basis Functions (RBF) and Multilayer Perceptrons (MLPs). Unlike specific learning algorithms, SBM is not tied to any particular learning method, making it a versatile and generalizable framework (Samunderu & Farrugia, 2022). In their study, Samunderu and Farrugia (2022) have shown that Similarity-Based Feature Selection algorithms exhibit excellent performance in both supervised and unsupervised learning tasks. However, it is important to note a key drawback: these methods struggle with handling feature redundancy. Despite their overall effectiveness, the inability to manage redundant features represents a limitation in their application.

Genetic Algorithms (GAs) typically initiate with a randomly selected and assessed population, where each individual represents a potential solution to the given problem. Subsequently, the following cycle is iteratively executed until a termination condition is met:

1. Selection of individuals for recombination,
2. Recombination of the traits of the selected individuals to produce descendants
3. Random mutation of the descendants' traits
4. Evaluation of the mutated descendants
5. Determination of the next generation

(Richter et al., 2022).

In of the scenarios that were presented in Farahzadi and Kioumarsi (2023), a mixed-integer linear and multi-objective model, specifically GA, was chosen to evaluate and optimize the performance of the supply chain within a closed loop comprising a four-level network. The outcomes obtained from the solution analysis indicated that an increase in the utilization of recycled aggregates has the potential to reduce the need for excavation in quarries during concrete production. In the same body of work, multiple other usage were presented, making GAs one of the main solution models to be used in supply chains' eco-friendly output prediction models. Being an exhaustive manner, and solving the neural network problem of getting stuck on local optimas, GAs are tend to be used coupled with NNs in prediction problems.

Convolutional Neural Networks (CNNs) belong to a category of sophisticated deep learning models crafted for handling structured grid data, notably images. These models have demonstrated remarkable success in various computer vision applications, including image classification, object detection, and image segmentation. Inspired by the organizational

principles of the animal visual cortex, CNNs excel at extracting hierarchical features from visual data, showcasing their adaptability and effectiveness in image-related tasks. For prediction tasks involving visual data especially, Tang et al., (2022) lists many usages in regards of railway health and fault detection, especially being used for prediction via pattern detection. Some of the usages actually can be interpreted into service systems, for example traffic planning. They actually divided this planning into two, being tactical planning having a micro scope, and strategical planning having a macro scale of longer timeframe.

LSTM, or Long Short-Term Memory, is a nonlinear model designed for analyzing time series data, allowing for the persistence of information through a memory state. It adeptly captures the dependencies between observations in a sequence. Arising from the limitations of Recurrent Neural Networks (RNN), LSTM addresses issues like the vanishing gradient problem and constraints on short-term memory. The vanishing gradient problem impedes the training of early RNN layers, and the short-term memory limitation restricts efficient information transfer across different time points, resulting in swift forgetting (Violos et al., 2020). Optimizing LSTM-RNN parameters involves minimizing an objective function. Commonly utilized optimization techniques for LSTM-RNN include Root Mean Square Propagation (RMProp) and Adaptive Moment Estimation (Adam). These approaches contribute to improving the training and efficacy of LSTM in capturing extensive dependencies within sequential data (Violos et al., 2020). Violos et al. (2020) used LSTM in order to then again for data center optimization, in sense of predicting Raspberry Pi performance in cloud technologies.

4.2.3 Other Methods

In the face of dynamic call center traffic, it is crucial for workforce management to adapt in real-time to changes in order to meet service level agreements. Therefore, in this section, we introduce a novel approach to adaptive workforce management utilizing reinforcement learning (RL). The reinforcement learning framework comprises two primary components: the environment module and the reinforcement agent module. This methodology aims to enhance the adaptability of workforce management strategies in response to fluctuations in call center traffic (Phikulngoen et al., n.d.). Reinforcement learning continues to provide value in two parallel process, first passively observing and labelling the data training data set but also provides control over the decided labels (Hernandez-Matheus et al., 2022).

Since the focus is also about predicting, whenever the researcher ends up with time series data, forecasting becomes available as a primary option. The cornerstone of effective forecasting lies in data quality (Matijašević et al., 2022). Based on the initiation period of the forecast, three distinct categories are identified: Short-term, medium-term and long-term forecasts (Matijašević et al., 2022). While Matijašević et al., (2022) used time series forecasting for energy distribution systems, in detailed version to predict demand and supply of energy demands from consumers, and energy supply from prosumers. On the other hand, Hernandez-Matheus (2022) used forecasting in 4 different fields in regards of energy communities, them being demand forecasting, renewable generation forecasting, elasticity forecasting and price forecasting. However, it is worth mentioning that, forecasting developed greatly in the recent processes, and utilizes ML greatly.

4. Conclusion

In conclusion, the challenges associated with effectively supplying demand, especially in fields with high variability, necessitate a strategic approach to workload management and prediction. Internationally defined metrics, such as Quality of Experience (QoE) and Quality of Service (QoS), guide technological fields in achieving supply satisfaction. The failure to adequately allocate personnel and capital resources can lead to disadvantage in many fields, in forms of fines, penalties and reputation loss. Addressing these challenges requires a thoughtful integration of AI and ML solutions, alongside traditional statistical and heuristic approaches. AI, defined as a field encompassing all elements enabling entities to execute tasks or make decisions, proves essential for discerning patterns and optimizing resources. ML, with distinct paradigms such as supervised, unsupervised, semi-supervised, and Reinforcement Learning, plays a crucial role. Deep Learning methodologies, particularly Neural Networks (NNs), gain popularity in ML, providing automatic feature learning and representing a cornerstone in the development of robust systems.

In this review, following SLR methodology, a review starting with 997 articles and 24 full text reviews were conducted. As a result, more than 15 different ML and non-ML solution and their respective drawbacks and success is presented. Finally, there are unfortunately some limitations to this study. The review has some constraints. It exclusively relied on specific databases, namely Scopus, Web of Science, and ScienceDirect, for collecting articles. There may be additional papers addressing prediction problems, especially in service sector in other databases and resources. Lastly, exclusion criteria is kept tight, which limited the scope of the study.

References

Ajayi, O., & Heymann, R. (2021). Data centre day-ahead energy demand prediction and energy dispatch with solar PV integration. *Energy Reports*, 7, 3760-3774.

Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., & Inman, D. J. (2021). A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications. *Mechanical systems and signal processing*, 147, 107077.

Carracedo-Reboredo, P., Liñares-Blanco, J., Rodríguez-Fernández, N., Cedrón, F., Novoa, F. J., A., ... & Fernandez-Lozano, C. (2021). A review on machine learning approaches and trends in drug discovery. *Computational and structural biotechnology journal*, 19, 4538-4558.

de Oliveira, E. L., Torres, J. M., Moreira, R. S., & de Lima, R. A. F. (2019). Absenteeism prediction in call center using machine learning algorithms. *In New Knowledge in Information Systems and Technologies: Volume 1* (pp. 958-968). Springer International Publishing.

El-Shamy, A. M., El-Fishawy, N. A., Attiya, G., & Mohamed, M. A. (2021). Anomaly detection and bottleneck identification of the distributed application in cloud data center using software–defined networking. *Egyptian informatics journal*, 22(4), 417-432.

Farahzadi, L., & Kioumarsi, M. (2023). Application of machine learning initiatives and intelligent perspectives for CO2 emissions reduction in construction. *Journal of Cleaner Production*, 384, 135504.

Fernández-Cerero, D., Troyano, J. A., Jakóbik, A., & Fernández-Montes, A. (2022). Machine learning regression to boost scheduling performance in hyper-scale cloud-computing data centres. *Journal of King Saud University-Computer and Information Sciences*, 34(6), 3191-3203.

Ghimire, S., Deo, R. C., Casillas-Pérez, D., Salcedo-Sanz, S., Sharma, E., & Ali, M. (2022). Deep learning CNN-LSTM-MLP hybrid fusion model for feature optimizations and daily solar radiation prediction. *Measurement*, 202, 111759.

Hernandez-Matheus, A., Löschenbrand, M., Berg, K., Fuchs, I., Aragüés-Peñalba, M., Bullich-Massagué, E., & Sumper, A. (2022). A systematic review of machine learning techniques related to local energy communities. *Renewable and Sustainable Energy Reviews*, 170, 112651.

Jun, S. P., Lee, J. S., & Lee, J. (2020). Method of improving the performance of public-private innovation networks by linking heterogeneous DBs: Prediction using ensemble and PPDM models. *Technological Forecasting and Social Change*, 161, 120258..

Kamila, N. K., Frnda, J., Pani, S. K., Das, R., Islam, S. M., Bharti, P. K., & Muduli, K. (2022). Machine learning model design for high performance cloud computing & load balancing resiliency: An innovative approach. *Journal of King Saud University-Computer and Information Sciences*, 34(10), 9991-10009.

Keon, Y., Kim, H., Choi, J. Y., Kim, D., Kim, S. Y., & Kim, S. (2018). Call Center Call Count Prediction Model by Machine Learning. *Journal of advanced information technology and convergence*, 8(1), 31-42.

Lim, A. J., Lim, L. J., Ooi, B. N., Koh, E. T., Tan, J. W. L., Chong, S. S., ... & Lee, C. G. (2022). Functional coding haplotypes and machine-learning feature elimination identifies predictors of Methotrexate Response in Rheumatoid Arthritis patients. *EBioMedicine*, 75.

Matijašević, T., Antić, T., & Capuder, T. (2022). A systematic review of machine learning applications in the operation of smart distribution systems. *Energy Reports*, 8, 12379-12407.

Matoussi, W., & Hamrouni, T. (2022). A new temporal locality-based workload prediction approach for SaaS services in a cloud environment. *Journal of King Saud University-Computer and Information Sciences*, 34(7), 3973-3987.

Mohammed, R. A. (2017). Using personalized model to predict traffic jam in inbound call center. *EAI Endorsed Transactions on Scalable Information Systems*, 4(12), e2-e2.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group*. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*, 151(4), 264–269.

Phikulngoen, S., Kumwilaisak, W., Thatphithakkul, N., Hansakunbuntheung, C., & Phantachat, W. Call Center Workforce Management with Neural Networks and Reinforcement Learning.

Richter, L., Lehna, M., Marchand, S., Scholz, C., Dreher, A., Klaiber, S., & Lenk, S. (2022). Artificial intelligence for electricity supply chain automation. *Renewable and Sustainable Energy Reviews*, 163, 112459.

Samunderu, E., & Farrugia, M. (2022). Predicting customer purpose of travel in a low-cost travel environment—A Machine Learning Approach. *Machine Learning with Applications*, 9, 100379

Tang, R., De Donato, L., Besinović, N., Flammini, F., Goverde, R. M., Lin, Z., ... & Wang, Z. (2022). A literature review of Artificial Intelligence applications in railway systems. *Transportation Research Part C: Emerging Technologies*, 140, 103679.

Theodoropoulos, T., Makris, A., Kontopoulos, I., Violos, J., Tarkowski, P., Ledwoń, Z., ... & Tserpes, K. (2023). Graph neural networks for representing multivariate resource usage: A multiplayer mobile gaming case-study. *International Journal of Information Management Data Insights*, 3(1), 100158.

Valle, M. A., Varas, S., & Ruz, G. A. (2012). Job performance prediction in a call center using a naive Bayes classifier. *Expert Systems with Applications*, 39(11), 9939-9945.

Violos, J., Psomakelis, E., Danopoulos, D., Tsanakas, S., & Varvarigou, T. (2020). Using LSTM neural networks as resource utilization predictors: The case of training deep learning models on the edge. In S. Kounev, D. Hales, & J. Dominguez-Sal (Eds.), *Economics of Grids, Clouds, Systems, and Services: 17th International Conference, GECON 2020, Izola, Slovenia, September 15–17, 2020, Revised Selected Papers 17* (pp. 67-74). Springer International Publishing.

Zancanaro, A., Todesco, J. L., & Ramos, F. (2015). A bibliometric mapping of open educational resources. *International Review of Research in Open and Distributed Learning*, 16(1), 1–23.

26. The Snowball Effect of Multiple Value Co-Creation (VCC) Interactions Leading to Value Co-Destruction (VCD)
(Research-in-Progress)

Mohammed Azmy Ateia
German University in Cairo (GUC)
mohammed.azmy@dba.guc.edu.eg

Nermeen Fawzy Mohamed
German University in Cairo (GUC)
nermeen.mohamed@dba.guc.edu.eg

Abstract
This work-in-progress research initiative aims to explore the adverse effects of the extensive involvement in various Value Co-Creation (VCC) instances on the part of the customer as they engage in a multitude of value streams offered from multiple service providers, thus presumably resulting in Value Co-Destruction (VCD). This research work is of practical importance for professional managers who must protect their investments in digital Value Co-Creation (VCC) Platforms. Thus, they can continue to leverage the benefits that VCC offers such as enhanced customer engagement, increased customer loyalty, customization and personalization, cost savings, efficiency, and brand differentiation. This research effort addresses a gap in the literature, as the extant literature focuses on the different sources of Value of Co-Destruction arising within a single value stream. Such a research enquiry will be linked to contemporary phenomena, applicable scientific theories and standard performance metrics that are well established in the literature. Additionally, a question begs to be asked whether an authoritative organization should act on behalf of the customer/end-user and govern the "implied" involvement in numerous VCC streams. These proposals align with the principles of Industry 5.0 and Society 5.0, thus aim to serve the purpose of augmenting the role of human-centricity along with sustainability and resilience in a technology-centric world of Industry 4.0.

Keywords: *Value Co-Creation (VCC), Value Co-Destruction (VCD), Service-Dominant Logic (SDL), Industry 5.0 (I5.0), Society 5.0.*

1. Introduction
Picture a day whereby a person has to navigate various applications for social networking (multiple apps), entertainment (multiple apps), appointments (healthcare, or other), applications (VISA, education, job opportunity, or other), deliveries (groceries, food, apparel, home supplies, or other), bookings and reservations (hotel, train, flight, hail riding or other), bill payments (various utilities, various subscriptions, or others), and banking services. Add to this, multiple work-related cloud-based collaborative apps for document management and

communication. Can one have too much of a good thing? A survey concluded that 46% (forty-six percent) of the respondents counted 5-9 apps open at the time that they undertook the survey, and a significant 19% (nineteen percent) had 10-14 apps open (Harmon.ie, 2017).

2. Value Co-Creation (VCC)

The foundational concept of Value Co-Creation (VCC) traces back to the concept of Service-Dominant Logic (SDL), which was first developed by Vargo & Lusch (2004). They proposed a paradigm shift from a goods-centric to a service-centric view, one where value creation is an interactive process between providers and consumers. Building on this concept, Prahalad and Ramaswamy (2004) first introduced Value Co-Creation (VCC) as a new business model that deviated from the traditional supplier/customer relationship.

"Value processes (are) the value-related procedures, tasks, mechanisms, activities and interactions that may result in Value Co-Creation or Value Co-Destruction." (Plé, 2017). While Value Co-Creation is defined as "the joint, collaborative, concurrent, peer-like process of producing new value, both materially and symbolically." (Galvagno & Dalli, 2014).

Value Co-Creation is indeed evolving. In practice, managers need to mitigate the risk of customers' abandoning their applications and thus rendering the digital investments useless. Along these lines, product designers incorporate chatbots to simulate human interaction (Bassano et al., 2020), and resort to Ecosystem Collaboration (Pera, Occhiocupo & Clarke, 2016). Apps are linked together via Application Programming Interfaces (APIs) to simplify the user journey and avoid ping-pong between different apps, thus avoiding the app overload. Value Co-Creation is also enabled via Crowdsourcing and Open Innovation Platforms that enable companies to tap into the collective intelligence and creativity of external communities; customers, enthusiasts, and experts (Liu, Zhao, & Sun, 2018).

2.1. Benefits of Value Co-Creation (VCC)

On one hand, Value Co-Creation serves multiple benefits for the Organization. (1) Thanks to the collaborative nature of Value Co-Creation, organizations become more innovative as they are able to tap into the diverse perspectives, ideas, and feedback from customers, and incorporate these valuable customer insights into the development of products, services, and business processes (Barile et al., 2020). (2) Moreover, Value Co-Creation nurtures Customer Loyalty and Retention. As customers contribute and exert effort to the creation of value, this builds a sense of ownership, and customers become more likely to feel connected to the brand. This creates a positive brand authenticity (Banik & Rabbanee, 2023; Kumar & Kandoi, 2018).

(3) In addition, Value Co-Creation allows organizations to adapt quickly to changing market trends and customer preferences and expectations, thus leading to higher levels of customer satisfaction. Satisfied customers become advocates, promoting the organization through positive word-of-mouth. (4) Augmented by continuous collaboration, real-time Value Co-Creation enables adjustments to offerings, ensuring they stay relevant and aligned with evolving customer needs (Vega-Vazquez et al., 2013). (5) Furthermore, joint Co-Creation helps organizations optimize resource allocation by focusing on features, services, or improvements

that are valuable from the customers' viewpoint and Co-determined with customers themselves. This reduces the risk of investing in aspects that may not resonate with the target audience. (6) This also facilitates agile decision-making about resource allocation, as organizations receive immediate feedback, enabling them to make informed decisions and respond promptly to market changes or unexpected challenges (Agrawal & Rahman, 2015).

On the other hand, Value Co-Creation also provides many benefits for the Customers. (1) Interactive relationship fosters a sense of partnership and demonstrates to the customers that their opinions are valued. (2) As organizations open-up their capabilities in terms of tools and systems to their customers, customers benefit from personalized offerings in the form of products and services tailored to their specific needs and preferences. Thus, Value Co-Creation allows customers to influence the development of offerings, resulting in a more individualized experience (Zine et al., 2014). (3) Furthermore, Value Co-Creation empowers customers by giving them a voice in the design and improvement of products and services (Moussafir & Qmichchou, 2021). This involvement enhances customers' engagement with the brand, making them feel more valued and connected. (4) It also contributes to a sense of ownership and a feeling that their preferences have been considered. Overall, the customers who actively participate in Value Co-Creation are more likely to be satisfied with the end product or service (Alexander & Jaakkola, 2015).

(5) In addition, the co-created offerings often have a higher perceived value for customers because they directly contribute to the design process. They customize products or services according to their preferences. This flexibility ensures that the offerings meet individual needs, leading to a more satisfying and relevant experience. This resulting sense of involvement and contribution enhances the overall value proposition (Akter et al., 2022; Solakis et al., 2022; Lalicic & Weismayer, 2021). (6) Moreover, the customers involved in Value Co-Creation may gain access to exclusive features, early releases, or special privileges. This adds an element of exclusivity and enhances the overall value proposition for actively engaged customers. (7) Additionally, Value Co-Creation fosters a sense of community and social connections among customers who share similar interests or preferences. This community-building aspect enhances the overall customer experience, creating a network of individuals with shared values (Chew et al., 2013).

2.2. Challenges to Value Co-Creation (VCC)

While Value Co-Creation offers numerous benefits, it also comes with its own set of challenges.

(1) Establishing and maintaining trust between businesses and customers is crucial for effective Value Co-Creation. Overcoming initial skepticism and building a transparent, trusting relationship can be challenging (Järvi et al., 2018). (2) Successful Co-Creation often requires significant time, effort, and resources from both the business and customers. Resource constraints may limit the feasibility and sustainability of Value Co-Creation initiatives (Plé & Chumpitaz Cáceres, 2010).

(3) Co-Creation initiatives can become complex, involving multiple stakeholders, technologies, and processes. Managing this complexity and ensuring a smooth collaborative experience can be a significant challenge (Buhalis et al., 2020). (4) Businesses may be

concerned about protecting their intellectual property, and customers may be hesitant to share valuable ideas or data. Clear agreements and communication are needed to address these concerns (Pera et al., 2021).
(5) Balancing the need for control with the desire for openness and collaboration can be challenging for businesses. Finding the right level of control to maintain quality and consistency while encouraging Co-Creation is crucial (Buhalis et al., 2020). (6) Sustaining customer engagement over time can be a challenge. Customers may lose interest or face competing priorities, leading to a decline in participation (Nangpiire et al., 2022). (7) Cultural diversity among customers and within organizations can pose challenges. Differences in communication styles, expectations, and approaches to problem-solving may need to be navigated for successful Co-Creation (Grott et al., 2019). (8) Defining, measuring, and evaluating the value generated through Co-Creation can be challenging. Traditional metrics may not fully capture the impact on customer satisfaction, loyalty, or long-term value- (Skaržauskaitė, 2013).

3. Value Co-Destruction (VCD)

Value Co-Destruction refers to situations where the interaction between a company and its customers results in a reduction or destruction of value rather than value creation. "Value Co-Destruction can be defined as an interactional process between service systems that results in a decline in at least one of the systems' well-being (which, given the nature of a service system, can be individual or organizational). During this process, these service systems interact either directly (person-to-person) or indirectly (via appliances such as goods) through the integration and application of resources." (Plé & Chumpitaz, 2010).

Several factors can contribute to Value Co-Destruction. (1) Mismatched Expectations, that arise when there is a significant gap between customer expectations and the actual product or service delivered, thus leading to dissatisfaction and perceived value loss. This is equally possible when products or services fail to meet quality standards or performance expectations, resulting in customer disappointment (Järvi et al., 2018; Järvi et al., 2020). (2) Ineffective Communication is another potential source, which may manifest as communication breakdowns, misinformation, or inadequate communication about product features, changes, or policies. These mal-interactions can lead to confusion and frustration among customers, negatively impacting their perception of the delivered value (Assiouras et al., 2023). (3) Ineffective communication may cause a Lack of Customer Engagement. Limited or ineffective engagement with customers can result in a loss of connection and understanding of their evolving needs. This lack of engagement may lead to customers feeling neglected or undervalued (Zhang et al., 2018). (4) Another potential factor that may contribute to Value Co-Destruction is Rigid and Inflexible Business Processes that do not adapt to customer needs or changes in the market. Customers may find it difficult to navigate or work with systems that are not user-friendly or accommodating (Assiouras et al., 2023). (5) A further consideration is Unresolved Issues and Failure to address customer complaints, concerns, or issues in a timely and satisfactory manner. This escalates dissatisfaction and erodes the perceived value of the product or service (Dolan, Seo & Kemper, 2019). (6) Technological Issues: in the form of Technical glitches, system failures, or issues with digital platforms can disrupt the customer experience and contribute to Value Co-Destruction (Zhang et al., 2018). (7) If a company fails

to innovate and keep up with evolving customer preferences, technological advancements, or industry trends, customers may perceive a decline in the value of the offerings (Čaić et al., 2018).

4. VCD arising from involvement in multiple value streams

Extant literature addresses the different sources of Value of Co-Destruction arising within a single value stream. However, there is a gap in the literature when it comes to addressing the adverse effects to the extensive involvement in various Value Co-Creation instances on the part of the customer as they engage in a multitude of value streams offered from multiple service providers. Accordingly, the following research question is postulated:

RQ1: Does excessive involvement in numerous successive Value Co-Creation instances – each created by a different organization – lead to value Co-Destruction on the part of the individual customer?

Such an enquiry may be linked to contemporary phenomena, applicable scientific theories and standard performance metrics. Excessive exposure to Digital VCC would ultimately lead to both Information Overload and Digital Overload (Marques & Batista, 2017). Withdrawal from a Co-Creation interaction in the context of health services that has been documented in the literature (Keeling et al., 2021). Evidently, Digital VCC interactions extend this behavior into other aspects of life in the form of Social Withdrawal or Hikikomori (Li & Wong, 2015).

As VCC has been augmented by the widespread of digital technologies in recent years, the following theories, models and frameworks can be adopted to assess the impact of excessive VCC engagements and whether VCD is indeed emerging: Diffusion of Innovation (DOI) Theory, developed by Rogers, Singhal & Quinlan (2014); Technology Acceptance Model (TAM) by Davis(1989) and Davis, Bagozzi & Warshaw (1989); and The Unified Theory of Acceptance and Use of Technology (UTAUT) by Venkatesh, Morris, Davis & Davis (2003).

The relevant standard performance metrics include: Customer Effort Score (CES) that was coined by Corporate Executive Board (CEB Global), now known as Gartner, (Dixon et al., 2010); Net Promoter Score (NPS) created by Reichheld (2003) a partner at Bain & Company, and Customer Satisfaction (CSAT) score which was first introduced by the University of Michigan as it established the American Customer Satisfaction Index (ACSI), (Fornell et al., 1996).

On a final note, “implied user consent” in VCC seems to be an issue worth investigation. Analogies may be drawn with data privacy, security and confidentiality concerns that were attended to with initiatives such as General Data Protection Regulations (GDPR) and The right to be forgotten (RtbF), (Mantelero, 2013). This thinking has led the authors to pose the additional research question:

RQ2: Is there a need for an authoritative body to act on behalf of the customer/end-user and govern the “implied” involvement in numerous VCC on the part of customer/end-user?

The Authoritative body may be entrusted with the following responsibilities: Standardization and Quality Assurance, Regulatory Oversight, Risk Management and Mitigation, Conflict Resolution, Education and Awareness, Industry Collaboration and Coordination, and Public Trust and Confidence (Alaassar, Mention & Aas, 2020).

These research questions align with the principals of Industry 5.0 and Society 5.0 that aim to serve the purpose of augmenting the role of human-centricity along with sustainability and resilience in a technology-centric world of Industry 4.0 (Barata & Kayser, 2023; Grabowska, Saniuk & Gajdzik, 2022; Kasinathan et al., 2022). In contrast to Industry 4.0 that was very much technology-centric, both Industry 5.0 and Society 5.0 aim to adopt a more human-centric approach to the exploitation of digital technologies. Both emphasize VCC and collaboration between humans and machines. Nevertheless, we are sailing in unchartered waters, and the risk of VCD is still imminent; associated with the misuse or unintended consequences of nascent & unprecedented humans-machine interactions. The possible applications and use-cases are endless. Numerous novel value streams will emerge, and the risk of VCD from excessive VCC streams becomes more potent. Needless to say, when considering technology-based investments, decision-makers need to be conscious to the associated return-on-investment, impact on reputation value, and conformity to external norms and expectations (Polykarpou et al., 2018; Salge et al., 2022; Salge, Kohli, & Barrett, 2015).

3. Research methodology

- **Survey design**
  To ensure validity, a structured questionnaire derived from the literature will be used to collect quantitative data from a representative sample of customers across different industries.
- **Samplin**
  Using a stratified random sampling to ensure representation from diverse demographics, industries, and geographic regions.
- **Sample size**
  It will be determined as soon as the research tool & no. of variables is finalized.
- **Data collection**
  Administering the survey online or through other appropriate channels, collecting data on participants' experiences with Value Co-Creation interactions and any instances of Value Co-Destruction. Careful as to not breach any confidentiality or privacy requirement, the following parties will be approached in order to reach-out for the appropriate invitees to the survey.
  - Technology Platforms & Apps owners have information regarding the intensity and frequency of usage of their resources, per user.
  - In addition, telecom service providers have similar information regarding users who use multiple Apps and Technology Platforms.
- **Variables**
  - Dependent Variable: Occurrence of Value Co-Destruction resulting from multiple Value Co-Creation interactions.

- Independent Variables: Frequency and intensity of Value Co-Creation interactions, customer satisfaction, perceived value, trust, perceived risk, etc.

- **Data analysis**
  Utilize statistical techniques such as regression analysis to examine the relationships between variables and identify factors associated with Value Co-Destruction.

References

Agrawal, A. K., & Rahman, Z. (2015). Roles and resource contributions of customers in value Co-Creation. *International Strategic Management Review*, *3*(1-2), 144-160.

Akter, S., Babu, M. M., Hossain, M. A., & Hani, U. (2022). Value Co-Creation on a shared healthcare platform: Impact on service innovation, perceived value and patient welfare. *Journal of Business Research*, *140*, 95-106.

Alaassar, A., Mention, A. L., & Aas, T. H. (2020). Exploring how social interactions influence regulators and innovators: The case of regulatory sandboxes. *Technological Forecasting and Social Change*, *160*, 120257.

Alexander, M., & Jaakkola, E. (2015). Customer engagement behaviours and value Co-Creation. In *Customer engagement* (pp. 3-20). Routledge.

Assiouras, I., Vallström, N., Skourtis, G., & Buhalis, D. (2023). Exploring how interaction leads to value Co-Creation, value co-destruction, and value no-creation during service mega-disruptions. *Current Issues in Tourism*, 1-18.

Banik, S., & Rabbanee, F. K. (2023). Value co-creation and customer retention in services: Identifying a relevant moderator and mediator. *Journal of Consumer Behaviour*.

Barata, J., & Kayser, I. (2023). Industry 5.0–Past, Present, and Near Future. *Procedia Computer Science*, *219*, 778-788.

Barile, S., Grimaldi, M., Loia, F., & Sirianni, C. A. (2020). Technology, value Co-Creation and innovation in service ecosystems: Toward sustainable co-innovation. *Sustainability*, *12*(7), 2759.

Bassano, C., Barile, S., Saviano, M., Pietronudo, M. C., & Cosimato, S. (2020). AI technologies & value Co-Creation in luxury context.

Buhalis, D., Andreu, L., & Gnoth, J. (2020). The dark side of the sharing economy: Balancing value co-creation and value co-destruction. *Psychology & Marketing*, *37*(5), 689-704.

Čaić, M., Odekerken-Schröder, G., & Mahr, D. (2018). Service robots: value Co-Creation and co-destruction in elderly care networks. *Journal of Service Management*, *29*(2), 178-205.

Chew, E., Semmelrock-Picej, M. T., & Novak, A. (2013, January). Value Co-Creation in the organizations of the future. In *Proc. of the European conference on management, leadership & governance* (pp. 16-23).

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management science*, *35*(8), 982-1003.

Dixon, M., Freeman, K., & Toman, N. (2010). Stop trying to delight your customers. *Harvard business review*, *88*(7/8), 116-122.

Dolan, R., Seo, Y., & Kemper, J. (2019). Complaining practices on social media in tourism: A value Co-Creation and co-destruction perspective. *Tourism Management*, *73*, 35-45.

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American customer satisfaction index: nature, purpose, and findings. *Journal of marketing*, *60*(4), 7-18.

Grabowska, S., Saniuk, S., & Gajdzik, B. (2022). Industry 5.0: improving humanization and sustainability of Industry 4.0. *Scientometrics*, *127*(6), 3117-3144.

Grott, E. M., Cambra-Fierro, J., Perez, L., & Yani-de-Soriano, M. (2019). How cross-culture affects the outcomes of Co-Creation. *European Business Review*, *31*(4), 544-566.

Järvi, H., Kähkönen, A. K., & Torvinen, H. (2018). When value Co-Creation fails: Reasons that lead to value co-destruction. *Scandinavian Journal of Management*, *34*(1), 63-77.

Järvi, H., Keränen, J., Ritala, P., & Vilko, J. (2020). Value co-destruction in hotel services: Exploring the misalignment of cognitive scripts among customers and providers. *Tourism Management*, *77*, 104030.

Kasinathan, P., Pugazhendhi, R., Elavarasan, R. M., Ramachandaramurthy, V. K., Ramanathan, V., Subramanian, S., ... & Alsharif, M. H. (2022). Realization of Sustainable Development Goals with Disruptive Technologies by Integrating Industry 5.0, Society 5.0, Smart Cities and Villages. *Sustainability*, *14*(22), 15258.

Keeling, D. I., Keeling, K., de Ruyter, K., & Laing, A. (2021). How value Co-Creation and co-destruction unfolds: a longitudinal perspective on dialogic engagement in health services interactions. *Journal of the Academy of Marketing Science*, *49*, 236-257.

Kumar, S., & Kandoi, M. (2018). Co-Creation a new method of building brand loyalty & innovation. *International Journal of Management (IJM)*, *9*(6).

Lalicic, L., & Weismayer, C. (2021). Consumers' reasons and perceived value Co-Creation of using artificial intelligence-enabled travel service agents. *Journal of Business Research*, *129*, 891-901.

Li, T. M., & Wong, P. W. (2015). Youth social withdrawal behavior (hikikomori): a systematic review of qualitative and quantitative studies. *Australian & New Zealand Journal of Psychiatry*, *49*(7), 595-609.

Liu, Q., Zhao, X., & Sun, B. (2018). Value Co-Creation mechanisms of enterprises and users under crowdsource-based open innovation. *International Journal of Crowd Science*, *2*(1), 2-17.

Mantelero, A. (2013). The EU Proposal for a General Data Protection Regulation and the roots of the 'right to be forgotten'. *Computer Law & Security Review*, *29*(3), 229-235.

Marques, R. P. F., & Batista, J. C. L. (Eds.). (2017). *Information and communication overload in the digital age*. IGI Global.

Moussafir, K., & Qmichchou, M. (2021). Customer Empowerment: An antecedent and a consequence of customer engagement in value Co-Creation. *International Journal of Accounting, Finance, Auditing, Management and Economics*, *2*(5), 305-323.

Nangpiire, C., Silva, J., & Alves, H. (2022). Customer engagement and value Co-Creation/destruction: The internal fostering and hindering factors and actors in the tourist/hotel experience. *Journal of Research in Interactive Marketing*, *16*(2), 173-188.

Pera, R., Menozzi, A., Abrate, G., & Baima, G. (2021). When cocreation turns into codestruction. *Journal of Business Research*, *128*, 222-232.

Pera, R., Occhiocupo, N., & Clarke, J. (2016). Motives and resources for value Co-Creation in a multi-stakeholder ecosystem: A managerial perspective. *Journal of business research*, *69*(10), 4033-4041.

Plé, L. (2017). Why do we need research on value co-destruction?. *Journal of Creating Value*, *3*(2), 162-169.

Plé, L., & Chumpitaz Cáceres, R. (2010). Not always co-creation: introducing interactional co-destruction of value in service-dominant logic. *Journal of services Marketing*, *24*(6), 430-437.

Polykarpou, S., Barrett, M., Oborn, E., Salge, T. O., Antons, D., & Kohli, R. (2018). Justifying health IT investments: A process model of framing practices and reputational value. *Information and Organization*, *28*(4), 153-169.

Prahalad, C. K., & Ramaswamy, V. (2004). Co-Creation experiences: The next practice in value creation. *Journal of interactive marketing*, *18*(3), 5-14.

Reichheld, F. F. (2003). The one number you need to grow. *Harvard business review*, *81*(12), 46-55.

Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. In *An integrated approach to communication theory and research* (pp. 432-448). Routledge.

Salge, T. O., Antons, D., Barrett, M., Kohli, R., Oborn, E., & Polykarpou, S. (2022). How IT investments help hospitals gain and sustain reputation in the media: the role of signaling and framing. *Information Systems Research*, *33*(1), 110-130.

Salge, T. O., Kohli, R., & Barrett, M. (2015). Investing in Information Systems. *Mis Quarterly*, *39*(1), 61-90.

Skaržauskaitė, M. (2013). Measuring and managing value Co-Creation process: overview of existing theoretical models. *Socialinės Technologijos*, *3*(01), 115-129.

Solakis, K., Pena-Vinces, J., & Lopez-Bonilla, J. M. (2022). Value Co-Creation and perceived value: A customer perspective in the hospitality context. *European Research on Management and Business Economics*, *28*(1), 100175.

Vargo, S. L., & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. *Journal of marketing*, *68*(1), 1-17.

Vega-Vazquez, M., Revilla-Camacho, M. Á., & Cossío-Silva, F. J. (2013). The value Co-Creation process as a determinant of customer satisfaction. *Management Decision*, *51*(10), 1945-1953.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.

Zhang, T., Lu, C., Torres, E., & Chen, P. J. (2018). Engaging customers in value Co-Creation or co-destruction online. *Journal of Services Marketing*, *32*(1), 57-69.

Zine, P. U., Kulkarni, M. S., Chawla, R., & Ray, A. K. (2014). A framework for value Co-Creation through customization and personalization in the context of machine tool PSS. *Procedia CIRP*, *16*, 32-37.

The false promise of the app economy. A study of how enterprise apps are used in the modern workplace (2017). Harmon.ie.

27. The Role of Information Systems Strategy in Driving Successful Digitalization: an IoT perspective

Hany Abdelghaffar
The German University in Cairo
hany.ismail@guc.edu.eg

Mona Ali Ali
The German University in Cairo
mona.ali@guc.edu.eg

Mohamed Sherif
The German University in Cairo
mohamed.sherif@student.guc.edu

Abstract

The Internet of things (IoT) has become one of the most significant topics in information systems (IS) industry yet; it still faces certain complications in its implementation process. The estimation of unsuccessful implementations are up to 60% Cisco (2017). Unsuccessful implementation could be due to many issues such as human resources, collaboration between information technology function (IT) and other business functions, cultural issues within organizations, etc... However, these issues are overcome with an appropriate IS strategy Arora et al. (2017). Although efforts were exerted to understand the factors of successful adoption of IoT, less efforts have been directed towards understanding these factors from an information systems strategy point of view. This study aims to identify the most important driving factors in an information systems strategy leading to successful implementation of IoT. A preliminary framework presenting the 10 IS driving factors leading to successful implementation of IoT is deduced from the literature. The developed framework is tested in two case studies within leading real estate developers who have successfully implemented IoT in various functions. Face to face semi structured interviews were administered and the results revealed that six main factors and nine sub factors were required to guide the IS strategy to successful adoption of IoT projects within an organizational context.

Keywords: IoT, IS strategies, Case study, Driving factors

1. Introduction

The concept of IoT was first proposed in 1999 by Kevin Ashton (Elgazzar, et al., 2022). IoT devices are a multitude of physical objects, equipped with sensors, actuators, and/or computing power connected to the Internet via communication technology, and enabling interaction with and/or among those objects (Bayer, Gimpel, & Rau, 2020). These sensors can listen and talk in ways that humans are incapable of doing which will be quite enriching for humanity at large. Sensors within an IoT system would enable companies to save money, forecast weather, predict events, and will provide an overall better environment for businesses (Chen, 2012). IoT devices include all devices that can connect to the internet including toys, mobile phones, home appliances and many others (Green, 2015). The connection to internet is either wired or wireless (Yang et al., 2013). Devices need an open standard to ensure the integration of different devices from different vendors for communication (Andersson & Mattsson, 2015).

International Telecommunication Union (ITU) defines the Internet of Things as "a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies" (Chehria, Zimmermann, Schmidtc, & Masuda, 2021).
Although, IoT has become one of the most significant topics in information systems industry, it still faces certain complications in its execution process. The estimation of unsuccessful experiments are up to 60% (Cisco, 2017). It is mainly due the shortcomings of human resources, collaboration between IT and business side, cultural issues within an organization, etc. (Cisco, 2017). A research on Telco organizations concluded that to succeed in IoT implementation, organizations must possess a strategic focus to define effective IS methods for its IoT (Arora et al. 2017). This evoked managements to tackle the use of IoT from a different perspective and that is through the IS strategies itself (Inoue, Takenaka, Kasasaku, Tamegai, & Arai, 2023). In addition, IoT is a new technology for most organizations hence, the outcome cannot be properly evaluated (Ahsan et al., 2016). This creates a scarcity of data crippling many businesses in implementing IoT (Kamin, 2017). Accordingly, the main motivation of this research stems from this gap. Accordingly, the research answers the question: *"how could IS strategy drive successful implementation of IoT?".* In order to answer this question two objectives are attempted:

- *Propose a strategic framework for the successful implementation of IoT*
- *Test the strategic framework for validation of factors*

2. Literature review

There is no consensus in literature about the definition of IS strategy, in fact the term is somewhat interchangeable with Information Technology (IT) strategy. Other scholars think that the two terms (IS and IT) should be replaced with a more agile terminology such as digitalization strategy. On the other side, there is consensus between scholars about why IS/IT strategy exists. The main purpose of IS/IT strategy is to answer the questions who is responsible for IS/IT functionality and how are they guiding the IS/IT towards achieving business strategies (Teubner & Stockhinger, 2020).

There are different success factors for adoption and implement IOT. The literature focuses on interdepartmental collaboration, building technology focused culture (from top–down the organization), acquire IoT expertise (internal or external) and, building an engaging ecosystem with IoT partners at every stage of the IoT implementation. Furthermore, It has been found that organizational realignment is one of the key factors in the success rate of IoT field (Arora et al., 2017). Accordingly, it is important to identify which parts of the IS strategy needs to be updated with the IOT components and architecture. The following ten factors were identified within the literature.

Organization structure

IoT equipment connected affects all the functions within the organization, from IT, R&D to sales, marketing, development, supply chain, human resources and many others (Brou, Janssen, & Herder, 2020). Recent researches identified the collaboration between the organization business units for successful IoT services. In big organizations, however, it is difficult to

change the structure so it can build the IoT business unit inside the organization to achieve this target (Arora et al., 2017).

Organization structure

IoT equipment connected affects all the functions within the organization, from IT, R&D to sales, marketing, development, supply chain, human resources and many others (Porter & Heppelmann, 2014). While recent researches identified the collaboration between the organization business units for successful IoT services (Brou, Janssen, & Herder, 2020). In big organizations, however, it is difficult to change the structure so it can build the IoT business unit inside the organization to achieve this target (Arora et al., 2017).

Human resources

IoT is a new field of study and application, which means limitation in finding proper talented staff (Arora et al., 2017). Therefore, organizations found it an essential element to conduct in-house training to equip its employees or hire already trained and talented personnel from the global market Arora et al. (2017). Further, the talents should have different skills sets as opposed to previous ones (Porter & Heppelmann, 2014). According to Kranz (2017), in sum, proper application of IoT entails new technical skills and collaboration both internally and externally.

Alignment with business

IoT is more than connecting devices it is, also, about connecting customers, suppliers, and partners (Kranz, 2017). Therefore, organizations need to consider IoT as a more collaborative approach (Brou, Janssen, & Herder, 2020). Amata-McCoy and Deena (2016) recommended defining business needs and build IoT solutions around them too.

Innovation

Innovation is a vital part of IoT as it creates value. The IoT can help firms to innovate and to address new business opportunities. However, many companies face difficulties in developing products and services based on the IoT with a clear and valid value proposition (Molling & Klein, 2022). End users are important source of innovation thus, organizations may need to build an innovative culture with focus on the outcome, encourage success, tolerance failure, the motivation to follow through on newly generated idea (Arora et al. 2017).

IS Infrastructure

Communication is one of the critical features of the IoT as all the devices should be connected to perform specific tasks like sensing, communicating, and information processing (B.B. Gupta, 2018). Therefore, any infrastructure setup must consider the operational requirements of IoT which require organizations to build entirely new infrastructure of multiple layers (Technology stack). It consists of the devices layer where the IoT hardware components and software to manage the functionality of the IoT hardware (Porter & Heppelmann, 2014). The connectivity layer links the different things and communicates to other layers using communication protocols. At the cloud layer management software and IoT application for monitor and control, analytics software, and to analyze data gathered (B.B. Gupta, 2018). In addition, Information Security is considered at each layer of the architecture. This architecture gives IoT solutions capabilities which allow them report about their status and the surrounding

environment providing data about their performance and use. Also, allows remote users to manage and customize complex operations of the system (Porter & Heppelmann, 2014).

Business process

The application of IoT is to improve the efficiency of the business processes and automate inter-organization integration (Acaulay et al., 2015). Also, (Soori, Arezoo, & Dastres, 2023) confirmed that adopting the IoT advances business process. IoT can support organizations to further the business process redesign (BPR) by process standardization, automation, empowerment of employees, creating values, information analysis, process automation and control Ferretti & (Schiavone, 2016).

Partnership with vendors

It is not only around building partnership for building IoT solutions but, also, bringing innovative ideas from the eco-system like open source communities, developers, partners, etc. (Arora et al., 2017). According to Cisco's survey results, organizations with the most successful IoT projects, focused strongly and build partnership with their eco-system at every stage of the project (Cisco, 2017).

Management strategic intention and support

The top management decision (strategic intention) is very important to the IS strategy formulation and IoT adoption; either the management requires the organization to be a leader in the IoT context (Innovator) or follower (Conservative) (Garrett et al., 2009). Thus, if the decision is to be innovator in the IoT market; organizations need to locate themselves as the first mover through developing innovative applications and gain an innovative reputation (Frynas et al. 2006). However, during this time organizations decide to be followers (Conservative) in the IoT market, study the market demand and copy successful market leaders in the IoT Covin et al. (2000). Moreover, two more factors (security and privacy) emerged from the literature about the IoT which states it also requires focus on the IS strategy. Security and privacy are important factors that are affecting the success of IoT and hence need consideration in the proposed framework.

Security and privacy

IoT security and privacy are main issues which prevent organizations from implementing IoT (Kim et al., 2015). The problem is IT leaders lack the knowledge of the IoT security and privacy strategies to enable the implementation of IoT. Inadequate IoT security resulting from extending the traditional IT security architecture is considered a high risk for organizations implementing IoT especially on a large-scale implementation (Chehria, Zimmermann, Schmidtc, & Masuda, 2021). Hence, in order to protect the organization, strong security services such as cloud services should be considered while developing IoT solutions (Vermesan & Friessm2014). It should secure the IoT architecture on all layers (from application to end nodes). On the other hand, privacy issues with IoT should be considered. IoT solutions have the ability to gather massive amounts of data that can breach the privacy of the organization and everyone within it. Sharing unauthorized access or unnecessary collection of personal information exposes personal privacy and puts it at risk Miessler & Smith (2015). Thus, organizations willing or aspiring to adopt IoT solutions are required to explore security and privacy strategies to assert that data is secured and address privacy issues Kamin (2017).

Agility

To be flexible in implementing the requirements of the customer, which means that organizations should have processes and tools to collect feedback from customers Arora et al. (2017). Therefore, the adoption of agile processes in IoT projects can lead to more customer satisfaction (Bernsteiner, Hall, & Fruhling, 2023).

3. Proposed Framework

Resulting from literature analysis, the proposed framework summarized in figure 1 consists of 10 factors that need to be considered for the IS strategy in organizations planning to adopt and implement IOT.

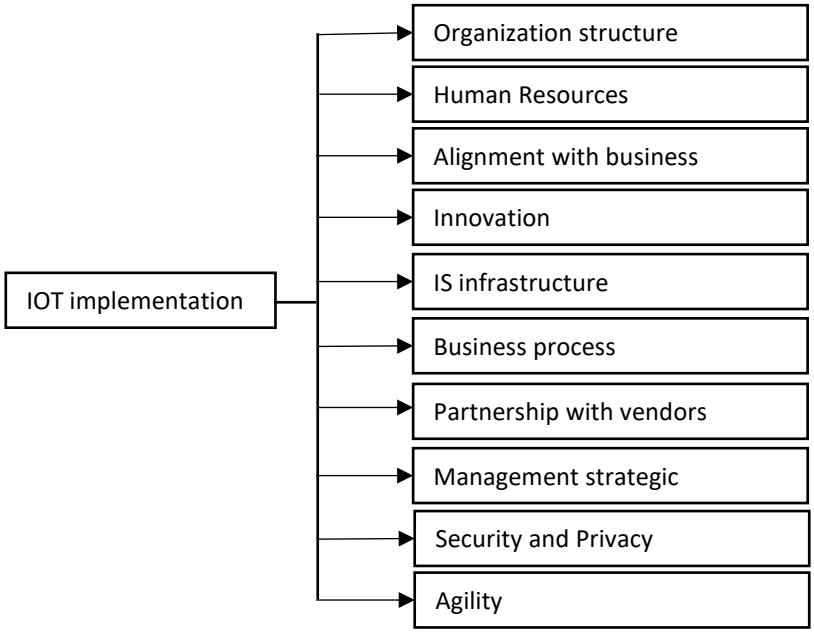


Fig. 1: Proposed Framework: The IS strategy Factors required for IoT implementation

4. Methodology

In order to reach the objectives of this study, a case study is used as the main method to direct the research. Case studies allow the researcher to understand the phenomena under study from the participants' point of view Boblin et al. (2013). Hence, interviewing the participants about their point of view regarding IS strategy, its impact on IoT, and understanding their experience help in unveiling this phenomenon. This requires deep understanding of the organization's IS strategy in the selected cases which will help to answer the research question. The focus of this case study is to explore IS factors necessary for successful adoption of IoT. Consequently, confirming the proposed framework that can work as a guideline for organizations planning to implement IoT in their organizations

4.1. Case selection

This research primarily focuses on two real estate organizations operating in Egypt. The criteria for selecting the organization under study was that the organization should have implemented

IoT solutions, consider IoT in the IS strategy. In addition, the availability of data and ability to conduct interviews, access to organization documents. A review of the current real estate market in Egypt demonstrates that few organizations have implemented IoT and are considering changes in the IS strategy to enable the adoption of IoT. Organizations were selected as they have completed the implementation of successful IoT projects. In addition, management in the organization took some steps in the IS strategy formulation to enable the adoption of IoT.

Company A: the first organization is an international organization and its head office located in Dubai. It has a branch in Cairo as well. Its project in Cairo consists of mixed-use project of: residential area, commercial offices, retail mall, educational institute, hospitality, automotive area, entertainment area, and utilities area like district cooling plant, sewage station, electrical substation, and water station. The organization also has implemented IoT in the different areas in their project; some of which implemented IoT projects are sensors installed on most of the utilities within the project to monitor the equipment and send notifications whenever there is an issue that requires attention. In some cases, the application managing the sensors is smart to take corrective actions and report them; which lead to saving time, energy, reduce the down time. Another example is free Wi-Fi inside the mall. The Wi-Fi access points and the mobiles acts as IoT devices that is able to collect user's data and send it to analytics software that analyzes the customer behavior and provide management with dashboards about the rush hours, number of visitors, footfalls heatmap, most visited stores and many others. This helps in designing marketing campaigns to increase retailers' revenue and leads to more customer satisfaction. One of the selected organizations implemented a unified application that is integrated with all the IoT devices and imbedded software in one dashboard that monitors all the aspects inside the projects and is capable of collecting data from all the devices into one platform. This application provides a dashboard for IoT operations and can respond to any issue either by sending notifications or take actions and reports back the data to head office for further data analytics. This application is open platform that can integrate with any other platforms or business applications, so data can be shared among other applications. The IoT projects automate some business processes like preventive and corrective maintenance, customer portal to request services from city management.

Company B: The second organization is local organization, which has many projects across Egypt including residential, commercial and planning for retail malls. This organization is still evolving in the field of IoT. The organization has implemented some IoT solutions in some areas in their projects, like smart homes, IPTV (Internet Protocol Televisions) that is integrated with the CRM (Customer Relationship Management) application and ERP. Security cameras CCTV (Closed Circuits Televisions) that connected to the internet and integrated with the access control. Moreover, to ensure success, both organizations management made some changes in the IS strategy to effectively implement IoT projects.

4.2. Data collection techniques

The characteristic of the population in the case study is related to the participants experience with the phenomena Berger (2015). To define the study population, inclusion and exclusion criteria can be used (Robinson, 2014). Thus, the population of this study is all stakeholders who take decisions about the IS strategy or management who have experience with IoT

projects. The population included the IT directors, IT managers, city management, general manager, country IT manager, managing director and IT infrastructure manager. To select appropriate samples, it was important to select participants who have best knowledge about the research topic (Kish & Verma, 1986). In this study the selection criteria for participants included:

- Currently working within the selected organizations
- Working in managerial positions
- Working with IT department that have experiences in IoT projects
- Participated in IS strategy formulation and decisions making

A suitable sample size is related to the study's data saturation. The degree of data saturation could be reached by conducting in-depth interviews from six to twelve interviews Guest et al. (2006). Also, interviewing participants who have had direct contact and knowledge of the phenomena can help in the reduction the sample size required for data saturation. Semi-structured face to face interviews are the most effective tool for information gathering in qualitative research as it is very flexible in designing, refining and conducting the interview Irvine et al. (2013). Also, using the same set of questions in face-to-face interviews can lead to data saturation Fusch & Ness (2015). Face-to-face semi-structured interviews was used as the data gathering technique until no new information or themes is available to reach data saturation. Some organizational secondary documents were referred to in order to cement the interviewees' claims.

5. Data Analysis

Content analysis is used to analyze the data collected from interviews and organization documents. The analysis of data began by searching the data collected for answers to the research question on how IoT affects the IS strategy. Data categorized into themes and developed understanding on the IS strategy factors that is considered by the participants and impacted by the adoption of IoT from the participants prospective.

Coding is used to categorize data into small meaningful chunks Leech & Onwuegbuzie (2007). Coding, also, allows searching for explanations, patterns, relationships of the data. During the data analysis, following these steps:

- Listened to the interview recording, reading scripts, and review all organization documents
- Generated list of codes representing the data and repeated the process until no new codes identified.
- Organized data and place them in tabular form.
- Used codes to search for and identify themes, patterns and relationships in alignment with the research question.
- Categorized the codes and determine major themes.
- Repeated the above steps until no new themes and codes were found and had explanation for the research question

Categorizing data helps researchers correlate the major themes in the literature and the proposed framework and answer the research question Boblin et al. (2013). Data is searched

for patterns and recurring themes that are related to the IS strategy factors that were identified from the literature and formulated the proposed framework Bansal & Corley (2011).

6. Findings and Discussion

The purpose of this study is to provide the management of organizations planning to implement IoT with a guideline for the factors required in IS strategy to implement successful IoT projects. The first objective was achieved since a proposed framework was developed using factors found in the literature. The validation of these factors was concurred through two case studies. The results show the below findings.

6.1. Security and privacy factor

The security and privacy factors were the most important factor for all the participants as it could hinder the implementation of IoT. Whatever the technology used for security and privacy; it should be considered in the IS strategy. According to, the security issue in IoT needs the collaboration of all stakeholders from companies involved in IoT by using experts to integrate security best practices with IoT products development, educating customers and end user with policy and procedures. Hence, the results of the case studies were in alignment with the literature. Security and privacy are a significant factor impacting IoT implementation and should be considered while adopting IoT. In addition, security and privacy policies should be in place as well as educating users about the security policies.

6.2. Organization structure factor

Organization structure was confirmed to be an important strategic factor to be considered when implementing IoT. Some organizations proposed new roles responsible for IoT based on business requirements. Others recommended reshaping the different departments to deal with the IoT. Only one participant mentioned that organizational structure is not required for the IoT. This is in accordance to Arora et al. (2017), where larger industries with more complex organization structures require building IoT business units within the organization to achieve its targets. The human resources planning had different point of views in the participants' responses. Some participants recommended the outsourcing of skilled human resources and others recommended up-skilling in-house resources due to lack of skilled resources in the Egyptian market. However, there is an agreement that planning for skilled human resources in the IS strategy, either internal or external and transfers knowledge to support the IoT is required. Cisco survey about the success factors for IoT demonstrated that one of the success factors is to have IoT expertise either internal or external Cisco (2017). For IoT, organizations require the development of necessary skills in-house and/or acquire the talents from the global market Arora et al. (2017). This confirms that human resources are a significant factor and requires consideration in the IS strategy as it is impacted by the IoT implementation.

6.3. Alignment with business factor

The alignment with business strategy factor was also confirmed from the analysis as it was required to meet business needs. Hence, managements should build their IS strategy in alignment with the business needs and goals. Recent researches identified collaboration between the organization businesses units as a needed factor for successful IoT services Neu & Brown (2005). Thus, the alignment is obligatory between the IoT projects and the business requirements. As for innovation, it was found that building an innovative culture within the

organization is a strategic factor required for the success of IoT implementation. The responses recommended encouraging innovation by building reward systems, plan events for innovation such as hackathons. This is in alignment with the literature, which highlights the importance of building innovation culture with focus on the outcome, encouraging success, tolerating failure, encouraging idea generation Arora et al. (2017).

6.4. IS resources and infrastructure

In reference to the IS resources and infrastructure factor, most of the participants agreed that planning for the resources and infrastructure should be considered in the IS strategy with all its components. Even most of participants proposed layered structure to be considered for the IoT; which aligned with the literature review as IoT requires organizations to build entirely new infrastructure of multiple layers (Porter & Heppelmann, (2014). Five layers of Infrastructure is proposed. 1) The hardware is one of the IoT infrastructure layers confirmed by participants that it is composed of sensors that can collect and transfer the data through the network layer. IoT hardware layer and sensors are able to collect and exchange data from the environment and between devices. This layer should be able to communicate with each other and save energy (B.B. Gupta, 2018). Some participants recommended that the hardware should be smart to take decision independently using imbedded software based on specific actions. Hence, planning for the IoT hardware in the IS strategy is required to enable the adoption of the IoT as it is impacted by the IoT implementation in alignment with the literature review. 2) The network is another component of the IoT infrastructure which created a debate between participants either to implement new networks for the IoT or to use the existing network if it can handle the traffic of IoT. However, in general, there is an agreement that network is necessary for the adoption of IoT. Thus, proper planning is required for the network either a new or an existing one that can accommodate the requirements of the IoT operations. Network is one of main features of the IoT that the hardware must be interconnected to perform its tasks (B.B. Gupta, 2018). Hence, the network also was confirmed that it is significant for IoT and planning for it should be considered in the IS strategy for current and future projects. 3) The data is one of the IoT infrastructure layers that need consideration in the IS strategy. From the analysis there is almost agreement that data is very important in IoT which requires gathering, and analytics. As IoT generates massive amounts of data, hence using big data technology is mandatory for IoT. Using IoT increases the sources of data, the IoT device themselves, which generate real-time and huge amounts of structured and unstructured data collected from different applications (Porter & Heppelmann, 2014). In addition, the IoT data is not useful as long as it is not analyzed according to Intel report (Chen et al., 2014). That was aligned with the participants' answers to obtain valuable data. 4) Application is the fourth layer of the IoT infrastructure. Most of the participants' recommended unified software that can monitor the IoT solution, integrate all the IoT systems, take decisions in some situations and automate some of the business processes. Few participants recommended that IoT end devices should be smart and have imbedded software that can take decisions and perform actions based on data collected. Yet, in general planning for the applications (scattered, unified or both) is important. In addition, (Truong & Dustdar, 2015) views that the IoT software is scattered and recommends a unified application to integrate all the software in one platform so that it will work. Planning the IS strategy should be for smart applications or software that is capable of analyzing the data, making decision, and performing actions based on the data collected. 5) The cloud emerged in the analysis as integral part of the IoT infrastructure especially with hardware, data, and applications layers.

It was recommended to host all these layers to benefit from the computing power and massive storage of the cloud. However, the cloud was not significant for some of the participants in the current stage, because the existing on-premises infrastructure can accommodate the IoT. Organizations also invested huge amounts on the on-premises infrastructure, thus moving existing infrastructure to cloud is not required in the current stage. However, most of participants agreed that planning for cloud is important for the future. This contradicted with the literature as some IoT operators realized that the success of IoT depends on the integration between big data and cloud computing (Chen et al., 2014). IoT changes the demand for more data storage, higher management and processing power. The cloud services are flexible to deliver the requirements for IoT and to accept the increasing workload of IoT (Porter & Heppelmann, 2014). Hence, the cloud was not significant and was not confirmed for the proposed framework. However, most of participants recommended planning for the cloud for the future, hence it can be considered in the IS strategy as future plan.

6.5. Business process

As for the business process factor, IoT can help organizations in the business process enhancement, automation and can support business process reengineering. Providing efficiency for the business process and automate some actions without human interference. The application of IoT improves the efficiency of the business process and automates inter-organization integration (Acaulay et al., 2015). It also creates values, enables information analysis, process automation and control (Ferretti & Schiavone, 2016). In addition, some participants recommended a standardized business processes required for enhancement. The standardization of organization processes enables process redesigning (BPR), automation and empowerment of employees. Hence, this factor is significant from the analysis and from the literature review and again it was confirmed in the proposed framework.

6.6. Building partnership with vendor

Although points of view vary around the purpose of the partnership, but it was confirmed that building partnership is required and management should consider it in the IS strategy. According to Cisco survey results, organizations with the most successful IoT projects, focused and strongly built partnership with their eco-system at every stage of the project (Cisco, 2017). Participants agreed that partnership is essential to either implement IoT solutions, bring innovative ideas or to have skilled human resources for IoT projects. In addition, review of the literature confirmed that partnership is not only around building IoT solution but also brining innovative ideas from the eco-system (Arora et al., 2017). Hence, both participants responses and literature review were aligned which confirmed the partnership with vendors as one of the IS strategy factors that is impacted by the IoT.

6.7. Management intention and support

The management intention and support for IoT factor was confirmed by all of the participants. As maintaining the management support for technology and innovation in general and the IoT projects in specific is mandatory for the success of IoT. Thus, obtaining management support on the IS strategy in regards to IoT is important. This is also aligned with the literature review as the top management strategic intention is very important to the IS strategy formulation and IoT adoption whether the management requires the organization to be a leader in the IoT context (Innovator) or follower (Conservative) (Garrett et al., 2009).

6.8. Agility

The last confirmed factor was agility. From the analysis, agility was important for managing the IoT projects as it maintains flexibility in responding to the customer requirements and collecting their feedback. There was agreement between participants about using agile methods in implementing IoT projects within their organizations. This will lead to more customer satisfaction and successful IoT projects that meets business needs. This result is in accordance with the literature, which proposes that the organizational agility could lead to more customer satisfaction (Upadhyay et al., 2015). (Bernsteiner, Hall, & Fruhling, 2023) recommend shift from traditional planning for more flexible IS Strategy planning (Agile) that maximize the business benefits from technology which is also confirmed the agility as one the proposed frame factors.

The analysis revealed the below framework as a comprehensive roadmap to successful adoption of IoT in organizations thus, providing factors necessary within an IS strategy to adopt and implement IoT projects within an organizational context:

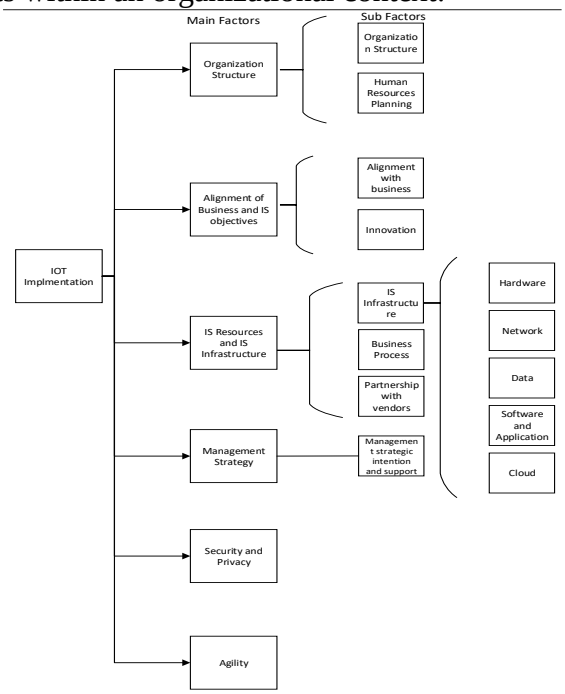


Fig. 2. Final Framework: The IS strategy Factors required for successful IoT implementation

7. Implications

7.1. Theoretical Implications

The rich data in case study can provide new ideas in theory building Siggelkow (2007). The findings of this study contribute to the understanding of the impact of the IoT implementation on the IS strategy factors. This study also contributes in answering the question "how could IS strategy drive successful implementation of IoT?" The proposed framework highlights the changes required in the IS strategy factors by organization management to enable the adoption of IoT. The findings of the case study suggest a framework for IS strategy factors enabling

successful adoption of IoT. There is a dearth of literature within this domain as outlined in section 2.0 especially for development purposes. Thus, this study aims to fill this gap and possess a consolidated view of the IS strategy factors that need consideration by top managements to enable the adoption of IoT.

7.2. Practical Implications

One of the practical contributions of this research is the detailed insight provided by the case study. The case study reveals that organizations' IT management and top management should refer to the proposed framework for the IS strategy planning to adopt the IoT. The information gained from this framework enables management in organizations to ensure appropriate processes and strategy are in place prior to IoT adoption. This study's findings were significant to business practices in several ways. The research highly recommends setting a security and privacy policies for IoT and considers them with all the levels of the IoT projects and layers regardless of the technology used. It also, demonstrates that it is not necessary to change any organizational structure to have successful IoT implementation, unless the organization does not have the required skilled resources to implement IoT. Hence, planning to develop the internal skills or acquiring them from the market is necessary. This research also demonstrates that it is required to understand the business needs before implementing of IoT hence add value to the business. Furthermore, it is imperative to encourage innovation within organizations in order to implement innovative IoT projects.

The research contributed in the planning of the IS resources and infrastructure with its different components from hardware, network, data, software, and business processes to adopt technologies that support the IoT operations and enable make use of the IoT by implementing data analytics and decision-making applications. Also, encourage building relationships with IoT partners to benefit from their experience in implementing the IoT and transfer knowledge. Using agile methods for the IoT is favorable than using traditional methods for the IoT projects to have the flexibility for the rapid changes in technology and meeting the customer needs.

8. Limitations and Future Work

The limitation of this research is due to limited number of case studies, which is related to two organizations working in the real estate industry. This may have led to participant bias based on their experiences in the company. Future research should focus on different industries align there IS strategy with IOT adoption and implementation. In addition, the data gathered for this study was cross sectional. An analysis on how these factors affect the adoption of IoT on a longitudinal study will enhance the understanding of the dynamic relationships between these factors and the phases of IoT implementation.

9. Conclusion

The purpose of this study was to explore IS strategy factors that influence the implementation of IoT projects. A framework was proposed based on the factors found in the literature. A case study research was conducted within two organizations in the real estate field to confirm the framework. After analyzing the collected data, six factors and nine sub factors were found to be essential for successful implementation of IoT projects.

References

Ahsan, M., Talib, M. R., Sarwar, M. U., Khan, M. I., & Sarwar, M. B. (2016). Ensuring interoperability among heterogeneous devices through IoT middleware. International Journal of Computer Science and Information Security, 14(4), 251.

Amata-McCoy, & Deena M., (2016). A retail journey. Retrieved from Chain Store Age: https://www.chainstoreage.com/news/retail-journey/

Andersson, P., & Mattsson, L. G. (2015). Service innovations enabled by the “internet of things”. IMP Journal, 9(1), 85-106.

Arora, P., Mawa, R., Diwvedi, V., Kathuria, K., & Sharma, N. (2017). IoT Strategy for Telcos: Critical Success Factors to Win in the Changing Business Landscape. Telecom Business Review, 10(1), 31.

Bansal, P., & Corley, K. (2011). The coming of age for qualitative research: Embracing the diversity of qualitative methods. Academy of Management Journal.

Bayer, S., Gimpel, H., & Rau, D. (2020). IoT-commerce - opportunities for customers through an affordance lens. Electronic Markets, 31, 27-50.

B.B. Gupta, M. Q. (2018). An overview of Internet of Things (IoT): Architectural aspects, challenges, and protocols. Concurrency and Computation: Practice and Experience, 32(21).

Berger, R. (2015). Now I see it, now I don’t: Researcher’s position and reflexivity in qualitative research. Qualitative Research, 15(2), 219-234.

Bernsteiner, R., Hall, M., & Fruhling, A. (2023). Enhancing IoT Project Success through Agile Best Practices. ACM Transactions on Internet of Things, 4(1), 1-31.

Boblin, S. L., Ireland, S., Kirkpatrick, H., & Robertson, K. (2013). Using Stake’s qualitative case study approach to explore implementation of evidence-based practice. Qualitative Health Research, 23(9), 1267-1275.

Brou, P., Janssen, M., & Herder, P. (2020). The dual effects of the Internet of Things (IoT): A systematic review of the benefits and risks of IoT adoption by organizations. International Journal of Information Management, 51.

Brown, I. T. (2004). Testing and extending theory in strategic information systems planning through literature analysis. Information Resources Management Journal, 17(4), 20-48.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. Mobile Networks and Applications, 19(2), 171-209.

Chehria, A., Zimmermann, A., Schmidtc, R., & Masuda, Y. (2021). Theory and Practice of Implementing a Successful Enterprise IoT Strategy in the Industry 4.0 Era. 25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2021). Szczecin, Poland: Elsevier.

Chen, Y. K. (2012). Challenges and opportunities of internet of things. In Design Automation Conference (ASP-DAC), 17th Asia and South Pacific (pp. 383-388). IEEE.

Cisco. (2017). Cisco Survey Reveals Close to Three-Fourths of IoT Projects Are Failing. Retrieved from https://newsroom.cisco.com/press-release-content?articleId=1847422

Covin, J. G., Slevin, D. P., & Heeley, M. B. (2000). Pioneers and followers: Competitive tactics, environment, and firm growth. Journal of Business Venturing, 15(2), 175-210.

Elgazzar, K., Khalil, H., Alghamdi, T., Badr, A., Abdelkader, G., Elewah, A., & Buyya, R. (2022). Revisiting the internet of things: New trends, opportunities and grand challenges. Front. Internet Things, 1- 18.

Ferretti, M., & Schiavone, F. (2016). Internet of Things and business processes redesign in seaports:The case of Hamburg. Business Process Management Journal, 22(2), 271-284.

Flowers, S., Mateos-Garcia, J., Sapsed, J., Nightingale, P., Grantham, A., & Voss, G. (2008). The New Inventors: How users are changing the rules of innovation. National Endowment for Science, Technology and the Arts (NESTA).

Frenken, T., Spiess, P., & Anke, J. (2008). A flexible and extensible architecture for device-level service deployment. European Conference on a Service-Based Internet (pp. 230-241). Springer, Berlin, Heidelberg.

Frynas, J. G., Mellahi, K., & Pigman, G. A. (2006). First mover advantages in international business and firm-specific political resources. Strategic Management Journal, 27(4), 321-345.

Fusch, P. I., & Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. The Qualitative Report, 20(9), 1408-1416.

Garrett, R. P., Covin, J. G., & Slevin, D. P. (2009). Market responsiveness, top management risk taking, and the role of strategic learning as determinants of market pioneering. Journal of Business Research, 62(8), 782-788.

Greene, J. (2015). TIM Lecture Series-The Internet of Everything: Fridgebots, Smart Sneakers, and Connected Cars. Technology Innovation Management Review, 5(5), 47.
Guest, G., Bunce, A., & Johnson, L. (2006). How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. Field Methods, 18(1), 59–82. https://doi.org/10.1177/1525822X05279903
Inoue, Y., Takenaka, T., Kasasaku, T., Tamegai, T., & Arai, R. (2023). How to design platform ecosystems by intrapreneurs: Implications from action design research on IoT-based platform. Electronic Markets , 33(9).
Kamin, D. A. (2017). Exploring security, privacy, and reliability strategies to enable the adoption of IoT (Order No. 10640866). Retrieved from ProQuest Dissertations & Theses Global. (1972093788). Retrieved from https://search.proquest.com/docview/1972093788?accountid=120728
Keri, E., Pearlson, S., & Saunders, C. (2006). Project Management. In Managing and Using Information Systems: A Strategic Approach. Wiley.
Kim, H., Lim, J., & Lee, K. (2015). A study of K-ISMS fault analysis for constructing secure internet of things service. International Journal of Distributed Sensor Networks, 11(9), 474329.
Kranz, M. (2017). Success with the Internet of Things Requires More Than Chasing the Cool Factor. Harvard Business Review Digital Articles, 2–5.
Leech, N. L., & Onwuegbuzie, A. J. (2007). An array of qualitative data analysis tools: a call for data analysis triangulation. School Psychology Quarterly, 22(4), 557.
Lim, M. K., Bahr, W., & Leung, S. C. (2013). RFID in the warehouse: A literature analysis (1995–2010) of its applications, benefits, challenges and future trends. International Journal of Production Economics, 145(1), 409-430.
Liu, P., & Peng, Z. (2014). China's Smart City Pilots: A Progress Report.Computer, (10), 72-81.
Irvine, A., Drew, P., & Sainsbury, R. (2013). 'Am I not answering your questions properly? Clarification, adequacy and responsiveness in semi-structured telephone and face-to-face interviews. Qualitative Research, 13(1), 87-106.
Macaulay, J., Buckalew, L., & Chung, G. (2015). Internet of Things in Logistics: A collaborative report by DHL and Cisco on implications and use cases for the logistics industry,". DHL Trend Research and Cisco Consulting Services.
Miessler, D., & Smith, C. (2015). Top 10 2014-I5 Privacy Concerns. Retrieved from https://www.owasp.org/index.php/Top_10_2014-I5_Privacy_Concerns
Molling, G., & Klein, A. Z. (2022). Value proposition of IoT-based products and services: a framework proposal. Electronic Markets, 32, 899-926.
Mytelka, L. K. (2000). Local systems of innovation in a globalized world economy. Industry and Innovation, 7(1), 15-32.
Neu, W. A., & Brown, S. W. (2005). Forming successful business-to-business services in goods-dominant firms. Journal of Service Research, 8(1), 3-17.
Porter, M. E., & Heppelmann, J. E. (2014). How smart, connected products are transforming competition. Harvard Business Review, 92(11), 64-88.
Ruiz-de-Garibay, J., Campo, T., Álvarez, M., & Ayerbe, A. (2011). Flexible and agile architecture for Internet of Things gadgets. RFID-Technologies and Applications (RFID-TA), 2011 IEEE International Conference on (pp. 50-55). IEEE.
Siggelkow, N. (2007). Persuasion with case studies. Academy of Management Journal, 50(1), 20-24.
Soori, M., Arezoo, B., & Dastres, R. (2023). Internet of things for smart factories in industry 4.0, a review. Internet of Things and Cyber-Physical Systems, 3, 192-204
Teubner, R. A. & Stockhinger, J. (2020) Literature review: Understanding information systems strategy in the digital age. The Journal of Strategic Information Systems. 29(4)
Trkman, P. (2010). The critical success factors of business process management. International Journal of Information Management, 30(2), 125-134.
Truong, H. L., & Dustdar, S. (2015). Principles for engineering IoT cloud systems. IEEE Cloud Computing, 2(2), 68-76.
Turner, M. (2015). How to secure the internet of things. Retrieved from Computer Weekly: https://www.computerweekly.com/opinion/How-to-secure-the-internet-of-things
Upadhyay, P., Matharu, G., & Garg, N. (2015). Modeling agility in Internet of Things (IoT) architecture. Information Systems Design and Intelligent Applications (pp. 779-786). Springer, New Delhi.
Vermesan, O., & Friess, P. (Eds.). (2014). Internet of things-from research and innovation to market deployment (Vol. 29). Aalborg: River publishers.

Vishvakarma, N. K., Sharma, R. R. K., & James, W. (2015). Relating 'Internet of Things' (IoT) Architectures to Strategy Types of Organizations: A Conceptual Framework. Journal of International Management Studies, 15(1).

Ward, J., Griffiths, P. M., & Whitmore, P. (2002). Strategic Planning for Information Systems (Vol. 3). Chichester: Wiley.

Yang, L., Yang, S. H., & Plotnick, L. (2013). How the internet of things technology enhances emergency response operations. Technological Forecasting and Social Change, 80(9), 1854-1867

28. A Student-Dominant View of the Readiness to Use Metaverse in Education
The TRI-F Framework

Malcolm Garbutt
University of the Western Cape
mgarbutt@uwc.ac.za

Ilhaam Ismail
University of the Western Cape
3694782@myuwc.ac.za

Raeez Adams
University of the Western Cape
4072419@myuwc.ac.za

Calvineo Juries
University of the Western Cape
3945939@myuwc.ac.za

Abstract
This paper reports on students' readiness to use Metaverse for Education in a university in a developing country facing infrastructure and poverty challenges. COVID-19 forced many universities to adopt a hybrid approach to teaching and supervision. While online meeting technologies have become commonplace, there is a lack of connectedness in face-to-face meetings, for which Metaverse is promoted as a solution. We pose the question of students' level of readiness to use Metaverse technologies. Thematic analysis of students' self-reflections on their experience of supervision in a 2D virtual world revealed the usefulness of the technology readiness index model, from which an extension to the model is proposed to include facilitators for the application of the technology that may mediate the motivators and inhibitors when assessing readiness to use Metaverse in education settings.

Keywords: Education, Metaverse, Virtual Worlds, Self-reflection, Technology Readiness.

1. Introduction
The COVID-19 pandemic severely disrupted the educational environment, significantly impacting students, teachers, and institutions worldwide. Academic institutions were forced to explore new methods of online teaching, such as using virtual and augmented reality technologies to create immersive learning experiences (Johnson, Veletsianos, Seaman, and Hanes, 2021). This aided in catalyzing a transition to online learning, thus opening education to the viability of the online world and accelerating the adoption of virtual learning (Smith, 2020). It has also facilitated the adoption of the Metaverse represented by virtual worlds. According to Tlili et al. (2022), the term "metaverse" refers to the joining of the words "universe," which describes a parallel or virtual environment connected to the real world, prefixed with "meta," which implies transcendence (Tlili et al., 2022). It offers an immersive and interactive environment where users can experience a wide range of activities, including gaming, socializing, entertainment, education, and commerce (Dwivedi et al., 2022). Recent technological advancements have made the Metaverse more accessible than before. Today, people can participate in virtual activities from the comfort of their homes at a level of convenience not previously possible (Cheah & Shimul, 2023).

Although the Metaverse is increasingly advocated as an immersive solution for distance education (Smith, 2022), and students have a positive attitude toward immersive Education (Dass, Dabbagh, and Clark, 2011), adopting these technologies in developing countries may be impractical due to high costs, limited connectivity, and power outages. Even without these concerns, an unknown factor is students' acceptance of this form of education, which may hamper the adoption of Metaverse. Before exploring solutions to the cost, connectivity, and power issues, it is prudent to test the students' appetite for using virtual worlds and educators. To this end, we undertook a limited study with three students and a supervisor to explore student's feelings and acceptance of supervision in virtual worlds in preparation for a broader survey on live education with a live lecturer prior to investigating the use of NPCs (non-player combatants) or AI-based (artificial intelligence) supervisors.

With this objective, we posed the question: What are higher education students' impressions of supervision in immersive virtual environments?

The research study focused on the reflections of three tertiary-level students in South Africa, their behavior, perceptions, and how they could adapt to an immersive environment for supervisory meetings. For our initial investigation, we tested a 2D virtual world that uses low bandwidth and runs in a browser. Individual opinions and perceptions are essential to understanding the feasibility of transitioning to a new technology. By acquiring individual feedback and user experience, the study provides insights into the viability of the readiness of students to use an immersive environment for learning using student-supervisor meetings.

The paper is organized as follows. In the next section, the background of the study is presented, followed by the research design. In section four, the findings are shown and discussed in section five. Section six presents a conceptual framework based on the analysis of the findings. Section seven concludes the paper and presents the limitations and suggestions for future research.

2. Background

The swift progress in technology has made virtual instruction a credible replacement for traditional classroom-based teaching. As a result of the flexibility virtual education provides, those living in non-urban areas, are the first in their family to study in higher education, have a disability, come from diverse cultural backgrounds, are older, have families, or are employed now have much stronger representation than they did in the past (Burke and Larmar 2021; Stone and O'Shea 2019). Immersive online education has become increasingly popular. It offers learners an engaging and transformative learning experience that allows students to interact with their peers and instructors in real-time, making online learning more interactive and personalized. It also enables instructors to customize their teaching approach and create a more interactive learning environment.

2.1 Immersive Online Education

Immersive online education refers to leveraging advanced technology to create a learning environment that simulates real-world experiences. For example, learners can interact with and manipulate objects in virtual environments using Virtual Reality (VR) and Augmented Reality

(AR) technologies, improving their understanding of complex concepts and increasing retention.

Immersive virtual environments often use gamification for educational enrichment. Gamification includes incorporating game-like elements and strategies into real-life situations like education. By introducing elements like badges or leaderboards designed to engage students, a more interesting and inspiring learning atmosphere may be created. However, there is a fundamental misunderstanding among educators that gamification is a pedagogical strategy to improve low motivation and engagement among students (Chee and Wong 2017). Rather, the intent behind gamification is to design environments that cultivate motivation and learning capability in students by utilizing game mechanics (Baydas and Cicek 2019; Seaborn and Fels 2015).

According to Twining (2010), choosing the best strategy to assist learners in reaching their objectives is an educational issue that all educators face. Virtual worlds hold the potential to revolutionize how students learn through an immersive experience that cultivates a presence of accessibility and flexibility that designs engaging, interactive, and individualized learning experiences.

2.2 Virtual Worlds

A virtual world can be described as a computer-based, online multi-user environment used to simulate real or fictional life that users can experience using their avatars, which are graphical representations of themselves (Loke, 2015). Worlds can be defined by three main ideas: (1) A shared space occupied and molded by its inhabitants. (2) How users interact and experience the world through physical and psychological responses. (3) A shared experience and space interacting with objects and other individuals is how individuals construct an understanding of the world (Girvan 2018). Thus, it can be concluded that individuals, depending on their physical attributes, personality, and interaction with other individuals and objects, would constitute how users view and define a world. Girvan (2018) defines “virtual” as a simulated environment and a virtual world as a multi-user simulated environment in which users can interact with one another objects and perform activities. Avatars represent individuals who control their actions, such as movements, socializing, and creations.

2.3 Evolution of Virtual Worlds

Even though virtual environments are less common than physical environments for educational purposes, the first virtual reality experience can be dated back to Morton Heilig’s Sensorama in 1962 (Boas 2013). This was a prototype that made use of 3D visuals, audio, haptic, olfactory stimuli, and wind to improve the immersive experience of the user. Sensorama’s objective was to create an experience for a user riding a motorcycle and make it feel as real as possible. Today, tele-immersive technologies have been introduced into the educational sector. Defined by Kuna, Hašková, and Borza (2023) as “immersion in an on-screen environment”, tele-immersive technology allows users to fully immerse themselves into a virtual world with VR glasses and headsets. Advanced technologies such as virtual reality (VR) and augmented reality (AR) in online learning environments provide immersive and interactive experiences that improve student engagement and comprehension (Johnson, Veletsianos, Seaman, and Hanes, 2021) for a wide range of disciplines. VR simulations provide realistic, hands-on experiences

in fields such as science, medicine, and engineering, allowing students to practice their skills and apply their knowledge in virtual environments (Brown and Costello, 2020). AR applications can overlay virtual elements onto the physical world to create interactive and dynamic learning experiences (Smith, 2020). Integrating these technologies helps to bridge the gap between physical and digital learning spaces and allows students to interact with content and peers in more tangible ways.

2.4 Virtual World Technical Features

There are two types of virtual environments. The first is a virtual world where a user can only interact visually. The second one is being fully immersed in the virtual world, which requires additional hardware.

In order to meet the requirements for a visual-only virtual world, a user would only need a desktop computer monitor (preferably with stereo capabilities) (McMenemy and Ferguson, 2007) with lower physical and psychological stress than a fully immersed virtual environment (Dalgarno 2002) for which a computer, head-mounted displays, headphones, and motion-sensing gloves are needed. In terms of student accessibility, making use of this technology in a developing country could be more challenging.

2.5 Virtual Worlds in Education

Gregory and Bannister-Tyrrell (2017) recommend 3D worlds as potential online teaching tools. Teachers create assignments and tasks beforehand in the environment. Students, who are represented by their avatars, are able to interact with one another in real-time by using audio or text. Additionally, students feel at ease communicating with their peers in real-time, whether to confirm instructions given by teachers or acquire assistance, which can decrease the feeling of loneliness or anxiety of remote students. In 2015, (Baleni, 2015) conducted a qualitative study to determine how students feel regarding online assessments. The results showed that most students preferred this assessment method compared to the traditional classroom method. The benefits of online assessment include (1) Immediate feedback so that students can determine their areas of improvement. (2) Formal assessments enhance student understanding, especially if students are allowed to submit the assessment multiple times and an average score is used as a final mark. (3) It adds value to lecturers as they can determine what their students struggle to grasp and where they should provide more clarity (Baleni 2015).

2.6 Supervision

The thesis writing phase for postgraduate students is critical to completing their degrees. Almeatani et al. (2019) reported that 50% of graduate students do not obtain their postgraduate degrees, and 25% drop out before completing their thesis. Completing a thesis requires students to conduct extensive research and formulate arguments and conclusions. Therefore, student and supervisor meetings are essential to the thesis writing process and are beneficial to students as it allows them to obtain guidance and feedback on their research and writing skills. Some of the challenges students face when collaborating with their supervisors include a lack of communication and supervisors not having the time or capacity to meet with their students (Almeatani et al., 2019). Although Almeatani et al. (2019) focus more on using a mobile application to solve communication issues between students and their supervisors, we concentrated on a structured supervision platform.

On the other hand, a study conducted by Rasool et al. (2022) was done to determine the perceptions of students and supervisors on online supervision during COVID-19. Concerns surrounding behavioral concerns and technical issues were recorded. Behavioral concerns include (1) no face-to-face interaction and social bonding, (2) lack of facial expressions and understanding of body language, and (3) students feeling tired and restless in front of their desktop computers when sitting for long periods. Technical concerns include (1) an unstable internet connection and (2) a lack of technical knowledge to handle devices. Technical concerns would be a hindrance in a developing country, especially if students opt to collaborate with their supervisors solely online.

3. Research Design

This qualitative study explored students' self-reflection of the experiences of educational supervision in a 2D and a 3D virtual world to answer the research question, What are higher education students' impressions of supervision in immersive virtual environments?

The unit of analysis for the study was the three individual students whose self-reflections were guided by Gibbs' Reflective Model (Gibbs, 1998) and analyzed using Thematic Analysis (Braun and Clarke 2019) by their supervisor. The students (authors 2, 3, and 4) met with their supervisor (author 1) in the 2D virtual world, Gather Town. They also met face-to-face to introduce the 3D virtual world of Second Life.

The three students reflected on their experiences in the virtual worlds, and the supervisor analyzed the three students' reflections using thematic analysis based on Braun and Clarke (2006). During initial readings by the supervisor, motivators and inhibitors were noticed, which resonated with the Technology Readiness Index (TRI) (Parasuraman and Colby 2015). TRI provided a theoretical framework to guide the analysis of the three students' reflections.

3.1 Reflective Model

For this research study, Gibbs' Reflective Model (Gibbs, 1998) was used to reflect on a set of questions concerning our experiences of supervision in virtual worlds. The questions include: What happened? What were you thinking and feeling? What was good and bad about it? If it arose again, what would you do? What else could you have done? What sense can you make of the situation?

3.2 Thematic Analysis

According to Braun and Clarke (2019), thematic analysis can be deductive, inductive, or a combination of the two. While this study started with a deductive approach guided by the Technology Readiness Index (TRI) (Parasuraman and Colby 2015), the generation of additional themes added an inductive component, which resulted in our positing an extended TRI conceptual framework for investigating virtual worlds and the Metaverse.

The following approach recommended by Braun and Clarke (2006) for analyzing qualitative data guided the data analysis of this study. Thematic analysis is an iterative process that has six basic steps.

- Familiarization with the data

- Generating initial codes
- Constructing potential categories
- Revising the categories
- Defining and naming themes
- Producing a report.

3.3 Theoretical Framework

Acknowledging that non-conscious automatic cognition (System 1) is more robust than rational cognition (System 2), which is at best a partial mediator of adoption (Khatri, Samuel, and Dennis 2018), we were guided in our analysis of the self-reflections by the technology readiness index (TRI) of (Parasuraman and Colby 2015). We used Gibbs' Reflective Model to identify underlying non-conscious cognition through self-reflection (Gibbs, 1998). As we are not always aware of our biases and non-conscious thoughts, we compared multiple self-reflections to provide reliability and validity of the underlying non-conscious cognitions, which are influenced by patterns of experience and personality traits (Khatri et al. 2018).

TRI is a framework that investigates individuals' intentions to adopt and use technology based on their state of mind rather than their innate skills. TRI uses four constructs to understand the individual's state of mind: optimism, innovativeness, discomfort, and insecurity. Optimism and innovativeness are grouped as motivators, and discomfort and insecurity are inhibitors (Parasuraman 2000).

- Optimism: Optimism is the view that technology allows users to attain their goals.
- Innovativeness: Innovativeness is the user's desire to be a leader in the use of technology and willingness to expand knowledge of new technology.
- Discomfort: Discomfort is the perceived lack of control over technology and feeling overwhelmed by it.
- Insecurity: Insecurity results from distrust of technology and uncertainty about its abilities.

Over the past two decades since Parasuraman (2000), TRI has been consistently used in research with respect to technology readiness in education, such as Summak, Bağlıbel, and Samancıoğlu (2010) in Turkey and Mwapwele et al. (2019) in South Africa.

3.4 Ethical Consideration

As the primary data for the study was the self-reflections of the authors, there were no ethical concerns for the study.

4. Findings

The last three authors reflected on their interactions in the virtual worlds (Gather Town, www.gather.town, and Second Life, secondlife.com) guided by Gibbs' reflexive model. The first author analyzed the three sets of reflections and summarised them. Thematic analysis was used following Braun and Clarke (2019), taking a deductive approach based on TRI 2.0 (Parasuraman and Colby 2015).

After reading the transcripts, the four constructs of TRI were used as coding categories and supplemented by in vivo coding. This provided an initial coding list of 34 codes. The codes

were reflected upon and reduced to 20 codes. These were categorized into 11 categories, from which five themes were generated.

Comparable codes were merged, and the merged codes were categorized. In addition to the TRI categories (optimism, innovativeness, discomfort, and insecurity), five other categories were identified. The first category concerned the dependent variable for the project, namely education. The final four were context, infrastructure, skills, and technology.

Through several iterations of code adjustment and categorization, four themes were determined. The TRI model guided the first two themes. These two themes were motivators (optimism and innovativeness) and inhibitors (discomfort and insecurity). Education was the focus of the study and linked to a theme entitled dependent variables. The last four categories (context, infrastructure, skills, and technology) were named facilitators as they deviated from the TRI "in mind" precept, being external to the students' minds and typically having an external locus of control. On the other hand, they are posited to facilitate the motivators and inhibitors. The final themes and categories are shown with examples in Table 1.

5. Discussion

Even with a small student cohort (n=3), the findings show the validity of using the TRI model as a guideline for technology readiness. Although the students felt some discomfort (n=5) and insecurity (n=5), their motivation was higher, revealing feelings of optimism (n=14) and innovativeness (n=12). Although an external factor of COVID-19 was a driver for the venture into the Metaverse, this was not reflected by the students. On the other hand, several drivers, considered facilitators for this study, were in evidence. The most commented category was technology (n=11), followed by skills (n=7), infrastructure (n=4), and context (n=1).

A review of concurrences revealed a pattern that the authors felt was worthwhile investigating. The starting point was to separate education as the dependent variable from the facilitator theme. Education showed co-occurrences to motivation (n=9), of which the co-occurrence to optimism was the greatest (n=7), and a co-occurrence to inhibitors (n=4) with co-occurrence with insecurity of n=3. Thus, the students were optimistic about using a 2D environment for education but also felt insecure.

Student R commented, "I would recommend [the 2D virtual world] to any fellow students looking for an immersive yet practical learning experience", given that initially, student R "was a bit confused as to how we would test the effectiveness of a platform."

The facilitator theme (n=14) showed co-occurrences to both motivators (n=10) and inhibitors (n=4). The most mentioned were skills with co-occurrence with motivation (n=6) but not with inhibitors. The most vital facilitator link was to optimism (n=4), for which the most mentioned facilitator was skill transfer through scaffolding. "our supervisor … has a passionate outlook on everything virtual. The notion I have gained from him is that he believes in finding a collaboration in virtual worlds and education to help us students" (Student C).

Theme	Category	Code	Quote
Motivators	Optimism		"As I was not as interested in the meetings initially, I found that with the interactiveness and immersive element now being bought by this new platform, I found a new excitement and enjoyment towards research being ignited, unlike traditional web-conferencing platforms." Student R
	Innovativeness		"I realized how interactive it actually was. The more I used it, the more I liked it. As a result, this increased my motivation, and I found myself feeling excited to start exploring this virtual world and taking advantage of its additional features." Student I
Inhibitors	Discomfort		"[The virtual world] seems like it is adding unnecessary features, which will not enhance the way we relate information to one another. I can acknowledge and appreciate how amazing the environment is; however, I'm sceptical regarding the value and benefits of using [virtual worlds]." Student I
	Insecurity		"Initially, I was a bit confused as to how we would test the effectiveness of a platform such as Gather Town; however, as time went on, I realized it was relatively straightforward, as any learning that occurred on the app would directly show how effective the platform was at hosting student-supervision virtual meetings." Student R
Facilitators	Context		"The reason as to why a more immersive environment is not a suitable choice is because of the current challenges South Africa is facing as a developing country." Student I
	Infrastructure		"I would have liked to test how well the platform runs on a slower, more outdated PC, a weaker wifi connection and a Cellphone device. As we are in South Africa, a lot of students do not have the resources used when I accessed Gather Town, so in order to have a fair opinion, I would have liked to run Gather Town on other devices." Student R
	Skills	Existing Skills	"Due to our very one-dimensional meetings, it did not give us a chance to effectively trial all the other features that could have potentially blown my mind in a positive way." Student R
		Platform Induction	"I would have liked to have a scheduled meeting once a month with everyone in - person to discuss the intricacies of what we have been experiencing together in Gather Town." Student C
		Scaffolding	"Furthermore, our supervisor … has a passionate outlook on everything virtual. The notion I have gained from him is that he believes in finding a collaboration in virtual worlds and education to help us students obtain a better gauge of the educational standpoint and how virtual worlds can be implemented effectively in that setting." Student C
	Technology	Alternatives	"Gather Town acts as an online web-conferencing room similar to Teams, Google Meets and Zoom. However, the similarities and differences are vastly different. For instance, Gather Town allows users to create and customize an avatar to then prompt into a virtual space that resembles a real-life classroom, board room, conference room or any other setting for users to utilize and interact with. These features are all in the hope to create an interactive environment prompting the realism factor of a traditional setting." Student C
		Exploring	"On the other hand, it would also be beneficial to explore other 2D environments in order to compare and evaluate which platform would work best for our objectives and goals." Student I
		Issues	"However, a few ironing out of certain features needs to be looked at. For example, a regular occurrence was the inability to hear one another over our respective microphones. The way Gather Town works is that when you step into a room that assesses the need for interaction, you will be prompted with a microphone option to unmute yourself and start having conversation." Student C
		Setup	"Another instance I picked up during our sessions is that we are "guests" and not "editors". This means we are limited to certain features … we will need … to upgrade our status to editor. In contrast, this may be damaging in a classroom environment or setting where the educator would be the editor and students the guests as the editor has higher authority in setting up certain restrictions that the students do not have the privilege to." Student C
Dependent Variable	Education	Basis for Learning	"My interest in the idea of using virtual worlds in education sparked because of its potential to increase learner engagement and generate immersive learning experiences. …. As I embarked on the journey, I was filled with excitement and curiosity." Student C
		Face-to-Face	"I found that with the interactiveness and immersive element now being bought by this new platform, I found a new excitement and enjoyment towards research being ignited unlike traditional web-conferencing platforms." Student R
		Hybrid	"The 2D virtual environment of Gather Town proved its validity. However, there is a limit to what you can and cannot … I would have liked to have a scheduled meeting once a month with everyone in - person to discuss the intricacies of what we have been experiencing." Student C
		Immersive	"As I was not as interested in the meetings initially, I found that with the interactiveness and immersive element now being bought by this new platform, I found a new excitement and enjoyment towards research being ignited unlike traditional web-conferencing platforms." Student R
		Virtual	"… any learning that occurred on the app would directly show how effective the platform was at hosting student-supervision virtual meetings." Student R
Research	Future Research		"… determine the validity of the online meetings through a different platform other than your typical Zoom or Google Meet." Student C
	Limitations		"However, there is a limit to what you can and cannot do as a student and supervisory role … we prefer the 2D virtual environment specifically because of the convenience it provides. The 3D virtual world, as enriching as it presented itself, posed as too intricate to get started and gave off an overwhelming feeling that would not go well with a lot of students if such an idea had to be ventured into." Student C

Table 1: Thematic analysis findings with supporting quotes.

Sometimes, students who wanted to explore the virtual world felt restricted. “Due to our very one-dimensional meetings, it did not give us a chance to effectively trial all the other features that could have potentially blown my mind in a positive way. However, all the additional out of the way, I still very much liked the use of this platform” (Student R). This highlights a potential challenge for students with existing skills, such as being au fait with online gaming. The net result is that the most observed construct for facilitating readiness for a 2D virtual world education environment was the development of a virtual world skillset (n=6).

Innovativeness (n=2) was the next most common observation. “The [virtual] space … came with work-space customization, a whiteboard for presentation, a private meeting room that allowed those who entered the room to access the audio of the room and the people uti-lizing it, as well as a virtual kitten, to accompany us during our meetings. Furthermore, the few features mentioned added to the realism factor and how effective it would be in stimulating student’s minds compared to that of a Zoom classroom or Google Meet” (Student C). Although skills were not shown to be an inhibitor, student C recognized that “there is a limit to what you can and cannot do as a student and supervisory role.” Nevertheless, a 3D virtual world was commented on in terms of “the consensus that we prefer the 2D virtual environment specifically because of its convenience. The 3D virtual world, as enriching as it presented itself, posed [to be] too intricate to get started and gave off an overwhelming feeling that would not go well with many students” (Student C).

Although technology (n=6) had a similar count to skills, there was a broader spread of co-occurrences. Optimism, innovativeness, and discomfort all reported n=2. Consequently, technology leaned towards being a motivator (n=4) but was tempered by the inhibitor count (n=2) associated with discomfort. Innovativeness was seen in how “amazing the [2D virtual world] environment is” (Student I). In contrast, optimism was seen in “recreating a realistic environment for students to feel a resemblance of a traditional brick and mortar classroom” (Student C) and the desire to explore “the other features the platform has on offer” (Student R). Discomfort was observed in relation to some features that did not operate smoothly at all times, such as an issue with the sound that required the user to “leave the meeting only to return in the hope that it would work after the attempt of leaving. In most cases, it does” (Student C). However, this is a challenge also observed in other online meeting software.

Infrastructure was a motivator (innovativeness, n=1) and inhibitor (insecurity, n=1). The observation was that although South Africa is facing infrastructure “challenges” (Student I) (inhibitor), student I reported that it was possible to connect to the 2D virtual world on the student’s “Macbook Air (2020) as well as my HP (2020) computer. I even tried making use of my mobile phone.” (innovativeness).

Finally, the context was an inhibitor due to insecurity (n=1), for example, the lack of electrical power and Internet resources as a reason why a “more immersive environment is not a suitable choice” (Student I)

6. Conceptual Model

The findings from this provisional study prompted us to create the TRI-F (technology readiness index with facilitators) conceptual model shown in Table 2. The model is populated with the findings counts to illustrate its use. The model encourages mapping external facilitators to the motivator and inhibitor mindsets of the potential adopter of technology to determine their readiness and the implications of external factors on their readiness.

The conceptual model expands the TRI model to add a dependent variable (in this case, education), which is impacted by all four categories and themes of TRI. Additionally, the model incorporates a set of Facilitators comprising four categories (Context, Infrastructure, Skills, and Technology), which showed varying degrees of implication for each of the TRI categories. This model is important in providing a more granular level of investigation for each category of the TRI Framework. Conversely, the TRI-F model allows a more granular approach to investigating facilitators of innovation such as context, infrastructure, skills, and technology through the tried and tested TRI Framework.

Theme	**Category**	**Motivators (n=20)**		**Inhibitors (n=8)**	
		Optimism (n=13)	**Innovativeness (n=7)**	**Discomfort (n=3)**	**Insecurity (n=5)**
Facilitators (n=15)	Context (n=1)	-	-	-	1
	Infrastructure (n=2)	-	1	-	1
	Skills (n=6)	4	2	-	-
	Technology (n=6)	2	2	2	-
Dependent Variable (n=13)	Education (n=13)	7	2	1	3

Table 2: The TRI-F (Technology Readiness Index with Facilitators) Conceptual Model.

7. Conclusion

While online meeting technologies have become commonplace, face-to-face meetings lack connectedness. The Metaverse is promoted for overcoming this problem and providing an immersive experience that fosters connectedness. The study confirmed the view of a predominant student-dominant feeling of optimism for education in the Metaverse (Dass, Dabbagh, and Clark, 2011), albeit with some insecurity. This finding aligns with the technology readiness index (TRI) of Parasuraman and Colby (2015). However, a third theme of facilitation was observed, which was posited as potentially mediating the motivator and inhibitor themes of TRI. This theme was incorporated into a conceptual model that we refer to as TRI-F (technology readiness index with facilitators). The conceptual model expands the TRI model to add a dependent variable, a set of facilitators which could provide a more granular level of investigation for each category of the TRI Framework.

As the model was developed inductively from a small sample and saturation was not necessarily reached, further research is needed (a) to validate the model with a larger sample and (b) to test it quantitatively.

Student C effectively summed up the paper and the need for further research. “In essence, I believed that the experience … gave me greater motivation to venture further into different platforms to determine whether the realism factor and interactivity will be beneficial or if the traditional brick and mortar classrooms are [preferable].”

References

Almeatani, M., Alotaibi, H., Alasmari, E., Meccawy, M., and Alghamdi, B. (2019). Thesis Supervision Mobile System for Enhancing Student-Supervisor Communication. *International Journal of Interactive Mobile Technologies (iJIM)* 13(06):4. doi: 10.3991/ijim.v13i06.9533.

Baleni, Z. G.. (2015). Online Formative Assessment in Higher Education: Its Pros and Cons. 13(4).

Baydas, O., and Cicek, M. (2019). The Examination of the Gamification Process in Undergraduate Education: A Scale Development Study. *Technology, Pedagogy and Education* 28(3):269–85. doi: 10.1080/1475939X.2019.1580609.

Boas, Y.A. G. V. (2013). Overview of Virtual Reality Technologies. 1–6. Bogota, Colombia.

Braun, Virginia, and Victoria Clarke. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* 3(2):77–101. doi: 10.1191/1478088706qp063oa.

Braun, V., and Clarke, V. (2019). Reflecting on Reflexive Thematic Analysis. *Qualitative Research in Sport, Exercise and Health* 11(4):589–97. doi: 10.1080/2159676X.2019.1628806.

Brown, M., and Costello, E. (2020). The Future of Learning and Teaching in the Era of COVID-19, *Postdigital Science and Education*, vol. 2, no. 3, pp. 923-932.

Burke, K., and Larmar, S.. (2021). Acknowledging Another Face in the Virtual Crowd: Reimagining the Online Experience in Higher Education through an Online Pedagogy of Care. *Journal of Further and Higher Education* 45(5):601–15. doi: 10.1080/0309877X.2020.1804536.

Cheah, I., and Shimul, A. S. (2023). Marketing in the Metaverse: Moving Forward – What’s Next? *Journal of Global Scholars of Marketing Science* 33(1):1–10. doi: 10.1080/21639159.2022.2163908.

Chee, C-M., and Wong, D, H-T. (2017). Affluent Gaming Experience Could Fail Gamification in Education: A Review. *IETE Technical Review* 34(6):593–97. doi: 10.1080/02564602.2017.1315965.

Dalgarno, B. (2002). The Potential of 3D Virtual Learning Environments: A Constructivist Analysis. *Electronic Journal of Instructional Science and Technology* 5(2):3–6.

Dass, S., Dabbagh, N., and Clark, K., (2011).Using virtual worlds: What the research says, *Quarterly Review of Distance Education*, 12(2).

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., Dennehy, D., Metri, et al. (2022). Metaverse Beyond the Hype: Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice and Policy. *International Journal of Information Management* 66:102542. doi: 10.1016/j.ijinfomgt.2022.102542.

Gibbs, G. (1998). *Learning by doing: A guide to teaching and learning methods.* Oxford Brooks University, Oxford.

Girvan, C. (2018). What Is a Virtual World? Definition and Classification. *Educational Technology Research and Development* 66(5):1087–1100. doi:10.1007/s11423-018-9577-y.

Gregory, S., and Bannister-Tyrrell, M. (2017). Digital Learner Presence and Online Teaching Tools: Higher Cognitive Requirements of Online Learners for Effective Learning. *Research and Practice in Technology Enhanced Learning* 12(1):18. doi: 10.1186/s41039-017-0059-3.

Johnson, N., Veletsianos, G., Seaman, J., and Hanes, D. (2021) Digital learning in the time of COVID-19: Understanding the transition to online education. *Internet and Higher Education*, 48, 100778.

Khatri, V., Samuel, B. S., and Dennis, A. R. (2018). System 1 and System 2 Cognition in the Decision to Adopt and Use a New Technology. *Information & Management* 55(6):709–24. doi: 10.1016/j.im.2018.03.002.

Kuna, P., Hašková, A., and Borza, L. (2023). Creation of Virtual Reality for Education Purposes. *Sustainability* 15(9):7153. doi: 10.3390/su15097153.

Loke, S-K. (2015). How Do Virtual World Experiences Bring about Learning? A Critical Review of Theories. *Australasian Journal of Educational Technology* 31(1). doi: 10.14742/ajet.2532.

Manegre, M., and Sabiri, K. A. (2022). Online Language Learning Using Virtual Classrooms: An Analysis of Teacher Perceptions. *Computer Assisted Language Learning* 35(5–6):973–88. doi: 10.1080/09588221.2020.1770290.

McMenemy, K., and Ferguson, R. S. 2007. *A Hitchhiker's Guide to Virtual Reality*. CRC Press.

Mwapwele, S. D., Marais, M., Dlamini, S., and Van Biljon, J. (2019). Teachers' ICT Adoption in South African Rural Schools: A Study of Technology Readiness and Implications for the South Africa Connect Broadband Policy. *The African Journal of Information and Communication (AJIC)* (24). doi: 10.23962/10539/28658.

Parasuraman, A. (2000). Technology Readiness Index (Tri): A Multiple-Item Scale to Measure Readiness to Embrace New Technologies. *Journal of Service Research* 2(4):307–20. doi: 10.1177/109467050024001.

Parasuraman, A., and Colby, C. L. (2015). An Updated and Streamlined Technology Readiness Index: TRI 2.0. *Journal of Service Research* 18(1):59–74. doi: 10.1177/1094670514539730.

Rasool, U., Zammad, M., Qian, J., and Hussein, S. (2022). The Effects of Online Supervisory Feedback on Student-Supervisor Communications during the COVID-19. *European Journal of Educational Research* 11(3):1569–79. doi: 10.12973/eu-jer.11.3.1569.

Seaborn, K., and Fels, D. I. (2015). Gamification in Theory and Action: A Survey. *International Journal of Human-Computer Studies* 74:14–31. doi 10.1016/j.ijhcs.2014.09.006.

Smith, N. (2020). The COVID-19 crisis: Implications for higher education, *Prospects*, 49(1-2), 1-6.

Smith, J. (2022). Exploring the Potential of the Metaverse for Immersive Distance Education, *Journal of Educational Technology*, 45(3), 432-448.

Stone, C. M. M., and O'Shea, S. E. (2019). My Children… Think It's Cool That Mum Is a Uni Student: Women with Caring Responsibilities Studying Online. *Australasian Journal of Educational Technology* 35(6):97–110. doi: 10.14742/ajet.5504.

Summak, Semih, M., Bağlıbel, M., and Samancıoğlu, M. (2010). Technology Readiness of Primary School Teachers: A Case Study in Turkey. *Procedia - Social and Behavioral Sciences* 2(2):2671–75. doi: 10.1016/j.sbspro.2010.03.393.

Tlili, A., Huang, R., Shehata, B., Liu, D., Zhao, J., Metwally, A. H. S., Wang, H., Denden, M. et al. (2022). Is Metaverse in Education a Blessing or a Curse: A Combined Content and Bibliometric Analysis. *Smart Learning Environments* 9(1):24. doi: 10.1186/s40561-022-00205-x.

Twining, P. (2010). Virtual Worlds and Education. *Educational Research* 52(2):117–22. doi: 10.1080/00131881.2010.482730.