

https://dspace.org/

Critical Success Factors for AI Cybersecurity Implementation in A South African State-Owned Entity: A Socio-Technical Approach

Mamane, Awonke

2025
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/47701

Mamane, Awonke. (2025). Critical Success Factors for AI Cybersecurity Implementation in A South African State-Owned Entity: A Socio-Technical Approach [Master`s dissertation, University of the Witwatersrand, Johannesburg]. WIReDSpace. https://hdl.handle.net/10539/47701

Downloaded from a DSpace repository. https://dspace.org/

CRITICAL SUCCESS FACTORS FOR AI CYBERSECURITY IMPLEMENTATION IN A SOUTH AFRICAN STATE-OWNED ENTITY: A SOCIO-TECHNICAL APPROACH

By

AWONKE MAMANE

(2138857)

SUPERVISOR: PROFESSOR RENNIE NAIDOO (PhD)

Research report

Submitted in partial fulfilment of the requirements for the degree of
Master of Commerce (Information Systems)
in the
School of Business Sciences,
Faculty of Commerce, Law and Management,
University of the Witwatersrand

Date of submission

27 February 2025

DECLARATION

Declaration regarding Plagiarism

I declare that this is my own unaided work and is, to the best of my knowledge and belief, original.

I have read and understood the Senate policy on plagiarism and am aware that plagiarism is the intentional or unintentional "failure to acknowledge the ideas or writing of another" or "presentation of the ideas or writing of another as one's own". In this context "other" means any other person including a student, academic, professional, published author or other resource such as the internet. Failing to acknowledge the use of ideas of others constitutes an important breach of the values and conventions of the academic enterprise.

I am aware that plagiarism offences will be dealt with in terms of Senate policy and may be subject to disciplinary action.

I (full names & surname):	Awonke Mamane
Student number:	2138857

Declare the following:

1. I understand what plagiarism entails and am aware of the University's policy in this regard.
2. I declare that this assignment is my own, original work. Where someone else's work was used (whether from a printed source, the Internet or any other source) due acknowledgement was given and reference was made according to departmental requirements.
3. I did not copy and paste any information directly from an electronic source (e.g., a web page, electronic journal article or CD ROM) into this document.
4. I did not make use of another student's previous work and submitted it as my own.
5. I did not allow and will not allow anyone to copy my work with the intention of presenting it as his/her own work.



Signature

27 February 2025

Date

ACKNOWELDGEMENT

I wish to thank my supervisor, Prof Rennie Naidoo for his guidance and for empowering me to do and complete my research report. I am thankful to my wife Axole for her unwavering support and encouragement. To my beloved daughters Aphakama and Sive, thank you for your love and patience. This work is dedicated as a tribute to my late son Kusekuhle, whose memory continues to inspire me.

LIST OF ACRONYMS AND ABBREVIATIONS

AI: Artificial Intelligence

CFS: Critical Success Factor

CIO: Chief Information Officer

CSIRT: Computer Security Incident Response Team

CVE: Common Vulnerability and Exposure

DLP: Data Loss Prevention

FPR: False Positive Rate

HCI: Human-computer interaction

HEI: Higher Education Institution

ICT: Information and communication technology

IS: Information Systems

LLM: Large language model

MIS: Management Information Systems

MTTD: Mean Time to Detect

MTTR: Mean Time to Resolve

NCPF: National Cybersecurity Policy Framework

NQF: National Qualifications Framework

SOC: Security Operations Center

SIEM: Security Information and Event Management

SOAR: Security Orchestration, Automation, And Response

STS: Socio-technical Systems

UEBA: User and Entity Behaviour Analytics

UX: User Experience

XAI: eXplainable AI

TABLE OF CONTENTS

CRITICAL SUCCESS FACTORS FOR AI CYBERSECURITY IMPLEMENTATION IN A SOUTH AFRICAN STATE-OWNED ENTITY: A SOCIO-TECHNICAL APPROACH 1

DECLARATION I

ACKNOWELDGEMENT II

LIST OF ACRONYMS AND ABBREVIATIONS III

TABLE OF CONTENTS IV

ABSTRACT 1

1 CHAPTER 1. INTRODUCTION 2

1.1 BRIEF BACKGROUND TO THE TOPIC 2

1.2 PROBLEM STATEMENT 4

1.3 RESEARCH QUESTIONS (AND OBJECTIVES) 5

1.4 DELIMINATIONS OF THE STUDY 6

1.5 IMPORTANCE (SIGNIFICANCE) OF THE STUDY 6

1.6 BRIEF CHAPTER OVERVIEW 7

2 CHAPTER 2. LITERATURE REVIEW 9

2.1 INTRODUCTION 9

2.2 INSIGHTS FROM IS IMPLEMENTATION RESEARCH 9

2.2.1 Understanding Implementation 9

2.2.2 Success and Failure of IS Implementation 10

2.3 AI IN CYBERSECURITY 11

2.4 IMPORTANCE OF THE SOCIO-TECHNICAL LENS IN AI CYBERSECURITY 12

2.5 SOCIAL FACTORS IN AI CYBERSECURITY 13

2.5.1 Change Acceptance 13

2.5.2 Top Management Support 14

2.5.3 Cybersecurity Awareness 14

2.5.4 Human Oversight 15

2.5.5 Usability 15

2.5.6 Ethical Considerations 15

2.6 TECHNICAL FACTORS AND THEIR INTERPLAY WITH SOCIAL DYNAMICS 17

2.6.1 Data Quality 17

2.6.2 Scalability 17

2.6.3 Automation 17
2.6.4 Efficiency 18
2.6.5 Socio-Technical Interplay 18
2.7 CRITICAL SUCCESS FACTORS FOR AI CYBERSECURITY, FROM A SOCIO-TECHNICAL PERSPECTIVE 20
2.8 EMPIRICAL THEORY DISCUSSION 20
2.8.1 Evaluation and Selection of The Theoretical Framework 20
2.9 DISCUSSION OF THE SELECTED THEORY (FRAMEWORK OR MODEL) 21
2.9.1 Socio-Technical Systems Theory 21
2.9.2 Critical Success Factors Theory 22
2.9.3 The AI In Cybersecurity Socio-Technical Critical Success Factors Model 23
2.10 CHAPTER SUMMARY 25
3 CHAPTER 3. RESEARCH METHODOLOGY 26
3.1 INTRODUCTION 26
3.2 RESEARCH PHILOSOPHY (PARADIGM) 27
3.3 RESEARCH APPROACH 30
3.4 RESEARCH METHOD 31
3.5 RESEARCH DESIGN 31
3.5.1 Case Study Background 32
3.6 RESEARCH POPULATION AND SAMPLING 35
3.7 DATA COLLECTION 36
3.8 DATA ANALYSIS 37
3.9 ETHICAL CONSIDERATIONS 38
3.10 CHAPTER SUMMARY 39
4 CHAPTER 4. PRESENTATION OF FINDINGS 40
4.1 INTRODUCTION 40
4.2 ANALYSIS AND PRESENTATION OF THE FINDINGS 40
4.3 SOCIAL FACTORS 41
4.3.1 Change Acceptance 41
4.3.2 Top Management Support 45
4.3.3 Cybersecurity Awareness 47
4.3.4 Human Oversight 49
4.3.5 Ethical Considerations 51
4.3.6 Usability 53

4.4 TECHNICAL FACTORS 54
4.4.1 Data Quality 54
4.4.2 Scalability 56
4.4.3 Automation 57
4.4.4 Efficiency 59
4.5 SOCIO-TECHNICAL INTERPLAY 60
4.5.1 Human Oversight and Automation 60
4.5.2 Top management support and scalability 61
4.5.3 Usability and Efficiency 61
4.5.4 Data Quality & Ethical Considerations 62
4.6 CHAPTER SUMMARY 62
5 CHAPTER 5. DISCUSSION OF THE FINDINGS 63
5.1 INTRODUCTION 63
5.2 SOCIAL FACTORS 63
5.2.1 Change Acceptance 63
5.2.2 Top Management Support 64
5.2.3 Cybersecurity Awareness 66
5.2.4 Human Oversight 67
5.2.5 Ethical Considerations 67
5.2.6 Usability 68
5.3 TECHNICAL FACTORS 69
5.3.1 Data Quality 69
5.3.2 Scalability 70
5.3.3 Automation 71
5.3.4 Efficiency 71
5.4 SOCIO-TECHNICAL INTERPLAY 72
5.4.1 Human Oversight and Automation 72
5.4.2 Top Management Support and Scalability 72
5.4.3 Usability and Efficiency 73
5.4.4 Data Quality & Ethical Considerations 73
5.5 CHAPTER SUMMARY 74
6 CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS 75
6.1 INTRODUCTION 75
6.2 SUMMARY OF FINDINGS 75

6.2.1 Objectives of The Study Reaffirmed 75
6.2.2 Social Factors 77
6.2.3 Technical Factors 78
6.2.4 The Socio-Technical Interplay 80
6.3 SUMMARY OF CONTRIBUTIONS 81
6.3.1 Theoretical Contributions 82
6.3.2 Methodological Contributions 82
6.3.3 Practical Contributions 82
6.3.4 Lessons Learnt 83
6.4 LIMITATIONS OF THE STUDY 84
6.5 CONCLUSION 84
7 REFERENCES 86
8 APPENDIX A: INTERVIEW PROTOCOL 107
9 APPENDIX B: PARTICIPANT INFORMATION SHEET (PIS) 109
10 APPENDIX C: CONSENT FORM 111
11 APPENDIX D: ETHICS CLEARANCE 112
12 APPENDIX E: PERMISSION TO CONDUCT THE CASE STUDY 113
13 APPENDIX E: TURNITIN REPORT 114

LIST OF FIGURES

Figure 1 Socio-Technical System interplay [adapted from Malatjie (2020)] 19
Figure 2 Socio-Technical Critical Success Factors (STCSF) Model 24
Figure 3 Research Methodology for The Socio-Technical Critical Success Factors of Implementing AI in Cybersecurity 26
Figure 4 Edu X ICT Division Organogram 33
Figure 5 AI Cybersecurity Architecture of Edu X 34
Figure 6 Sample codification word cloud 38
Figure 7 Thematic analysis map 41

LIST OF TABLES

Table 1 The five basic attributes of the interpretivist paradigm (adapted from Oates, 2006) 30
Table 2 Organisational roles included in the sample 35
Table 3 Duration of interviews 36

Table 4 Social Factors 78
Table 5 Technical Factors 80
Table 6 Socio-technical interplay 81

ABSTRACT

The integration of artificial intelligence (AI) into cybersecurity is crucial for mitigating increasingly sophisticated cyber threats. However, existing research often focuses exclusively on either technical or social factors, leading to an imbalance in AI cybersecurity implementation priorities. This study addresses this gap by exploring the critical success factors (CSFs) for AI cybersecurity implementation in a South African state-owned entity (SOE) by applying a socio-technical lens. Using an interpretive case study approach, this research employed in-depth interviews with technical experts and end-users in the SOE. Thematic analysis was used to identify the key critical success factors. The findings reveal that successful AI cybersecurity implementation depends on technological advancements and their alignment with human, environmental, and organisational factors. The key technical CSFs identified include *data quality, efficiency, scalability, and automation*, while the social CSFs encompass *change acceptance, top management support, cybersecurity awareness, human oversight, ethical considerations, and usability*. A socio-technical perspective provides practitioners with a more balanced understanding of AI cybersecurity implementation, ensuring that security solutions are not only technically sound but also socially acceptable and effectively integrated within organisational structures. Based on these findings, the study develops a framework to guide policymakers, IT security professionals, and AI developers in implementing AI-driven cybersecurity solutions effectively. This research advances the application of socio-technical theory in organisational cybersecurity by illustrating how the interplay between technological and social factors can influence the successful implementation of AI-driven security solutions. Future research should examine how human oversight can be effectively integrated with AI in cybersecurity to enhance reliability, ethical compliance, and user trust in AI-enabled security solutions.

Keywords: *artificial intelligence, case study, cybersecurity, socio-technical, critical success factors, state-owned entity, implementation.*

1 CHAPTER 1. INTRODUCTION

In recent years, there has been a growing number of cybersecurity attacks in South African government departments and state-owned entities (Pieterse, 2021). According to Microsoft Threat Intelligence (2024), there has been a noticeable rise in the speed, scope, and complexity of attacks, which coincides with the rapid evolution and adoption of AI technology. The effects of these cyber-attacks can be devastating, more especially in governments as they hold crucial national data (Li & Liu, 2021). Liu et al. 2021 assert that a cyberattack can affect the trust and image of a victim organisation, lead to revenue loss, and result in lawsuits. Lee (2021) made the same observation and argued that the benefits of investing in cybersecurity outweigh the cost of repercussions in case of a cyberattack. There are various cybersecurity safeguards that organisations can invest in to improve their security posture, and artificial intelligence (AI) has emerged as a powerful tool in fortifying against cyberattacks (Apruzzese et al., 2023).

According to a CIO report by Logicalis (2024) among 1000 Chief Information Officers globally, it emerged that AI in cybersecurity is one of the top priorities, and the majority (around 85%) of organisations are committing resources to the implementation of AI technologies for cybersecurity. In this study, the word ‘implementation’ is used to refer to the processes of planning, deploying, and post-deployment monitoring and evaluation of AI in cybersecurity. Apruzzese et al. (2023) assert that AI in cybersecurity leverages machine learning, deep learning, and natural language processing to swiftly identify and mitigate risks in real-time. Taddeo et al. (2023) posit that when implementing AI cybersecurity there are socio-technical factors that need to be taken into consideration that can impact its success. The term ‘socio-technical’ in AI cybersecurity implementation refers to the integration and interplay of human, organisational, and technological components to enhance cybersecurity (Pollini et al., 2022). This study sought to investigate the socio-technical critical success factors of artificial intelligence in cybersecurity in a South African state-owned entity through a case study.

1.1 BRIEF BACKGROUND TO THE TOPIC

Cybersecurity is a critical component of national security, playing a vital role in protecting a nation's digital infrastructure, safeguarding sensitive information, and ensuring the integrity and availability of essential services (Naidoo & Jacobs, 2023). The escalating number and nature of threats aimed at South African government departments and state-owned entities (SOEs) is a major concern (Pieterse,

2021; Timcke et al., 2023). Fonseca and van Wyk (2021) posit that the South African government is aware of the repercussions of cybersecurity and has a government-funded research centre such as the Council for Scientific and Industrial Research (CSIR) and ICT agencies such as the State Information Technology Agency (SITA) that assist state-owned entities with research and implementation of cybersecurity. This is evident with the National Cybersecurity Policy Framework (NCPF) that was published as far back as 2012, which mandates the cybersecurity hub as National Computer Security Incident Response Team (CSIRT) and the information regulator to regulate the protection of information in the country (Phahlamohlaka et al., 2024).

Previous studies concerning South African state-owned entities have found that the ability of these entities to successfully detect and defend against threats, and promptly share intelligence, is to an extent not effective and this is evident from the cyberattacks against state-owned entities like Transnet a custodian of rail, ports and pipelines in 2021; Department of Justice and Constitutional Development in 2021 and 2024; Companies And Intellectual Property Commission (CIPC) which is the registrar of companies in 2024; Government Employees Pension Fund with is the largest pension fund in Africa in 2024; International Trade Administration Commission of South Africa (ITAC) which is a trading regulatory body in 2024; Tshwane University of Technology (TUT) in 2023 (Brett Van Niekerk, 2017; Dagada, 2024; Pieterse, 2021; Timcke et al., 2023). Mtsweni & Thaba (2022) and Pieterse (2021) posit that social and technical factors like the lack of information and communication technology (ICT) skills, obsolete ICT Infrastructure, inadequate ICT funding, the inability of cybersecurity systems to detect attacks, lack of threat intelligence, and improper implementation of cybersecurity measures have been major contributing factors in the attacks on these SOEs.

It is estimated that the South African economy loses around 2.2 billion rands per year due to cybercrime (van Niekerk et al., 2023). In a review covering 2010–2020, Pieterse (2021) discovered that approximately 39% of cyber incidents involved data exposure, around 22% were related to compromised websites, and approximately 15% were cases of system intrusion, approximately 15% were cases of system intrusion, around 13% relate to cybercrime, approximately 11% were related to denial of service. According to the Cisco Cybersecurity Readiness Index (2024) in South Africa, 5% of organisations are at the Mature stage of cybersecurity readiness which is needed to be resilient against modern cyber security risks, 28% are at the Progressive stage, 57% are at the Formative stage,

and 10% are at the Beginner stage. Taddeo et al. (2023) posit that AI is imperative for fortifying cybersecurity, and Malatji and Tolah (2024) emphasise the importance of socio-technical factors in its implementation by arguing that focusing on technology alone does not assist and that there is a need to consider the social factors as well.

1.2 PROBLEM STATEMENT

Cyber-attacks are sophisticated, and cybercriminals are using AI techniques to formulate attacks that bypass traditional cybersecurity systems like endpoint protection, the defence then tends to depend on the cybersecurity team who also get overwhelmed by attacks (Alavizadeh et al., 2021). Bécue et al. (2021) argue that AI in cybersecurity helps with the detection and prevention of threats across an organisation, they further claim that AI's ability to continuously learn and adapt to new threats positions it at the forefront of cybersecurity efforts. As organisations increasingly rely on digital infrastructure, understanding and harnessing AI in cybersecurity is paramount in safeguarding sensitive information and maintaining trust in the digital ecosystem (Malik et al., 2024). Malatji and Tolah (2024) argue that the implementation of AI in cybersecurity is a double-edged sword that can have dire consequences for an organisation in the form of privacy and ethical violations and errors that AI can make due to learning wrong information. In addressing these concerns, it is imperative to consider the social and technical factors of AI in cybersecurity.

The implementation of AI cybersecurity is a complex exercise that requires consideration of both the technical artifacts and the social factors within the implementation environment. Sarker et al. (2019) emphasise the importance of balancing social and technical aspects within the IS discipline. They argue that although the discipline inherently integrates both aspects, scholars tend to focus predominantly on either technical or social whilst neglecting the other dimension. This imbalance may lead to technically good implementations that fail to consider social factors or systems well aligned with social factors but fail to meet the technical functionality. Malatji et al. (2020) argue that achieving a balanced focus is essential for the success and relevance of cybersecurity implementation. The balance is essential to ensure compliance with organisational and regulatory requirements while achieving technical expectations.

In recent years there have been extensive studies on artificial intelligence cybersecurity among scholars (Araya, 2022; Bécue et al., 2021; Guembe et al., 2022; Malik et al., 2024; Mohamed, 2023;

Sarker, 2023). However, in the extant literature scholars tend to narrowly focus on either the social or technical aspect of AI cybersecurity which leads to a socio-technical imbalance, and there is a scarcity of studies that investigate the socio-technical aspects and critical success factors of AI Cybersecurity. In their study Yeoh et al. (2022) focused on the general cybersecurity critical success factors (CSFs) from an IT capabilities perspective, in their study they do not cover the socio-technical aspect. Taddeo et al. (2023) explored the intricate relationship between artificial intelligence (AI) and cybersecurity within the context of socio-technical ecosystems however did not cover the critical success factors. Merhi (2023) explores the critical success factors for the implementation of Artificial intelligence in general and does not explore AI in cybersecurity.

The area that required exploration pertains to the socio-technical critical success factors of AI in cybersecurity implementation. To address this gap, this research focused on Edu X, a South African state-owned entity in higher education, which manages sensitive data related to South African qualifications and qualification holders, and is responsible for ensuring that qualifications have international comparability. This study aimed to identify the socio-technical CSFs for implementing AI in cybersecurity, offering a model to guide similar organisations in achieving effective and balanced cybersecurity solutions. Cybercrimes are not unique to the public sector institutions, the private sector also suffers the same fate, which has been exacerbated by digitisation and covid-19 pandemic(van Niekerk et al., 2023). The focus has been on a state-owned entity due to the criticality of the data that is stored by state-owned entities and the cyberattacks that the South African government departments and state-owned entities have suffered recently (Timcke et al., 2023; van Niekerk et al., 2023).

1.3 RESEARCH QUESTIONS (AND OBJECTIVES)

The research question sought to understand the critical success factors in the implementation of artificial intelligence in cybersecurity according to end-users and ICT professionals using a socio-technical lens.

The research question is:

What are the socio-technical critical success factors in the implementation of artificial intelligence in cybersecurity.

The sub-questions are:

i. What are the social and technical challenges encountered while implementing AI cybersecurity in state-owned entities?
ii. What critical factors contribute to the successful implementation of AI-based cybersecurity in South African state-owned entities?
iii. What lessons can be learnt from implementing AI cybersecurity in South African state-owned entities?

Research objectives

The research aims to bridge the existing knowledge gap and provide valuable insights into the socio-technical critical success factors for AI in cybersecurity. The organisation under study has implemented AI cybersecurity and is also continuously improving its implementation. In the context of these advancements, the organisation can derive substantial value from a model delineating success factors for AI in cybersecurity. Such a model can not only contribute to the enhancement of AI capabilities but can also prove invaluable in increasing the success rate of AI-cybersecurity implementation.

1.4 DELIMINATIONS OF THE STUDY

The case study on which this study is based was conducted on one South African state-owned entity which operates in the higher education space. The study focuses on the socio-technical critical success factors of implementing AI in cybersecurity, as well as the interplay between these social and technical factors. In this study a selected number of employees of Edu X were interviewed within a period of two months.

1.5 IMPORTANCE (SIGNIFICANCE) OF THE STUDY

The primary role of State-Owned Entities (SOEs) is to ensure that necessary developmental prerequisites are fulfilled for those segments where market provisions are inadequate (Gumede et al., 2024). Given the integral role of ICT systems in societal and economic landscapes, safeguarding against cybercrime has become a fundamental prerequisite for progress in the modern era (Kosutic &

Pigni, 2022; Teichmann et al., 2023). Consequently, cybersecurity emerges as a crucial developmental concern. While having dedicated cybersecurity infrastructure within government and updated legal frameworks are vital, they alone are insufficient for cultivating the requisite human expertise and technical prowess needed to ensure cyber safety (Timcke et al., 2023). Therefore, cybersecurity transcends from being merely an operational necessity; it evolves into a significant policy imperative (Pieterse, 2021).

This research aimed to provide a concise understanding of the socio-technical critical success factors of AI in cybersecurity implementation, this knowledge is crucial for overcoming critical success challenges i.e. change resistance, resource constraints, integration with existing systems, etc. The outcome of this research contributes to the practice and academic literature gap in studies that are focused on the socio-technical critical success factors of AI in cybersecurity. This study provides a framework for identifying and prioritizing socio-technical CSFs, which can guide Edu X and similar entities in implementing effective AI-driven cybersecurity measures. Even though this study focuses on a state-owned entity in South Africa, the findings can translate to other sectors and/or countries. For instance, challenges in South African SOEs can be similar to other developing nations that have resource-constrained organisations.

1.6 BRIEF CHAPTER OVERVIEW

The research detailed discussions are presented in the following chapters.

Chapter 1 Introduction: This chapter briefly gives a background for the study. The identified problem is discussed comprehensively, and the socio-technical critical success factors of AI in cybersecurity are identified as the area that requires exploration. The research question and the objectives of understanding the socio-technical critical success factors in the implementation of artificial intelligence in cybersecurity according to end-users and ICT professionals' perspectives are highlighted, and thereafter the importance of socio-technical CFSs on AI cybersecurity is discussed.

Chapter 2 Literature Review: The first chapter lays a foundation for chapter two which is the review of extant literature, guided by the research objectives. This chapter discusses the critical success factors of AI cybersecurity from a socio-technical perspective. Thereafter the potential theories for the study are discussed as well as where they have been previously used, and the reasons for not using

them in this study are also discussed. The selected theory, the Socio-Technical Critical Success Factors (STCSF) theory is discussed. This theory combines the elements of the socio-technical systems theory and the critical success factors theory. The final part of the chapter depicts and discusses the conceptual model which is the Socio-Technical Critical Success Factors.

Chapter 3 Research Methodology: This chapter discusses interpretivism as the research paradigm chosen for this study and the reasons behind choosing it amongst others. The chapter details the research approach, the strategy, the research method together with the data collection and data analysis techniques for this study.

Chapter 4 Presentation of findings: This chapter comprehensively analyses what has been found from the data collected. A thematic map of the findings is presented, and the social and technical critical success factors are analysed together with the interplay between the two.

Chapter 5 Discussion of the findings: This chapter discusses the findings from the data collected on the socio-technical CFSs. The findings are referenced back to the extant literature to find similarities and differences with the existing knowledge.

Chapter 6 Conclusions and recommendations: This chapter sums up all that has been discussed in the previous chapters with the aim of coming up with concluding remarks and recommendations. This chapter covers the limitations that have been identified in this study on socio-technical CFSs and provides a direction for future studies.

2 CHAPTER 2. LITERATURE REVIEW

2.1 INTRODUCTION

This chapter reviews the literature on AI cybersecurity, its implementation, and socio-technical critical success factors, The chapter starts with explaining the concept of implementation in the context of information systems (IS), it further explores the extant literature on studies about IS implementations. It introduces the concept of AI in cybersecurity and what the scholar have written about it. It then highlights the importance of the socio-technical perspective of AI cybersecurity, thereby elaborating on the social and technical factors of AI in cybersecurity, their interplay, and critical success factors. The potential empirical theories are discussed, and the conceptual model is presented.

2.2 INSIGHTS FROM IS IMPLEMENTATION RESEARCH

2.2.1 Understanding Implementation

Lapointe & Rivard (2007) describe implementation as a process that integrates new policies, technologies and practices into an organisation, they assert that this process is characterised by multiple, at times conflicting relations amongst the involved stakeholders, which then influence the results of the process. In the context of AI for cybersecurity, implementation involves multiple stages, including planning, deployment, adaptation, and post-deployment continuous monitoring. In each phase, there is a need for a socio-technical approach that combines the social and technical aspects to ensure effectiveness (Chatterjee et al., 2024).

A socio-technical perspective fits into all phases of implementation as follows:

i. **Planning Phase**: This phase ensures that there is alignment between the AI capabilities and organisational goals, it involves assessing readiness, communicating with stakeholders, and addressing potential resistance.
ii. **Deployment Phase**: This phase involved rolling out the AI cybersecurity solution to the organisation, ensuring that configurations are done correctly, integrating with existing infrastructure, and providing cybersecurity awareness and training. This phase includes adapting to the change and monitoring AI cybersecurity performance, ensuring compliance

with regulations, integrating ethical considerations, analyzing feedback, and resolving challenges.

iii. **Sustainability Phase**: Involves post-deployment monitoring and evaluation of AI in cybersecurity, and continuously updating AI models.

Pan et al. (2008) and Rajapakse & Thushara (2023) argue that Information systems implementation is a complex journey that requires careful planning and consideration of various elements, including change acceptance, user awareness, top management support, systems integration, and scalability. They further assert that implementation has almost equal chances of being a failure or success and that the onus is on understanding the interplay between the social and technical dynamics during the implementation. Aanestad & Jensen (2016) add to this argument by stating that the failure can be a result of social and technical factors, from poor planning, to change resistance and technology inefficiencies.

2.2.2 Success and Failure of IS Implementation

The implementation of information systems projects and interventions in an organisation is crucial is for its day to day business operations (Meredith, 1981). In pursuit of describing information systems success, Seddon (1997) & Jeyaraj (2020) assert that it is an interdependent multidimensional phenomenon that varies by stakeholder, who can view it from different delivery, net benefits, and ease of use lenses. Delone & McLean (2003) in pursuit of measuring the impact and success of information system implementation, divided IS success into six dimensions namely system quality, information quality, use, user satisfaction, individual impact, and organizational impact. Petter et al. (2008) found that in as much as failure can be glaring, it is difficult to measure implementation success, they argue that the impact of success is more often indirect. Wu & Wang (2006) argue that to measure success one should look at both the social and the technical factors of implementation and how they thrive.

Kim & Kankanhalli (2009) posit that implementation brings changes that disrupt the status quo and routine leading to user resistance. They argue that negative peer opinions and low self-efficacy contribute to its failure. Barker and Frolick (2003) argue that failure is exacerbated by poor communication and lack of employee involvement. Heeks (2002) asserts that misalignments and gap between the system and the region it is implemented on can lead to failure if the local improvisation

gap is not closed. Kim & Kankanhalli (2009) argues that the perceived value of the new system and organisational support can mitigate IS failures.

2.3 AI IN CYBERSECURITY

Cybersecurity refers to the practice of safeguarding digital systems, data, information, networks and devices from attacks, unauthorised access, manipulation and destruction (Li & Liu, 2021). Ozkan-Okay et al. (2024) & Okdem and Okdem (2024) argue that it involves managing risks, detecting and deterring threats, responding to incidents that pose a risk to the confidentiality, integrity and availability of data and systems. Cybersecurity encompasses a range procedures and technologies to counter attacks and further argue that to enhance the speed of countering attacks it leverages Artificial intelligence capabilities (Lim, 2022; Naik et al., 2022; Sadok et al., 2020).

Artificial intelligence is a discipline of computer science that simulates human intelligence through mathematical algorithms, it makes use of but is not limited to machine learning (ML), natural language processing (NLP), and deep learning (DL) to perform cybersecurity tasks (Kaur et al., 2023). 'Machine learning' is a process of AI continuously learning from past cyberattacks, adapting to current threats, and improving its ability to withstand new threats, 'Deep Learning' is a subset of ML that processes mainly unstructured large data used for making decisions. 'Natural language processing' analyses text-based threats mainly found in communications (Lauriola et al., 2022). Kaur et al. (2023) & Teo et al. (2024) posit that there are two main types of AI used in cyber security, which is predictive and generative AI. They further argue that cybersecurity leverages predictive AI to analyse, detect and forecast threats, and generative AI to produce attack simulations, images, text, code, etc.

Jada and Mayayise (2023) and Martin (2022) assert that, AI can bring about transformative enhancements on complex cybersecurity operations in organisational cyber environments, enable accurate analysis of data faster than humans, and improve human-machine collaboration and decision-making. The substantial potential of AI in cybersecurity arises from the shifting landscape, moving from manual identification, response, and mitigation towards automated mitigation (Kaloudi & Li, 2020).

Al-Suqri & Gillani (2022) argue that traditional cybersecurity solutions have limitations in timeously detecting threat vectors, due to the increasing data volume, velocity, and variety generated by emerging technologies and the sophistication of attacks generated with AI. According to Sarker (2024) implementing AI in cybersecurity addresses this gap and leads to time and effort conservation in the identification, investigation, and mitigation of incidents because AI learns and reapplies past patterns when identifying emerging threats. Sarker (2023) argues that AI in cybersecurity significantly reduces the costs and time associated with threat detection and response. Ramos-Cruz et al. (2024) argue that AI cybersecurity outperforms traditional algorithms of insider threat detection and improves defense strategies and the overall cybersecurity approach. Bécue et al. (2021) argues that while AI cybersecurity has advantages over traditional cybersecurity it is not entirely autonomous and may have a probability of encountering disruptions and might struggle to accurately assess novel situations which may then require human intervention.

In the context of organisations operating in the higher education space and state-owned entities, AI-driven Cybersecurity plays a critical role in in ensuring that the data that is protected from unsolicited activities (AL-Ghamdi et al., 2022; Siphambili et al., 2024). Wada et al. (2025) assert that higher education institutions (HEIs) are faced with a heightened cyber threats. They further argue that AI-driven cybersecurity has the capability, but assert that widespread adoption is limited by the implementation costs, and ethical concerns. Khan & Mkuzangwe (2022) argue that AI-driven cybersecurity is a critical component that South African organisations can leverage to improve their cybersecurity posture.

2.4 IMPORTANCE OF THE SOCIO-TECHNICAL LENS IN AI CYBERSECURITY

Chuang (2024) emphasise that AI cybersecurity should consider both the technical and social aspects. They propose a comprehensive framework that underscores the necessity of a holistic approach that combines AI capabilities, ethical considerations, organisational processes, and human expertise to enhance resilience, address biases and ensure cybersecurity efficiency and social alignment. A socio-technical approach ensures both technical and social elements are factored in the implementation of AI in cybersecurity (Kayworth & Whitten, 2010; Taddeo et al., 2023). Roba Abbas et al. (2023) argue that even though AI can enhance traditional cybersecurity solutions, its efficiency is reliant on the socio-technical context in which it operates. Díaz-Rodríguez et al. (2023), Taddeo et al. (2023) and Proudfoot et al. (2023) converge that the effective implementation of AI in cybersecurity can be

guided by established theoretical frameworks like the socio-technical framework, together with cybersecurity best practices and AI principles. This guidance is essential for an impactful and successful the implementation of AI-driven cybersecurity within an organisation (Sarker et al., 2019; Taddeo et al., 2023)

Yuan and Li (2022) acknowledge the impact and importance of artificial intelligence in cybersecurity, in doing so they look at it through the socio-technical lenses and argue that for its successful implementation, there needs to be consideration of the socio-technical factors at play. Sarker et al. (2019) highlight an uneven emphasis on the social and technical components in IS research, with some studies focusing predominantly on either the social or technical aspects and neglecting the other, they further suggest that there is a need for a more balanced approach that considers both components substantively. Thus, it is important to have balance and ensure that the social and technical dimensions are not neglected as they contribute substantially to the success of AI in cybersecurity (Taddeo et al., 2023).

Ahmad et al. (2019) assert that cybersecurity implementation does not merely involve the installation of threat prevention and detection software, however, it is a complex program that involves multifaceted organisational, technological, environmental, structural, and process factors. In support of this, Taddeo et al. (2023) contend that in understanding critical success factors of AI in cybersecurity in a socio-technical landscape, we have to look into the perspectives of various stakeholders including cybersecurity experts, end-users, and managers. Diesch and Krcmar (2020) posit that cybersecurity should ensure credibility, integrity and availability of data, however, they argue that it is difficult to measure the success of cyber security as it cannot be measured directly. There is no unified definition of success in AI cybersecurity as different stakeholders define success from their perspectives (Jeyaraj, 2020).

2.5 SOCIAL FACTORS IN AI CYBERSECURITY

2.5.1 Change Acceptance

The implementation of change necessitates a shift in organisational culture, as employees and security teams must adapt to new tools, automation, and decision-making processes, and this adjustment is crucial for the success of AI in cybersecurity (Lapointe & Rivard, 2007). However, Mwim & Mtsweni (2022) argue that implementing change is not always welcomed and there can be resistance to

adapting to the introduced change. Merhi & Ahluwalia (2019) argue that resistance to change can be due to a lack of understanding of the benefits of the change. Cram et al. (2024) add to this and argue that this can be attributed sometimes to how users feel at a given point in time, thereby advocating for an idiographic approach. Ayat et al. (2020) posit that proper planning and stakeholder engagement can better prepare the parties involved to manage resistance. Koch & Fortkord (2024) argues that as part of managing change, it is necessary to use peers to encourage each other on change adaption.

2.5.2 Top Management Support

Kosutic & Pigni (2022) posit that the top management has a strategic responsibility to lead cybersecurity in an organisation, they argue that senior leadership commitment is crucial for the success of AI in cybersecurity, asserting that they have the responsibility to ensure that AI cybersecurity initiatives are adequately supported and funded. Li et al. (2021) and Berg (2024) argue that underfunding AI cybersecurity poses a risk to the organisation. Gale et al. (2022) emphasise that senior leadership should champion AI adoption and motivate the cybersecurity culture. Uchendu et al. (2021) assert that without executive support, cybersecurity teams may struggle with change resistance, further arguing that their involvement is crucial. Wessel et al. (2021) counters this view by advocating that the users must be self-reliant even when they are not monitored.

2.5.3 Cybersecurity Awareness

Zwilling et al. (2022) argue that while AI has the capabilities to protect the organisation, another layer of protection through cybersecurity awareness is crucial. Mwim & Mtsweni (2022) acknowledge that in practice there is a bias towards the technical aspects of cybersecurity, and they advocate for a change to a holistic cybersecurity culture that accommodates the human counterpart. Chaudhary (2024) posit that a cyber-aware end-user base ensures that in the case of other cybersecurity measures failing, an organisation's defence is strengthened through the human firewall. Khando et al. (2021) and Taherdoost (2022) emphasise that, end-users should be trained through regular awareness campaigns and attack simulations to empower them to use AI cybersecurity responsibly, understand potential risks, and know how to avoid and report risks

2.5.4 Human Oversight

The extant literature indicates the concerns of scholars on the implementation of AI in cybersecurity (Jada & Mayayise, 2023; Martin, 2022; Zhang et al., 2022). Naik et al. (2022) raise concerns of AI replacing humans, particularly in areas involving automation. In addressing this Kaur et al. (2023) and Díaz-Rodríguez et al. (2023) argue that AI is not competing with or replacing humans however the two complement each other as AI needs human agency and oversight. The introduction of AI enhances human-machine collaboration in cybersecurity by automating tasks making it seamless to detect, prevent, and respond with high precision to security threats and cyberattacks (Aminu et al., 2024; Sarker et al., 2024). Brown and Anwar (2021) make mention of the human-in-the-loop concept which entails the fusion of human insight and machine intelligence, forming a crucial approach to harness the mutually beneficial strengths of both human and machine capabilities. Happa et al. (2021) argue that human judgments can be further enhanced through the supportive role played by AI cybersecurity.

2.5.5 Usability

Herrmann & Pfeiffer (2023) argue that AI cybersecurity should be implemented with the human element in mind for ease of use. In what they term ‘user, usage and usability’ (3U’s) Grobler et al. (2021) emphasise the importance of taking into consideration these three components in cybersecurity, they argue that humans are key players in the cybersecurity ecosystem. Shneiderman (2022) adds to this argument by stating that human-centred AI, can help users build trust in the system and ensure they don't simply bypass it out of frustration. Further arguing that usability influences user behaviour and if not taken care of could result in resistance. Usability plays an important role in the adoption of AI cybersecurity ensuring that users can interact with the solution. Baruwal et al. (2024), Alneyadi & Kassim (2023) , & Guerino & Valentim (2020) argue that AI cybersecurity should have as minimum disruption and false positives as possible, as these have a negative impact on the usability of AI in cybersecurity.

2.5.6 Ethical Considerations

The usage of computer technologies inevitably raises ethical concerns, and these concerns include several issues including security, authenticity and privacy. The use of AI in cybersecurity is not spared from these ethical concerns, and from its use, there is a rise in the number of ethical concerns

including transparency, fairness, and accountability (Ashok et al., 2022; Bauer et al., 2023; Kaur et al., 2023; Malatji & Tolah, 2024). Bauer et al. (2023), Galli et al. (2024) & Sarker et al. (2024) emphasise the criticality of eXplainable Artificial Intelligence (XAI) in aligning AI with human values to reduce ethical risks whilst fostering human-AI trust and teaming. Bruschi and Diomede (2023) assert that Artificial intelligence should conform to allowing room for human intervention, preventing harm and malicious use, promoting fairness and equity, and lastly they posit that its capabilities should be communicated to all affected parties before implementation. Choraś and Woźniak (2022) allude to these sentiments and purport that AI cybersecurity although it is necessary it should be used ethically and with caution, they caution that AI cybersecurity is a double edged sword that has the capability to commit and counter crime and they advocate for its accountable use in line with regulations, and that it must not be misused for adversary attacks which may lead to cyber war fare. Gordon et al. (2021), Moabalobelo et al. (2023) , Bharosa (2022) and Radanliev (2024) argue that AI cybersecurity should be used in a way that is considerate of technology and social implications, and they advocate for the alignment with established societal norms and regulations.

Poehlmann et al. (2021) emphasise the importance of regulations and principles for AI cybersecurity in a socio-technical landscape. Díaz-Rodríguez et al. (2023) allude to the importance of regulating AI and emphasise that artificial intelligence needs to be trustworthy and argue that trustworthy AI systems should adhere to legal compliance, ethical alignment, and be robust. Proudfoot et al. (2023) highlight the multifaceted role of cybersecurity regulations in organisations, offering valuable insights for both theoretical understanding and practical application in managing complex cybersecurity challenges, they further make emphasis on ensuring that regulations are followed in all cybersecurity domains including AI in cybersecurity to minimise repercussions. Edu X as a public entity in South Africa is required by law to operate within the confines of the law when it comes to cybersecurity (Fonseca & van Wyk, 2021). Nyakasoka and Naidoo (2024) are however of a different view from the usage of static regulations and framework that are lacking behind in keeping up with the everchanging threat landscape and suggest dynamic alternatives that will keep up with the changes.

2.6 TECHNICAL FACTORS AND THEIR INTERPLAY WITH SOCIAL DYNAMICS

2.6.1 Data Quality

Data quality and quantity are crucial in training AI cybersecurity, the training dataset is a critical factor that can impact the reliability of AI cybersecurity, the data it ingests is reciprocal to its performance (Merhi, 2023; Sarker, et al., 2024). Erhan et al. (2021) and Jabbarova (2023) emphasise that high-quality data is essential for truthful detections, arguing that it minimises false positives thereby ensuring efficiency. Kaushik et al. (2024) argue that AI has inherent biases that can be discriminatory in nature and that can affect its output. The AI systems are also susceptible to attacks that can affect the confidentiality, integrity and availability, this can also include data poisoning manipulation of data models which can cause algorithmic bias challenges like inaccuracy (Sarker, 2024; Shamsan Saleh, 2024). These concerns, if not attended to, can negatively impact the integrity of AI cybersecurity (Witanto et al., 2022).

2.6.2 Scalability

As technology and modern cybersecurity continuously advance, cyber threats are also advancing and becoming sophisticated (Shamsan Saleh, 2024). These technology advances are a cause of concern to some organisations with resource constraints, as keeping up with the advancements requires changes and monetary investments, this then gives rise to the concerns of the AI cybersecurity interoperability with the existing systems (Kaur et al., 2023). Okdem & Okdem (2024) argue that the implementation of AI in cybersecurity does not automatically warrant procurement of additional infrastructure, arguing that AI cybersecurity must be interoperable with existing systems. Kaur et al. (2023) emphasise the importance of flexibility in AI cybersecurity to accommodate organisational processes and structure changes.

2.6.3 Automation

According to Sarker (2024) and Naseer et al. (2024) automation is a critical component of Ai in cybersecurity, it revolutionises traditional cybersecurity by seamlessly handling the predictive analysis, detection, response and orchestration. Gafni & Levy (2024) argue that AI can automate both technical and managerial tasks of cybersecurity, with full automation of the technical aspect and partial automation of managerial tasks, thereby affording the cybersecurity personnel time to focus on more strategic aspects of their work. Nguyen et al. (2024) argues that AI-driven cybersecurity

tools like automated threat intelligence, security information and event management (SIEM) and security orchestration, automation, and response (SOAR) that employ User and Entity Behaviour Analytics (UEBA) have the capability to predict emerging threats, and handle and resolving incidents on their own. Automation of cybersecurity tasks significantly reduces the costs of hiring cybersecurity personnel and minuses the turnaround time of resolving incidents (Sarker, 2023).

2.6.4 Efficiency

Artificial Intelligence (AI) brings about transformative enhancements to cybersecurity through the use of machine learning, deep learning, reinforcement learning and natural language processing, these components of AI contribute to its knowledge and training on how to handle cybersecurity (Lauriola et al., 2022). Erhan et al. (2021) posit that optimizing algorithmic performance contributes to AI cybersecurity threat-handling capabilities, efficiency, and resilience. They argue that the efficiency of AI cybersecurity is more reliant on its algorithmic performance which can be evaluated based on reliability, accuracy, and resolution speed. According to Ozkan-Okay et al. (2024) the key metrics to measure efficiency are the false positive rate (FPR), mean time to detect (MTTD) and mean time to resolve (MTTR). Merhi (2023) emphasises the importance of insisting on efficiency rather than price when procuring AI cybersecurity solutions, highlighting the importance of functionality and scalability. Overall the accuracy, frequency of learning and relearning, return on investment, and predictive and adaptive capabilities contribute to the efficiency of AI in cybersecurity (He et al., 2024; Ozkan-Okay et al., 2024).

2.6.5 Socio-Technical Interplay

Taddeo et al. (2023) assert that socio-technical consideration is indispensable and critical, for a successful implementation of AI in cybersecurity in socio-technical settings like state-owned entities where there is an interplay between social dynamics and technical factors. Malatji et al. (2020) posit that the identification of critical success factors is imperative for achieving optimal outcomes in cybersecurity enhancement through AI integration. They further argue that the socio-technical perspective of cybersecurity in an organisational setup is concerned with the interplay of the social and technical factors which may include the organisation, its employees, the operation environment, the underlying technology, processes, and methods.

Taddeo et al. (2023) further extend the discussion by arguing that the interplay between the social and technical factors is so intense to an extent that it increasingly becoming difficult to differentiate between the two, because technical factors may also include social factors like usability and user acceptance. In the context of cybersecurity, the socio-technical theory (STS) provides a framework for understanding the interplay between organisational components, technology, and humans (Malatji et al., 2020). Figure 1 illustrates the socio-technical interplay in the implementation of AI in cybersecurity.

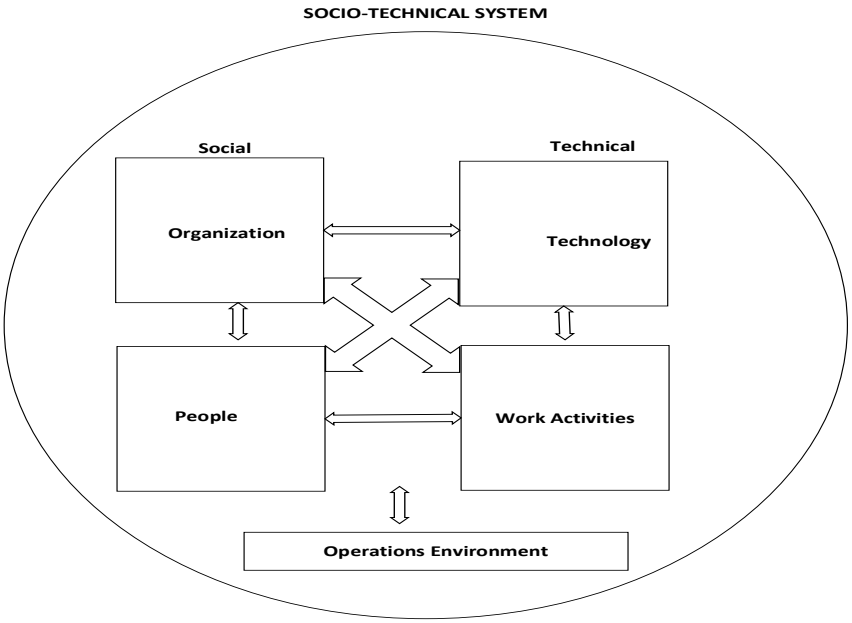


Figure 1 Socio-Technical System interplay [adapted from Malatjie (2020)]

Figure 1 illustrates the ecosystems that AI in cybersecurity operates on. The ecosystem is socio-technical in nature as it combines the social factors like the organisation, work activities, people and the operations environment together with the technology and its capabilities. The effective socio-technical interplay in the implementation of AI in cybersecurity requires collaboration across human expertise, AI capabilities, work flows, and ethical practices. AI cybersecurity should be designed and implemented in such a way that it is accommodative the ecosystem it operates on.

2.7 CRITICAL SUCCESS FACTORS FOR AI CYBERSECURITY, FROM A SOCIO-TECHNICAL PERSPECTIVE

Critical success factors (CSFs) are the key areas or elements that need to be fulfilled or satisfied in order to realise success (Freund, 1988). In AI cybersecurity, when looking at it from the socio-technical lens establishing the success evaluation criteria is not straightforward particularly in relation to the social elements, while the return on investment may be attempted on the technical part it is more difficult to measure its impact on improving the quality of staff work-life improvement from its implementation, the inclusion of the human factor in cybersecurity goes beyond merely putting human resources (Baxter & Sommerville, 2011; Ebert et al., 2023; Van Haastrecht et al., 2021). Yeoh et al. (2022) posit that organisational culture, strategic direction and leadership, cyber awareness and training, technological capability, ethical considerations and governance play a pivotal role in cybersecurity success. Empirical theory discussion

2.8 EMPIRICAL THEORY DISCUSSION

2.8.1 Evaluation and Selection of The Theoretical Framework

When exploring the critical success factors of AI in cybersecurity, there are several theoretical frameworks in place that can provide valuable guidance and structure to the study. These theoretical frameworks include the Protection Motivation Theory (PMT, the DeLone and McLean Model of Information Systems Success, the IT Capability Theory (ITCT), and the technology acceptance model (TAM).

Khan et al. (2023) used PMT to study the cybersecurity protection motivation of end-users, however this theory may not be relevant for studying socio-technical success factors of AI in cybersecurity because it primarily focuses on individual behavioural responses to perceived threats and coping mechanisms, which are more suited to understanding personal security behaviours (Haag et al., 2021). Therefore, it lacks the comprehensive scope needed to evaluate the socio-technical success factors critical to AI in cybersecurity. Zibak et al. (2021) employed the DeLone and McLean Model of Information Systems Success to study the cybersecurity platforms' success, however, this model may not be ideal for this study because it primarily focuses on traditional metrics like system quality and user satisfaction, it is static and does not account for the dynamic nature of cybersecurity, overlooks the complex socio-technical environment (Delone & McLean, 2003; Jeyaraj, 2020). Yeoh et al.

(2022) conducted a study on the critical success factors of cybersecurity using the IT Capability theory. While the IT capability theory can be used to study CFSs of cybersecurity, Wade and Hulland (2004) and Yeoh et al. (2022) posit that the IT Capability theory is designed to understand an organisation's ability to leverage its IT resources for a competitive advantage, and it is not entirely relevant to this study. The TAM theory has been used by Tam et al. (2024) in their study on what influences people's adoption of cybersecurity. TAM is designed to understand and forecast users' acceptance and adoption of novel technologies, considering factors like perceived usefulness and ease of use (Camilleri & Falzon, 2020). TAM's constructs, particularly perceived usefulness, and ease of use may not seamlessly apply to answering the research question. These models are less suitable for a qualitative study that seeks to explore the complex interplay of social and technical aspects, which are critical to understanding the implementation of AI cybersecurity in a state-owned entity.

After a thorough evaluation of these frameworks and considering the specific context of this study, the Critical Success Factors (CSF) theory and Socio-Technical Systems (STS) theory were found to be ideal for this study. The CFS theory is ideal for focusing on the most crucial areas for achieving objectives (Cooper, 2008). The STS theory enables a thorough analysis by incorporating technical, social, organisational, and contextual elements, and it is ideal for qualitative research, enables a deep exploration of the complex interactions and processes in AI cybersecurity implementation (Baxter & Sommerville, 2011; Malatji et al., 2020).

2.9 DISCUSSION OF THE SELECTED THEORY (FRAMEWORK OR MODEL)

2.9.1 Socio-Technical Systems Theory

The term Socio-technical is a combination of two words 'socio' which may be associated with the society in general inclusive of humans, the term 'technical' refers to machines and technology (Walker et al., 2008), this is a term rooted from the 1950s as it came about from the psychology work that was done by the Tavistock Institute (Baxter & Sommerville, 2011; Mumford, 2006). Baxter and Sommerville (2011) argue that Emery and Trist coined the term Socio-technical systems in the 1960s to refer to systems that holistically take into consideration the complex interaction of humans, the environment, machines, and technology for the functionality of a system. In literature there has been notable inconsistency in the use of the term as some scholars call it 'sociotechnical' whilst some call it 'socio-technical', also when referring to theory some scholars call it 'The socio-technical theory' whilst others call it 'The socio-technical systems theory', however the terms are interchangeable and

identical (Baxter & Sommerville, 2011; Mumford, 2006; Walker et al., 2008). This study refers to this theory as ‘The socio-technical systems theory’ and to the term as ‘socio-technical’.

Baxter and Sommerville (2011) argue that to this present-day majority of enterprise systems are characterised by the interaction between individuals, the environment and the machines or technology, they further argue that more often than not systems fail not because of technology failures rather due to neglecting the social-technical ecosystem aspect of where these systems are implemented. In the context of cybersecurity, it can be alluded that AI in cybersecurity requires more than technical consideration (Roba Abbas et al., 2023). Socio-technical theory in the context of AI in cybersecurity is not different from other contexts as it is a perspective that emphasises the interconnectedness of social, technical, and environmental factors, it is a holistic approach that considers both the social and technical aspects of security within an organisation (Roba Abbas et al., 2023). Taddeo et al. (2023) argue that it focuses on the impact of AI in cybersecurity on individuals, organisations, and society as a whole, and the need for an interdisciplinary approach to address the complex challenges and opportunities presented by AI in cybersecurity. The theory also highlights the importance of human requirements and values, inclusive of end-user capacity and awareness, organisational capability and skills development, and stakeholder interests (Cram et al., 2024). This theory provides a valuable perspective for developing a holistic enterprise AI cybersecurity implementation that considers both the social and technical aspects within contextual environments. It is adopted in this research with the aim of answering the research question.

2.9.2 Critical Success Factors Theory

Freund (1988) defines Critical Success Factors as the things that should be done to realise success. Research in Information Systems (IS) has extensively explored the fundamental factors contributing to organisational success, with notable studies such as that by Delone and McLean (2003), their work influenced the development of CSF theory and provided a framework for understanding the factors that are critical for the success of information systems within organisations (Jeyaraj, 2020). The CSF theory is a framework that is used to identify the key factors that are essential for the success of the implementation of a particular strategy or initiative within an organisation (Williams & Ramaprasad, 1996). These factors are the essential activities and processes that must be performed well for an organisation to be successful. The CSF theory emphasises the importance of focusing on the most critical aspects of an organisation's operations and aligning resources and efforts to ensure success in

those areas (Cooper, 2008). The theory helps organisations prioritise their efforts and resources, leading to more effective strategic planning and decision-making
(Li, 1997). By identifying and addressing critical success factors, organisations can improve their performance and achieve their desired outcomes.

The critical success factors for AI in cybersecurity are essential areas that, when successfully addressed, can lead to effective cybersecurity implementation in an organisation. These factors are crucial for organisations to thrive and achieve their business objectives. The purpose of the study was to gain a clear understanding of the most influential and critical factors of AI in cybersecurity implementation success. The problem statement highlights the challenges organisations have with regard to cybersecurity. Many organisations struggle to implement counter technology for their defence leading to a high failure rate of their cybersecurity implementation characterised by successful cyberattacks (Pieterse, 2021). The uncertainty of how to successfully implement AI in cybersecurity, together with the substantial financial investment required poses a significant challenge for organisations.

2.9.3 The AI In Cybersecurity Socio-Technical Critical Success Factors Model

By combining the STS theory and CSF theory, this study developed the Socio-Technical Critical Success Factors (STCSF) Model. The STCSF model highlights the importance of considering both social and technical components to ensure the successful implementation and operation of technological systems in organisations. This study uses this model to address the complex critical success factors and the interplay between social, technical, and environmental factors in the context of AI in cybersecurity. Figure 2 depicts the conceptual framework for this study, this model has been adapted from Malatji et al. (2020) and Kayworth & Whitten (2010).

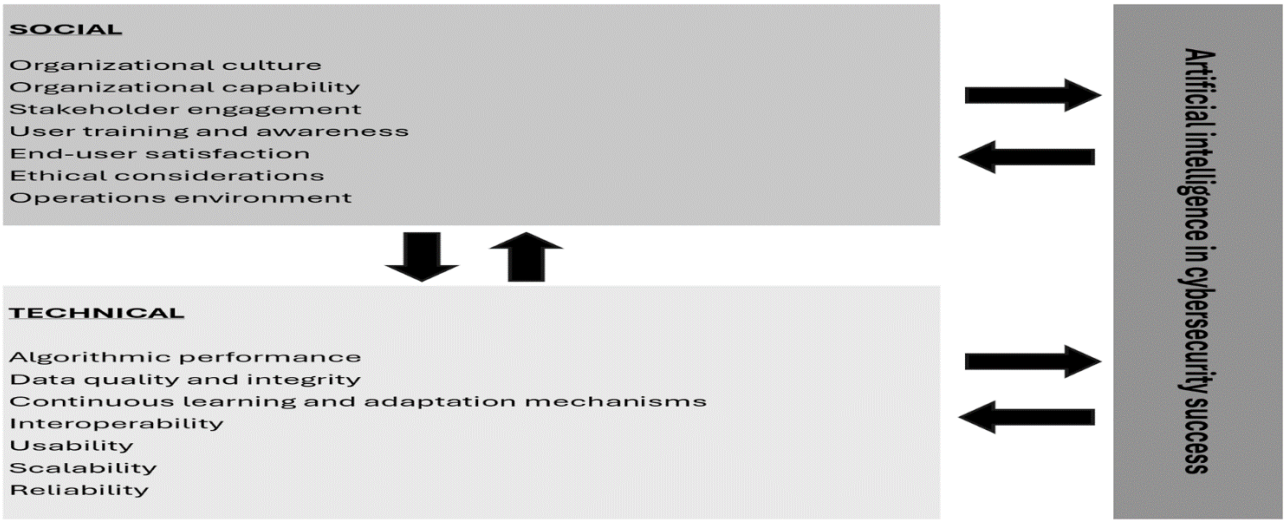


Figure 2 Socio-Technical Critical Success Factors (STCSF) Model

This model identifies and integrates key success factors from both the social and technical domains. There is an interplay between these two domains and some elements overlap between the domains. Both these domains feed into the success or failure of AI in cybersecurity, the outcome also feeds back to the domains. The identified social and technical factors for AI cybersecurity success are discussed below.

Social factors: Organisational culture is the core of organisation operations, moulding employee behaviour and success in general, this can have influence on the success or failure of AI cybersecurity. And the organisational capability is based on how effectively and efficiently an organisation uses its resources, the utilisation of knowledge, processes, people, and technology resources has an effect in AI cybersecurity success (Mwim & Mtsweni, 2022; Uchendu et al., 2021). Communication is key and engagement with all the relevant stakeholders has an impact in the success and adoption of AI cybersecurity and user training and awareness when done in a good way can promote awareness, reduce cybersecurity fatigues and improve the AI cybersecurity success rate (Gordon et al., 2021; Mwim & Mtsweni, 2022; Sony & Naik, 2020). Aanestad & Jensen (2016) and Barker & Frolick (2003) argue that user satisfaction is reliant on technical and social factors like usability, stakeholder engagement, cybersecurity awareness and training. Chatterjee et al., (2024), Kaushik et al. (2024), and Lapointe & Rivard (2007) assert that ethical environmental considerations are crucial when implanting AI cybersecurity as it will impact social aspects like privacy concerns and it also involves

the regulations that influence an organisation's operations and can have an impact in the success of AI cybersecurity.

Technical factors: The algorithmic performance is more focused on the accuracy, the strain on operations and the timeliness of the AI cybersecurity systems to detect and act on anomalies (Erhan et al., 2021; Salem et al., 2024). The data quality and integrity are due to the criticality of reliable, accurate, timely and useful data for decision-making. And continuous learning and adaptation mechanisms need to be embedded in the AI cybersecurity system so that they can continuously learn, improve and adapt (Kaushik et al., 2024; I. H. Sarker, 2024; I. H. Sarker, Janicke, Mohsin, et al., 2024). Shamsan Saleh (2024), Ozkan-Okay et al (2024) and Alneyadi & Kassim (2023) argue that AI cybersecurity should blend and be Interoperable with the existing systems, should be usable and not be cumbersome to use, should scalable for growth or decline in the organisation and lastly should be reliable with very minimal false positives and negative.

2.10 CHAPTER SUMMARY

In this chapter, the literature was reviewed for the critical success factors of AI in cybersecurity form the socio-technical lenses. The potential empirical theories for this study and where they have been used in cybersecurity studies are discussed, and the reasons for not using them for this study provided as they have some shortcomings. The chosen empirical theory was discusses and the Socio-Technical Critical Success Factors (STCSF) Model that was developed. The research methodology will be discussed in the next chapter.

3 CHAPTER 3. RESEARCH METHODOLOGY

3.1 INTRODUCTION

This chapter outlines the methodology describes the methodology employed to investigate the critical success factors for AI cybersecurity implementation from a socio-technical perspective within a South African state-owned entity operating in the higher education sector. This chapter discusses in detail the research method, research paradigm, and research design. Furthermore, it explores the research method, data collection and data analysis used with the aim of understanding the socio-technical critical success factors of implementing AI in cybersecurity and answering the research question. Lastly, it discusses the sample demographics, and the ethical considerations thoroughly addressed throughout the research process. The research methodology followed in this study is presented in Figure 3.



Figure 3 Research Methodology for The Socio-Technical Critical Success Factors of Implementing AI in Cybersecurity

3.2 RESEARCH PHILOSOPHY (PARADIGM)

Myers and Klein (2011) describe research philosophy as underlying assumptions or beliefs of researchers that influence their philosophical view of the world, which influences the way they approach and conduct their research. Saunders et al. (2019) that recognising and understanding their philosophical views and preconceptions is crucial for ensuring the soundness and effective evaluation of their research endeavours (Saunders et al., 2019). The research philosophy is grounded in a specific set of opinions about the physical, social and the researcher's perspectives. These beliefs have a significant influence on the course of their research. The interplay among the three domains (physical, social, and personal) shows that each domain possesses distinct modes of existence and diverse epistemological potential: the physical world is external and autonomous from human influence; the personal world includes individual encounters, opinions, sentiments and perceptions; and the social world covers what we collectively partake in and share (Mingers, 2001; Stahl, 2012).

Hassan et al. (2018) assert that to differentiate between various research philosophies, attention is directed towards the underlying assumptions that set them apart, namely ontology, epistemology, and axiology. Researchers have different key views about the nature of the world (ontology) and what knowledge is, the nature of knowledge (epistemology) and the ways to gain knowledge (Oates, 2005). Axiology is more about the role of ethics and values applied in the research process, and it considers what the researcher should do and how they should do it in ensuring what values go into knowledge (Mingers, 2001). To further distinguish between various philosophies, researchers consider the research paradigm. Chen and Hirschheim (2004) posit that a research paradigm represents philosophical assumptions and principles that guide research design, and that some of the widely used paradigms in information systems include positivism, pragmatism, and interpretivism. This study adopts the interpretive paradigm.

Oates (2005), Walsham (1995), and Bhattacherjee (2012) argue that positivism is a scientific paradigm that emphasises the study of empirically observable events and aims to reveal knowledge that is both objective and quantifiable, it assumes that science or knowledge creation should be limited to what can be observed and measured. They posit that it relies exclusively on theories that are directly testable and tends to dismiss any effort to extend or reason beyond observable facts. They further assert It assumes that what can be understood about the world is independent of any individual, irrespective of their identity or personal experiences. The positivist researcher strictly focuses on the

empirical scientific method to generate facts and data which is independent of human interpretation or bias. This renders positivism less appropriate for studies of the social world which may include organisational and human behaviour (Bhattacherjee, 2012). While this paradigm provides valuable quantitative insights into this study, it is important to recognize that it could not fully capture the qualitative, subjective, and contextual aspects of this research, because this research explores human behaviour, stakeholder interaction, organisation and would also involve interpretation of the interviews results.

Goldkuhl (2012) posits that pragmatism is a research paradigm that places emphasis on achieving specific goals or outcomes through practical and effective means, it prioritises the utilization of methods, strategies, or solutions that have demonstrated success or utility in real-world scenarios. Pragmatism is associated with the notion of using knowledge to make impactful differences in real-world scenarios, it is more concerned with why things work or do not work and providing practical solutions (Baskerville & Myers, 2004; Bhattacherjee, 2012). Goldkuhl (2012) furthers the argument by asserting that the pragmatic researcher focuses on creating data and theories, often through design and action research. The role of the researcher in pragmatism is to promote change by the actions of the researcher, while in interpretivism, the researcher is engaged in understanding how the context influences the research problem (Myers & Klein, 2011). The pragmatism paradigm provides valuable insights into solving problems, however it is less appropriate for this study as this study explores the organisation, its processes, human behaviour, stakeholder interaction, as well as interpretation the data collected.

Interpretivism is a research paradigm that seeks to understand the subjective meanings of individuals in a particular area of study, it seeks to explore how individuals interpret and derive meaning from the events and phenomena in their surroundings, and it places significance on understanding the personal experiences, meanings, and interpretations that individuals and groups assign to what is being studied (Chilisa, 2019; Goldkuhl, 2012; Walsham, 1995). Interpretivist research in the field of information systems (IS) focuses on understanding human actions and interpretations surrounding the development and use of computer-based information systems. This approach may involve interviews, triangulation, questionnaires, and observations, where researchers gather data over a certain period (Goldkuhl, 2012; Myers & Klein, 2011; Walsham, 2006). After investigating the different research paradigms, this study adopted the interpretivist philosophy, mainly because it is

deemed appropriate for answering the research question, and exploring the critical success factors, meanings, perspectives, and interpretations that individuals associate with AI cybersecurity. For an example a cybersecurity implementors might view the implementation of AI cybersecurity as a success if it seamlessly integrates with existing systems, and an end-user might be of the view that the success is reliant on their daily activities being not disturbed. The five basic attributes of the interpretivist paradigm (Oates, 2005) can be identified in this research and are depicted in Table 1.

THE FIVE BASIC ATTRIBUTES OF THE INTERPRETIVIST PARADIGM		
1	Study of people in their natural social settings.	This study involved interviews with the purpose of understanding how the participants come across with AI cybersecurity in the socio-technical context in which they operate.
2	Multiple subjective realities.	Recognizing that different stakeholders have diverse perspectives, the sample included different stakeholder groups.
3	Dynamic, socially constructed meaning.	This study explored how meanings and practices change overtime, taking into consideration that AI in cybersecurity is continuously evolving, and its effectiveness and success is shaped by ongoing interactions among users, the technology, and the organisational context.
4	Multiple interpretations.	Acknowledging that different users and stakeholders may interpret the success factors in different ways. The study employed thematic analysis to identify common themes and variations in stakeholder interpretations of AI cybersecurity implementation.
5	Qualitative data analysis.	To generate detailed understanding of the socio-technical factors influencing the implementation success of AI in cybersecurity, informing more nuanced and effective strategies for implementation and improvement, this study employed the Atlas.ti tool for data analysis.

Table 1 The five basic attributes of the interpretivist paradigm (adapted from Oates, 2006)

Interpretivism assists in grasping the perspectives and comprehension of various stakeholders regarding AI cybersecurity critical success factors. The study examined the socio-technical elements that shape AI cybersecurity implementation. Additionally, in pursuit of proving the credibility of the research interpretations and findings to the research readers, the presented arguments are backed by evidence from collected data. The evidence data was collected in the form of interviews (Creswell, 2014).

3.3 RESEARCH APPROACH

A research approach is a crucial phase in the research process as it serves as a comprehensive roadmap outlining how a research study will be executed, this includes selecting a research method, operationalizing constructs of interest, and devising an appropriate sampling strategy, it provides a structured framework for the collection of data, allowing for an efficient investigation of the research hypothesis (Creswell, 2014). Rahi (2017) and Bhattacherjee (2012) assert that the choice of research approach should depend on the nature of the research phenomenon being studied, with different designs being more suitable for exploratory, descriptive, or explanatory research.

This study employed an exploratory research approach. This choice was considered appropriate for the study because the exploratory approach is adaptable and open-ended nature, which allows for a dynamic approach to designing and testing solutions, especially in situations where the problem or requirements are not clearly defined or understood at the outset (Fowler & Horan, 2007). At the heart of the exploratory design lies the iterative process, which permits continuous learning and adaptation (Creswell & Clark, 2011; Fowler & Horan, 2007). User-centred methodologies, including user research, interviews, and observations, hold a significant role in this approach, aiming to gather valuable insights into the needs, behaviours, and preferences of both end-users and ICT professionals. Through the exploratory design, the study aimed to attain a deeper understanding of the critical success factors for AI in cybersecurity within the organisational context.

3.4 RESEARCH METHOD

There are three primary methods that researchers use to collect and analyse the empirical data; the quantitative, qualitative and mixed methods. The method that the researchers select is determined by the research paradigm and the approach best suited to answer the research question (Creswell, 2014). Quantitative research method utilizes experiments with the aim of collecting numerical data with score or metrics where in the researcher is independent of the research setting (Bhattacherjee, 2012; Creswell & Clark, 2011). Whereas qualitative research usually utilizes methods which require rich description of individuals and culture and aims to gather more non-numerical data in the form of words such as interviews and observations and the data is collected in natural settings of participants (Creswell & Creswell, 2017) Mixed methods research combines both elements of quantitative and qualitative research and it is ideal for studies that aim to produce quantifiable and qualitative data (Creswell, 2014)

The study aimed to obtain valuable insights into end-users' and ICT professionals' socio-technical success factors, viewpoints, and hurdles concerning AI in cybersecurity implementation. This study employed qualitative research approach to gain a more comprehensive understanding of the research problem. The participants were engaged through semi-structured interviews to obtain valuable insights into their critical success factors viewpoints, and hurdles concerning AI in cybersecurity. Qualitative methods play a crucial role in understanding the complex experiences, perspectives, and behaviours of individuals and groups engaged with or impacted by the incorporation of AI in cybersecurity (Fujs et al., 2019).

3.5 RESEARCH DESIGN

The study adopted the case study research strategy. A case study is relevant for a studying the success factors of AI in cybersecurity implementation from a socio-technical perspective, owing to the complexity of AI adoption where social and technical factors intertwine, case studies capture these interplays in a real-world setup (Shakir, 2002). Morland et al. (1992) argue that case study research involves the examination of a phenomenon within its natural context over a specific time frame, gathering evidence through methods like interviews and observations. This approach proves advantageous as it allows the study to holistically uncover the socio-technical success factors, challenges and the lessons learnt implementing AI in cybersecurity, thereby enabling best practice identification and theory development Paré and Elam (1997). Case studies have been extensively used

over an extended period, yet certain researchers have raised concerns about their effectiveness as a research method due to misconceptions regarding their capacity to generate reliable and generalizable knowledge, most of the critics scrutinize the limited sample sizes typical in case studies (A. S. Lee, 1989; Saunders et al., 2019).

Walsham and Waema (1994) assert that relying on a case study to make inferences about a specific field of study aligns with an interpretive epistemological standpoint. From an interpretive perspective, the soundness of a conclusion from one or more individual cases is not dependent on their representativeness in a statistical context, but rather on the persuasiveness and soundness of the logical rationale employed in presenting findings from the case, as well as in arriving at conclusions based on those findings (Walsham, 1995). Interpretive case studies in IS research aim to contribute valuable insights and understanding of the human interpretations and meanings associated with computer-based information systems, and to encourage careful work on the conceptualization and execution of case studies in the information systems field (Creswell & Clark, 2011; Walsham, 1995). The most appropriate research strategy for this study was a case study at Edu X, the reason for choosing this organisation is based on its accessibility to the researcher. The research outcome was anticipated to be useful to Edu X in providing input into its cybersecurity implementation, management and the improvement of the organisation's cybersecurity posture.

3.5.1 Case Study Background

The organisation to be studied, Edu X is a state-owned entity operating in the higher education space. Edu X is responsible for but not limited to overseeing the national qualification framework (NQF) implementation which is a key national policy tool for South Africa’s education and training system. There has been particular interest in this institution due to its implementation of AI in cybersecurity to secure its operations and the critical data related to the learner records of the whole population of South Africa. Edu X’s ICT division at executive level is led by a chief financial officer, at senior management level it is led by a senior ICT manager responsible for ICT operations and a senior ICT manager responsible for ICT projects. At middle management and senior specialist level the organisation has an ICT manager, senior specialist cybersecurity, senior specialist business analyst , senior specialist software development. At specialist level the entity has a specialist software developer and a ICT specialist networks. At practitioner level it has two ICT technicians, and at administration level it has an office administrator. The organisation has adopted a culture whereby

cybersecurity is not solely the responsibility of the senior specialist cybersecurity rather a shared responsibility in the ICT division and in the organisation at large. Figure 4 depicts the ICT division organogram of Edu X.

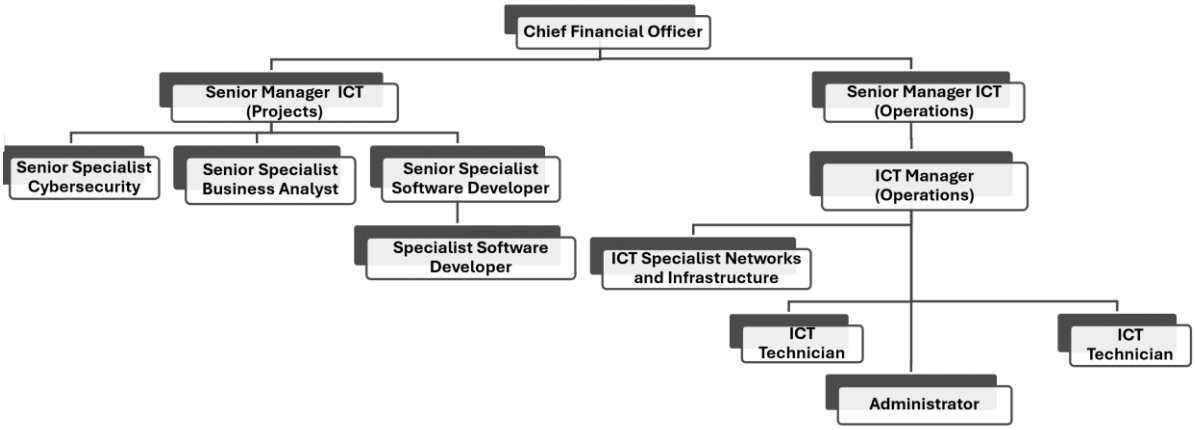


Figure 4 Edu X ICT Division Organogram

The case study focuses on understanding the critical success factors of Edu X's AI cybersecurity implementation. Edu X started the implementation of AI cybersecurity in its environment in October 2023 and is continually improving the implementation in pursuit of improving its cybersecurity posture. It has implemented an AI-based threat detection system responsible for the detection and resolution of threats, end-user cybersecurity awareness and training, anomaly detection systems responsible for identifying and resolving security incidents from outside and from within the organisation, and dependable authentication for ensuring that identities are not compromised, and users are verified. These systems use machine learning and deep learning to gather data from data sources like user behaviour, network activity, application activity, database activity and ICT infrastructure activity. The systems learn to detect anomalies and automatically resolve them, they also provide visual interfaces for notifications, dashboards and reports for human use and intervention. Using rules, the systems also learn the patterns from human interventions on how to deal with similar situations in the future. The implementation of the systems ranges from social factors such as learning use behaviours which needs to be done in an ethical manner, to technical factors such as security orchestration, automation and response. Figure 5 depicts the AI cybersecurity architecture of Edu X.

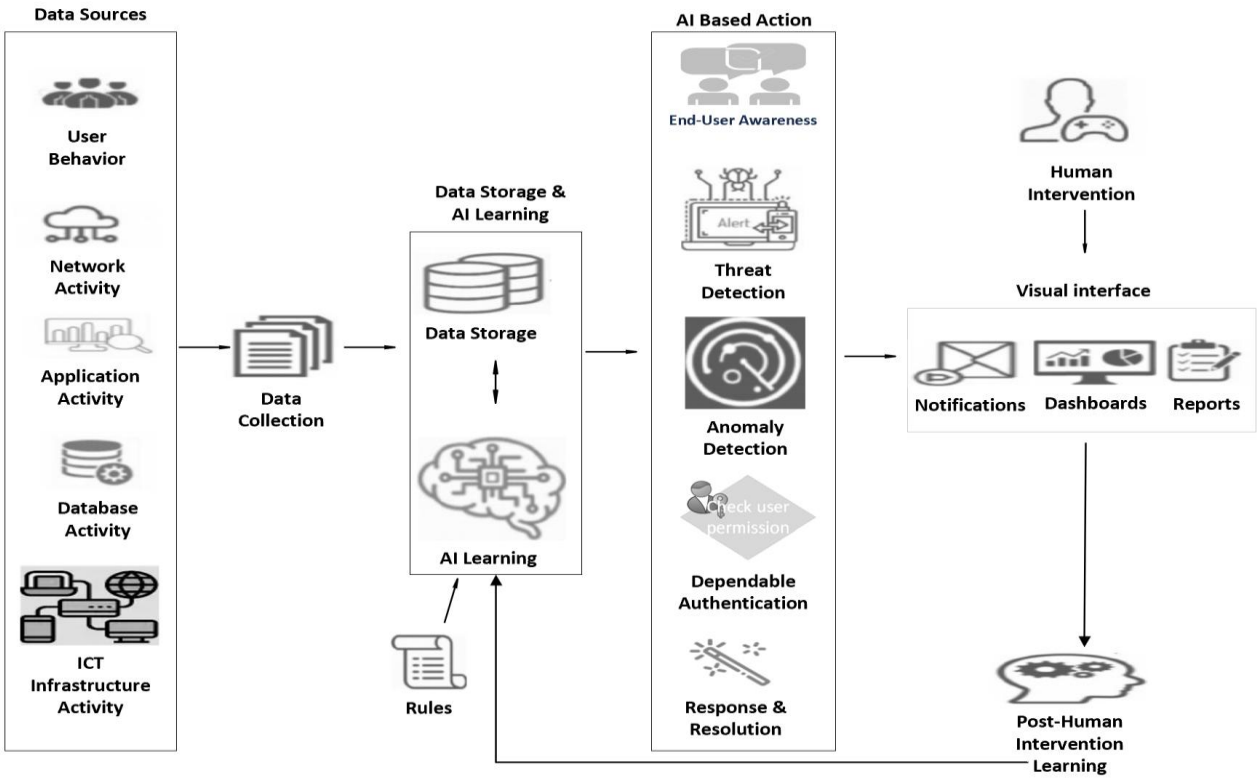


Figure 5 AI Cybersecurity Architecture of Edu X

The Edu X AI cybersecurity solutions ingest data from user behaviour, the network, applications databases and ICT Infrastructure activities. The collected data is then processed and stored for the purposes of learning and interpretation. The systems use the collected data to perform actions. For instance one of the solutions, the security information and event management (SIEM), learns the patterns and behaviour of the users or entities in the environment. It may learn that a particular user normally logs in to the organisation systems from specific locations and duration certain hours, should the SIEM tool detect the user tried to login from different countries within and impossible travel time between those countries it would classify that as an anomaly and log an incident. The SIEM would then invoke security orchestration, automation, and response (SOAR) to either temporally revoke the user's access, require multifactor authentication or require administrator (human) intervention. In the case of administrator intervention, the SIEM/SOAR would learn the intervention made and apply in similar situation. The SIEM makes use of dashboards, alerts and reports to visualise output of its operations.

3.6 RESEARCH POPULATION AND SAMPLING

The employees at Edu X are the target population for the study. Edu X has an approximate number of 130 employees. According to Myers (2013), in qualitative research, the objective is to choose participants who possess unique insights, experiences, or conditions that the study aims to understand. For this study, purposive sampling was be employed to enlist potential participants. Creswell (2014) describes purposive sampling as a non-random sampling technique where participants are selected based on specific characteristics or qualities that are relevant to the research study.

The goal of purposive sampling is to select participants who can provide valuable and in-depth insights into the phenomenon being studied (Creswell & Clark, 2011; Saunders et al., 2019). Purposive sampling is applied when the focus is on description rather than statistical generalization (Saunders et al., 2019). The use of purposive sampling aligns well with the objectives of this study and the sample for this study included managers within the organisation, ICT personnel, and end-users. The reason for this sample was to find get different viewpoints of different stakeholders in the organisation. The sample of the study consisted of seventeen participants, and table 2 below depicts the participants.

Roles in the organisation	**Number of participants**
Senior ICT Manager	2
Executive Manager	3
Software Developer	1
ICT Technician	1
ICT Specialist	1
Supply Chain Specialist	1
Senior Manager Risk & Governance	1
Manager Recognition	1
Verification Specialist	3
Manager MIS	1
Data Specialist	1
Total	17

Table 2 Organisational roles included in the sample

3.7 DATA COLLECTION

The study employed in-depth semi-structured interviews because they offer both the researcher and the participant greater flexibility (Saunders et al., 2019). Additionally, this format allows the researcher to pursue particularly interesting lines of inquiry that may arise during the interview, while enabling participants to provide a more comprehensive perspective. In this study, data was collected through interviews. Interviews were conducted, digitally recorded and transcribed using Microsoft Teams, and then arranged into themes and categorized for easy interpretation. An interview protocol was used to guide the data collection process, this interview protocol can be found in appendix A. The interviews were conducted between November and December 2024, the total duration of the conversations was 619 minutes 20 seconds. Table 3 show the duration of the interviews.

Description	Quantity
Number of interviews	17
Duration of interviews	619 minutes 20 seconds
Longest duration	66 minutes
Average duration	36 minutes 40 seconds
Shortest duration	17 minutes 58 seconds

Table 3 Duration of interviews

Interviews are a personalized form of data collection method that involves interviewers using a standardized set of questions to gather information from respondents (Creswell, 2014). Unlike questionnaires which do not necessarily require human interaction, interviews allow for special instructions for the interviewer, space for personal observations, and the opportunity to clarify issues or ask follow-up questions (Bhattacherjee, 2012). Semi-structured interviews are formally structured but also give allowance for the interview to explore more information (Creswell, 2014).

Despite its strengths, interviews are subject to biases such as non-response bias, sampling bias, social desirability bias, and recall bias and these can impact the reliability and credibility of the collected data (Bhattacherjee, 2012; Creswell, 2014).The study anticipated the biases that could arise in the collection and analysis of data, and for the credibility of the research outcome the study employed

data source triangulation. Yin (2017) posits that data source triangulation in research refers to the use of multiple sources or methods to validate and corroborate findings, in the context of qualitative research, data triangulation involves using a variety of data sources, such as interviews with different people, observations, documents, and surveys, to ensure the credibility and trustworthiness of the research findings. This approach helps to mitigate the limitations of relying on a single data source and enhances the overall rigor of the study (Bhattacherjee, 2012; Creswell, 2014).

Due to the sensitivity of cybersecurity, this study did not include other data source triangulations like observation and document analysis but rather employed the participant triangulation where participants from different levels and departments in the organisation were interviewed with the aim of getting varied insights from different groups thereby enhancing the credibility and richness of the data. The use of data triangulation and interview protocol eliminated the threats to data validity and reliability. Creswell (2014) posits that data validity and reliability are key factors in research design, data validity is more concerned with ensuring that necessary measures are taken to ensure accuracy in the collection and analysis of data, and data reliability is more concerned with the consistency in data collection and analysis (Saunders et al., 2019).

3.8 DATA ANALYSIS

Qualitative data analysis involves the interpretation and understanding of non-numeric data, such as text data from interview transcripts, focus group discussions, narratives, and other forms of qualitative data (Goodrick & Rogers, 2015). Unlike quantitative analysis, which is driven by statistical methods, qualitative analysis relies heavily on the researcher's analytical and integrative skills and their personal knowledge of the social context in which the data was collected (Bhattacherjee, 2012; Creswell, 2014; Goodrick & Rogers, 2015).

In this study, thematic analysis was employed to scrutinize the data. Nowell et al. (2017) assert that thematic analysis is a method of analysing qualitative data that involves identifying, analysing, and reporting patterns (themes) within the data, they further affirm that thematic analysis serves as a qualitative technique for recognizing, examining, and presenting recurring patterns (themes) within the data. They further argue that the initiation of the data analysis process involves the identification of themes and the exploration of relationships between these themes. Moreover, it is crucial to ensure that the chosen method of data analysis aligns with the research paradigm, the nature of the data, and

can effectively address the research question. The qualitative data from interviews was analysed using the Atlas.ti qualitative analysis system. The data that was regarded as having a potential to contribute to answering the research questions was marked with appropriate codes. These codes, when combined feed to the sub-themes, which also feed to the overarching themes. Figure 6 depicts a codification word cloud generated from Atlas.ti, containing thematic codes from the transcribed data.



Figure 6 Sample codification word cloud

Braun & Clarke (2006) argue that the outcome of thematic analysis should explain the primary patterns of meaning found in the dataset. The process of data analysis began with re-evaluating the data collected, looking for codes, sub-themes, themes and generated overarching theme.

3.9 ETHICAL CONSIDERATIONS

Permission was requested from Edu X to conduct research and collect data on the organisation and its employees, and the permission was granted which then allowed for the data to be collected and

this can be found in appendix E. The University of the Witwatersrand requires all studies involving humans to first go through the steps of receiving clearance from the ethics committee before the study is conducted, and the researcher applied and was granted an ethical clearance from the University of Witwatersrand School of Business Sciences Ethics Committee to conduct this study. This can be found in Appendix D.

All participants were informed about the aims and objectives of the study. The respondents were made aware of their right not to participate in this study and that they could withdraw their participation at any time should they have wished, also the participants had to give informed consent to the study. Obtaining informed consent implied acknowledgement of the goals of the study, the expected duration of the respondent's involvement, how the study was to be conducted, all the details of the study and the researcher's credibility to respondents (Saunders et al., 2019). The participant information sheet that gives the potential participant background of the study can be found in Appendix B and the consent form can be found in Appendix C.

To adhere to the research principles, all possible means accessible to the study of protecting the rights of the researcher's institution, the researched institution and the privacy of the study participants were applied according to the ethical guidelines. Pseudonyms were used to protect the identity of the participants, and Edu X is not the real name of the case study institution. The data that was collected and processed for the study of the socio-technical critical success factors of AI in cybersecurity implementation and is stored in a secure manner in a folder that is accessible through the use of user authentication, the data will be discarded in the form of deletion in line with the ethical clearance.

3.10 CHAPTER SUMMARY

This chapter discussed the research paradigm that was used in this study, which is the Interpretivism paradigm. It outlined the research approach and the methodology that was used for this study. The method of collecting data, participant sampling and the data analysis were discussed in this chapter. The research strategy which is case study research has been discussed and a background to the case study at Edu X has been discussed. Lastly this chapter discusses the ethical considerations. In the next chapter the findings will be presented and analysed.

4 CHAPTER 4. PRESENTATION OF FINDINGS

4.1 INTRODUCTION

This chapter will present and analyse the findings from the that was collected from the interviews with Edu X employees, which led to the creation of the thematic analysis map that visualises the overarching theme, themes, sub-themes and codes that were discovered in relation to the research question about the socio-technical critical success factors in the implementation of artificial intelligence in cybersecurity. The critical success factors that were identified from the participants' engagements are social factors, technical factors and the socio-technical interplay. These factors are outlined in the following sections.

4.2 ANALYSIS AND PRESENTATION OF THE FINDINGS

Sarker et al. (2019) and Kayworth & Whitten (2010) emphasise that paying attention to both technical and social aspects and the interplay between them leads to achieve a balanced success. An analysis of the data revealed key socio-technical factors that are crucial in the implementation of AI cybersecurity and the interplay between them. The thematic maps expands from and links to the STCSFs model that was presented in figure 2, the model highlights that social and technical components are interlinked, and to achieve a successful implementation of AI cybersecurity in an organisation it is important to balance both. Figure 7 presents the thematic analysis map developed for this study.

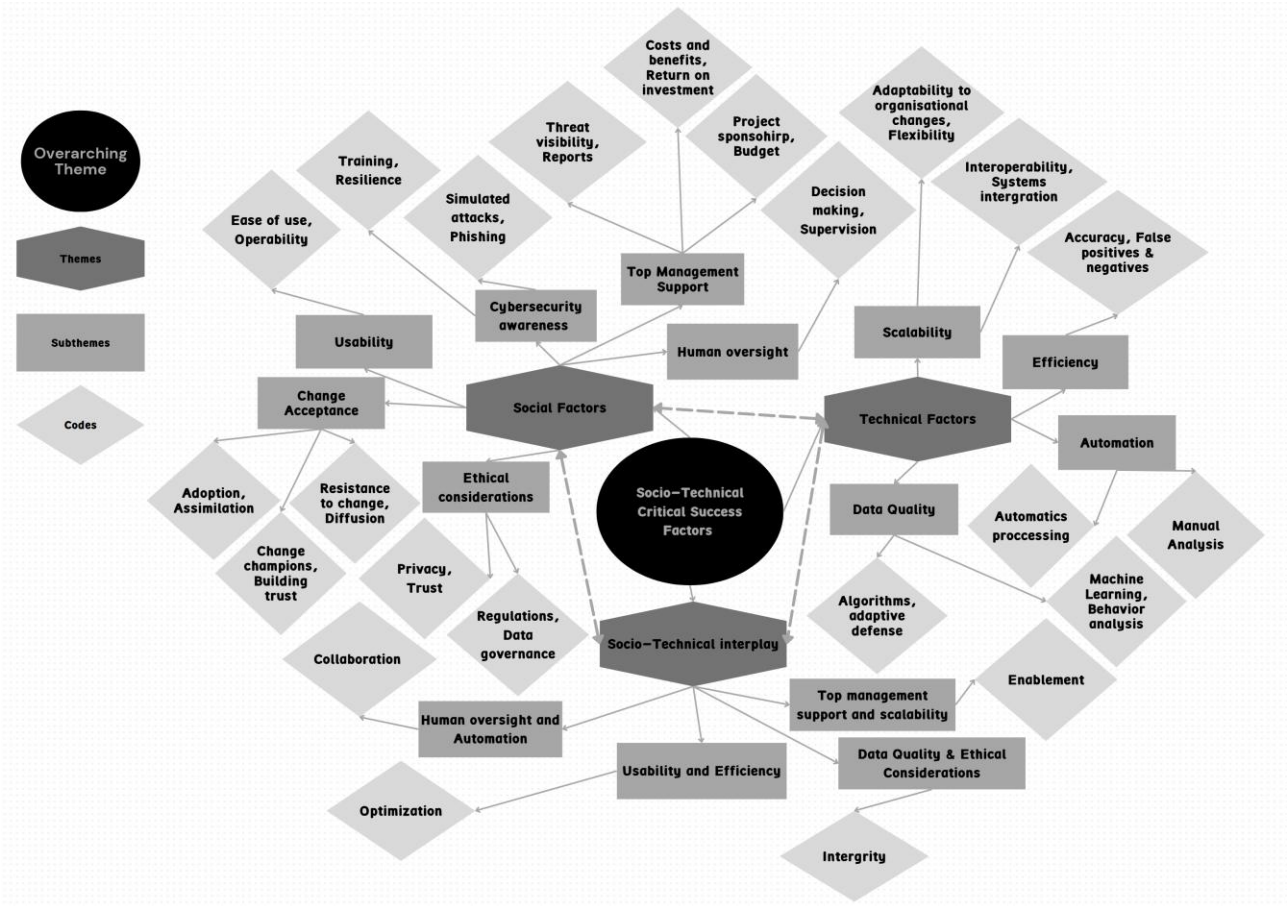


Figure 7 Thematic analysis map

4.3 SOCIAL FACTORS

This section discusses six key social factors critical for successfully implementing AI in cybersecurity namely change acceptance, top management support, cybersecurity awareness, ethical considerations, human oversight, and usability. Each factor is examined in detail to assess its role in the successful implementation of AI in cybersecurity.

4.3.1 Change Acceptance

The data analysis revealed that the implementation of AI in cybersecurity involves introducing changes in the organisation, the introduction of change has been noted a process that and it is prone to challenges and needs careful planning. Participant Alfred Manyika mentioned that they used a three-pronged approach when introducing AI in cybersecurity.

"Firstly, it is the people as people aspect. People didn't want to change, also in terms of the staff members that we had, there was not enough skill to support the changes, and a clear understanding of how the impact of the AI would be on the systems… And then the second aspect was the processes. We didn't have the clearly defined, processes and procedures. So we had to create new policies and new procedures… then lastly the technology part. We didn't have enough supporting technology"

The three elements of the three-pronged approach referred to by Manyika are the people, processes, and technology. This approach aligns with the findings of Malatji et al. (2019) who identified that the socio-technical cybersecurity implementation combines the technical and the social factors for a well-balanced result. The organisation also adopted a phased approach when implementing AI in cybersecurity, this approach meant not introducing changes all at once. Participant Alfred Manyika elaborates on this:

"The strategies that we used were sort of like an agile approach in which we had to break down structures into smaller bits and pieces and then implement them slowly so that the users can get used rather than for us just coming up with a big bang approach."

Manyika highlights the importance strategy of slowly integrating AI into cybersecurity so that the users would be accustomed to the changes as time went on thereby avoiding resistance from the impacted stakeholders caused by a sudden change. This would also afford time for the technical personnel to assess the impact on the technology part to ensure readiness for a full-blown rollout. Participant Rebamang Lekoelea mentioned the importance of assessing the environment for readiness before implementing changes.

"We have look at the readiness of the organisation in terms of implementing such solution, is the organisation ready for such or maybe is a big jump for an organisation to implement such a solution"

The analysis revealed that as part of stakeholder readiness preparation, the changes to the environment on enhancing cybersecurity were communicated with the users before the actual implementation took

place so that the end users might know what to expect. Participant Joice Mphahlele, a senior manager in ICT emphasised the importance of stakeholder engagement by saying:

> *"In terms of preparation, communication was prepared well in advance, to sensitize these users as to what to expect in the coming few weeks and leading right down to the actual cut-off date, and I think that has provided some level of comfort to the users"*

Participant Amelia Smith agreed with Joice Mphahlele:

> *"I remember, it was an e-mail from the senior manager IT to inform us that these interventions will be enrolled, and every time a new one is being enrolled"*

This reveals implementation of AI in cybersecurity is not just a technical change, it is also a social one, and communication is the cornerstone of this. It shows that communication fosters collaboration, understanding and trust among employees, it provides clarity on expectations, roles and responsibilities. Mphahlele highlighted clear communication and change management as some of the lessons learnt from this implementation.

> *"Communication is key, as well as organisational change management"*

Participant Alfred Manyika added to what Mphahlele said and emphasised the importance of stakeholder engagement for the success of the AI cybersecurity intervention:

> *"We need to prepare the stakeholders way before we start doing the implementation, one of the most important parts of the project success is actually heavily dependent on proper communication being given to the stakeholders, because once the stakeholders don't fully understand what exactly needs to happen. It becomes very much challenging, so stakeholder engagement is very important for us to succeed."*

Alfred Manyika emphasises the importance of planning and engaging with stakeholders to ensure change acceptance. The analysis found communication also involves soliciting feedback from end-

users on the experience of rolled-out solutions. Participant Joice Mphahlele pointed out during her interview:

> *"We also run surveys and based on those surveys as well, we're able to gauge the user's satisfaction on these projects that we roll out"*

This shows that requiring feedback from the impacted stakeholders provides the implementors with information on the areas of improvement. However, some solutions run in the background and are not noticeable by end-users and therefore do not warrant feedback unless errors arise. Participant Lynn Manuel pointed this out in her interview:

> *"I think my experience is probably more than I would know, because these things don't happen necessarily while I'm paying attention, some run in the background"*

The analysis found that implementing change to a work environment is not always well welcomed by people who are used to a certain way of doing things, and that to be successful in implementing cybersecurity changes that have an impact to end-users, the senior management needs to be involved. Participant Alfred Manyika pointed this out in his interview:

> *"People are naturally resistant to change, and we had to do a lot of education. We engaged with senior managers of the respective units in trying to make sure that everyone also does their part… We slowly changed the culture of the organisation"*

Analysis shown that there is a need for a change in the culture of doing things, and the change must be instilled and supported from top management level and be relayed down to junior levels. By so doing needs to be onboarded in cybersecurity culture change. The analysis revealed that as part of culture change the organisation also made use of the change champions. The term 'Change champions' refers to the end users who are in other business units outside of the ICT unit who are technology proficient. Participant John Mokwena who is a change champion elaborates this role:

"It is basically a group of people that were nominated from each division who are to basically soften the blow to the employees that they work with in terms of letting them know of any major changes that would be happening also giving feedback on the experience"

Data analysis found that change acceptance is critical to the success of AI cybersecurity implementation. Change implementation involves stakeholder engagements to ensure that the stakeholders are aware of their roles and responsibilities. Clear communication minimises resistance that is based on not knowing what is happening and what to do. From the quotes it is evident that implementing AI in cybersecurity needs to be approached in a way that will ensure that the people, processes, and technology dimensions are addressed.

4.3.2 Top Management Support

The data analysis found that top management is responsible for making strategic decisions in the organisation, and those strategic decisions are then filtered down the organisation into action items. Assent and support from senior management is crucial for the implementation of AI in cybersecurity and for improving the overall security posture of the organisation. Participant Lynn Manuel a chief executive officer, explained the role of top management and support on AI in cybersecurity initiatives:

"The executives are very much supporting and sponsoring these initiatives whilst not necessarily holding the technical competence. Certainly, from my side I am constantly pushing for further innovation, further security and further leveraging of AI to do things better… Certainly from the executive's perspective it's a non-negotiable, we sponsor and we enable these functionalities."

The analysis found that the support can range from culture enforcement to budget and resource allocation. The allocation of budget is led by top management of the organisation, they are the custodians of the budget and from that budget they sponsor the organisational operations. Participant Adrian Makwetla expressed his view on the importance of having approved budget:

"The lesson I learnt in this is that you must make sure that you got budget. Enough budget to be able to remediate cybersecurity issues. You need to need money to get enough people with expertise."

Makwetla highlighted a crucial enabler aspect by indicating that the success of AI in cybersecurity needs budget to buy the necessary technology, subscriptions, and source the right people with skills. It is worth noting that top management do not necessarily have the hands-on operational experience of cybersecurity and may at times have oversight on the importance cybersecurity implementation if not given the relevant information or visibility. Participant Alfred Manyika expressed this view in a statement:

> *"Some of the lessons that we've seen is that some of the sponsors don't really know how and what is the impact in terms of the things that you want to implement and also if these stakeholders don't fully understand, it becomes difficult to secure a buy in in terms of the project."*

The analysis found that the top management have the responsibility to account for the use of funds allocated to the organisation. As part of accounting, they have to weigh the costs and benefits of AI for cybersecurity implementation and try to achieve return on investment. Participant Alex Bhingwa, an executive for finance, explains the return on investment of this implementation from an executive point of view:

> *"The return on investment is huge… I see the value for money. We do not want to find ourselves being hacked and our data falling into the wrong hands, because the cost of not having this is massive and the organisation might not afford these ransomware and stuff."*

Besides providing resource-based support, it has been found that top management also play an important role in fostering culture change. The analysis found that the introduction of a culture change needs the involvement of top management of an organisation to use the power vested in them to ensure that culture change is implemented. Participant Joice Mphahlele pointed out the importance of top management support:

> *"The strategy was to make sure that we have top management buy in and also to ensure that when we implement all these interventions, we move on the same page with the organisation in terms of all the staff members"*

Participant Alex Bhingwa, also supported this view and made an example of top management intervention on people who were reluctant to participate in cybersecurity awareness.

> *"If there's no strong support from senior leadership AI for cybersecurity will not be successful. We have even gone to the extent of saying those people that would not have completed their awareness trainings should be noted"*

Data analysis revealed that artificial intelligence is a subject that is relatively new at Edu X. In pursuit of understanding AI and how it can enhance business processes, the organisation established a task team to study AI's broader use in the organisation, not only for cybersecurity. Participant Gauta Mothapo an executive responsible for organisational operations highlights the existence of the task team.

> *"We put a test team together for them to be able to look at what is the impact AI within our work environment"*

From the quotes above, it is revealed that top management play a critical role in sponsoring AI for cybersecurity and ensuring that measures are in place for its rollout throughout the organisation. The top management might not be technically savvy but with engagement from the parties responsible for the implementation they can have the visibility of the benefits.

4.3.3 Cybersecurity Awareness

The data analysis revealed that cybersecurity awareness is crucial for optimizing organisational security posture. As part of the implementation, all employees were introduced to cybersecurity awareness campaigns with the aim of enhancing and personalising efforts to educate the users about cybersecurity risks, safe practices, and threat prevention. Sarker et al. (2024) found that cybersecurity awareness addresses both the social and technical aspects of cybersecurity to ensure that users understand their responsibilities in maintaining compliance, security and efficiency, and therefore bridge the gap between technical system requirements and user behaviour. These cybersecurity awareness campaigns included training, educating the users about the implications of their behaviour, recognizing phishing attempts from phishing simulations, continuous feedback and testing,

gamification and risk scoring. Participant Andiswa Modise indicated that these trainings were educational.

> *"It has really helped a lot from the layperson's point of view, it has given an understanding that you cannot think that cyber threats and vulnerabilities are only touching on ICT people"*

Participant Amelia Smith also alluded to the need for cyber awareness.

> *"I think it is important to educate us the staff, especially those of us who are not in the IT space or thinking about IT security and make us aware of possible risks."*

The analysis found that AI in cybersecurity can learn what is normal and what is not normal in the environment and when it reaches a set threat threshold, it can automatically limit the functions of the end-user, and this may affect the end-user's operations. Participant, Pule Molefe indicated he noted that emails between his work and private email addresses were not being sent, he then recalled that this was due to impersonation prevention and data loss prevention (DLP) that was implemented in the organisation, and the awareness that the users should not mix official and private emails to prevent data leakage.

> *"I wanted to download some presentations I have seen online, and then I did e-mail them to myself using my personal e-mail and the e-mail couldn't come through to my official email and I got a notification that it's like a person is impersonating of me. So I recalled that this is one of the AI security tools that has been implemented. So that was an eye opening"*

The collaborative approach of integrating both AI and end-users strengths, bas been seen as essential for navigating the complex and ever-changing world of cybersecurity. A participant Anele Moloto who is an ICT technician pointed this out:

> *"We haven't had much of our users that will just click on a link that is forwarded to them or provide information to impersonators who pretend to be senior leaders"*

Moloto highlighted that before they implemented cybersecurity awareness and security enhancement of the email systems, she would notice a high number of sophisticated phishing emails impersonating managers and executives and she highlights that employees would most of the time fall for this and divulge information. And this was due to the integration of AI and cybersecurity awareness whereby suspicious emails are automatically withheld and users know how to spot and report unsolicited emails.

Analysis revealed that huge numbers of threat detections cause a lot of work for the ICT team and impact the cybersecurity operations, and implementation of AI cybersecurity alone without cybersecurity awareness does not decrease the number of detections. This meant that there is a direct relationship between cybersecurity awareness and lower threat detection. Research by Khando et al. (2021) supports this observation, emphasising that educated users generate fewer alerts, enabling AI systems to focus on genuine threats rather than false positives. It has been found that the introduction of cybersecurity awareness training has made the end-users to be alert of things that would raise detections and alarms. Another participant from the ICT unit, Jacob Motaung emphasised the importance of cybersecurity awareness.

> *"Cyber awareness is important so that we all are on the same page and we understand what is expected of us in order to not to be compromised because cause you can implement all the security measures but unaware users can lead to compromises or a large number of alerts."*

It is evident from the aforementioned quotes that effective cybersecurity awareness is crucial for the implementation of AI cybersecurity. It has been found As artificial intelligence also monitors user behaviour to detect anomalies, it is essential to make users aware that certain behaviours when using the organisation's systems will trigger alerts. Making users aware of the evolving threat landscape and empowering them to make informed security decisions can significantly strengthen an organisation's overall security posture.

4.3.4 Human Oversight

Data analysis revealed that humans play an important role in providing oversight on AI in cybersecurity by applying critical thinking and human judgment. Participant Anele Moloto highlighted the need for human decision-making in AI-driven cybersecurity:

"While automation can be useful in some areas, it's important to ensure that a human is still involved in the decision-making process."

Human oversight is an essential part of preventing or rectifying errors from AI cybersecurity. This was revealed by the data analysis, and it was found that the humans involved should have the right skills and expertise to operate AI in cybersecurity. Participant Rebamang Lekoelea stressed the importance of having the right skills and expertise:

"It is essential to look at the skills available within the organisation when implementing AI cybersecurity tools or solutions, the right skill set in terms of managing that and other organisational challenges"

Participants Alex Bhingwa also alluded to the criticality of having the right skills, he also highlighted that it is important for the humans responsible for championing the projects to have technical expertise and an understanding of the business needs.

"It is very critical for me to have skilled personnel. People that understand or have a deep understanding of both the technical and the business side AI projects"

The analysis has revealed that at times artificial intelligence has erroneously blocked emails that were coming from legitimate sources and stakeholders. This would necessitate human intervention, where a human would conduct security checks and if satisfied with the safety of the email would release it to the recipient's mailbox or reject it if an anomaly is detected. Participant Akani Mathebula an end-user attests to AI-human oversight.

"I think one incident, if not two, where one of the emails from one of our information partners was held and then we could not locate that e-mail. It was getting withheld and we were not able to find it, but then after the intervention from IT we were able to locate it"

Participant David Radebe an end user highlights his experience with accessing certain applications that are blocked by AI cybersecurity, the human intervention turnaround time to fixing that.

"It becomes a bit of a challenge where you know that an application is genuine or an application that you need or previously used is getting blocked… This gets resolved through communication to the ICT team and there's quick response time, if maybe there was no quick response time, then I would think that this would have had an impact on my performance or on the tasks that I had to perform on the day"

Data analysis revealed that although AI in cybersecurity can work on its own there are moments when a need arises for human oversight to make decisions on complex issues. Human-AI teaming through human oversight plays and important role in ensuring that where AI has shortcomings, a human can intervene. The timeliness of human oversight is essential for continued business operations. The right skills are needed for intervening human, to ensure that that the intervention is done in a correct way.

4.3.5 Ethical Considerations

Data analysis revealed that ethics and regulatory compliance are critical factors to consider for AI in cybersecurity. In the context of South Africa, organisation that collect and process data have to be compliant with the data protection regulations and data sovereignty is one of the important aspects of dealing with data. Participant Joice Mphahlele mentioned this in her interview:

"The tools that we use, some of them are tools that are originating from outside the country, so one would have to look at when they collect this data, where does it reside and all those kinds of things."

Mphahlele stressed the need to ensure that the collected data is stored within the perimeters of South Africa. A study by Bharosa (2022) emphasised this, arguing that state-owned entities should not give up sovereignty to the market parties and that the processing of information be in line with applicable laws and regulations. Participant Alex Bhingwa revealed that the organisation is mindful of ethics and regulations, and that it develops policies to accommodate AI cybersecurity.

"To avoid unintended negative outcomes, we are putting up ethical guidelines in the form of policies so that we can regulate this."

Mphahlele added to this and mentioned some of the regulations and best practices Edu X adheres to:

> *"We comply in terms of ISO standards in setting up our ICT infrastructure, particularly for cybersecurity related activities… We comply with regulatory requirements like PoPI, that require confidentiality of information or data and that it must be safeguarded at all times."*

The analysis found that ethical considerations for AI in cybersecurity are not just a tick-box exercise for regulatory compliance, they include transparency, privacy, trust, fairness and accountability. It was found that for the success of AI in cybersecurity there needs to be some form of trust between the systems and humans. A participant Johan De Bruyn, who is a developer raised trust concerns on the implementation of Artificial intelligence.

> *"The provider can be spying on you, I think they're not transparent on how data is used and now that is not ethical."*

De Bruyn was sceptical of the privacy in AI cybersecurity and is concerned of being watched on. The analysis revealed that AI cybersecurity can monitor user behaviour through User and Entity Behaviour Analysis (UEBA) to detect anomalies. UEBA collects various data including account profiles, access, roles, permission, geographical location, and activity. This can give rise to privacy concerns from the users, and they might feel that they are being watched on. A privacy concern was raised by participant Banele Mahlangu who works in the supply chain management unit, that when emails meant for supply chain management get held and require an intervention from ICT to release them. The concern is based on the probability of the person responsible for releasing the email having access to its contents:

> *"I have a concern with work related email getting held, there is a concern for a privacy because we are in different units and there's information from supply chain that must not be shared with other units."*

From the analysis it has been discovered that ethics and regulatory compliance are delicate matters that need to be handled with care to avoid repercussions that might arise from non-compliance. Transparency and fairness in the ingestion of data contributes to the trust of AI in cybersecurity.

Storing data within the shores of the country is an important aspect in complying with South African regulations.

4.3.6 Usability

The term ‘Usability’ in this context refers to the ease of using AI in cybersecurity or working in an environment with AI cybersecurity. The study found that a solution that seamlessly integrates with the organisational operations and enhances security while maintaining intuitiveness can be viewed as user-friendly. Highly sensitive systems can raise frequent false positives that can frustrate the users. Participant Lynn Manuel raised a concern about the usability of AI in cybersecurity at any time of the day without requiring ICT intervention to resolve false positives.

> *“When these things happen, it's late at night or it's on a weekend and then you have to revert to a human being when the AI is not helping you, it's inconvenient… It delays us because now I've got to wait for somebody to download before I can move forward and read the thing and approve it, etcetera… And the particular case I'm talking about was a weekend after hours. You don't want to bother people after hours. You would hope that the technology is helping you do these things without reaching out to the people, it does create delays and it does create a bit of frustration.”*

Manual expressed unhappiness with the need for human intervention form the ICT staff over non-working hours, and that this would negatively impact business operations. Participant Banele Mahlangu also highlighted a usability concern that would come with delays that need human oversight:

> *“There's been hiccups that we've come across with the systems, especially the system that is used to detect any threats on the emails because at times you'll find that we are working on an RFQ and then the e-mail gets blocked and then we are unable to access that e-mail. So we have to go back to ICT to request for access to that e-mail. So it does delay us in a way, but then I think it's good for the for the organisation to prevent any threats.”*

From a technical side, participant Anele Moloto indicated that she has been part of the team implementing these solutions and from her experience, she has not noted usability challenges.

"I was part of the team that implemented these cybersecurity solutions that utilize AI… I find it easy to view data, incidents and other resources"

The data analysis revealed usability of AI cybersecurity can mean different things to different groups of people. To the end-users it could mean the comfort of working without disturbance or delays raised by detected anomalies and for the technical staff it could mean ease of setting up and operating the AI cybersecurity to achieve set goals.

4.4 TECHNICAL FACTORS

This section discusses four technical factors critical for successfully implementing AI in cybersecurity namely data quality, scalability, automation and efficiency. These factors will be examined in detail to determine their effect on the implementation of AI in cybersecurity success.

4.4.1 Data Quality

Data analysis revealed that there can be inherent biases in AI cybersecurity that pose a risk of discriminative or unequal outcome of threat detections. Participant Lynn Manuel, noted the bias in the detection of potential malicious emails from the Edu X African counterparts.

"People that develop these cutting-edge tools have a certain perspective. That perspective is fed into the AI algorithms, and the African marginalized people are the ones that generally take the brunt of that. So I think we must use AI very circumspectly and we must always be watching out for those biases… like AI blocking all African things because Africa is bad."

Participant Alex Bhingwa alluded to this in his interview:

"I've always worried about data integrity issues when we're looking at this thing. I mean it has inherent biases, if we over depend on this AI this could give problems."

Manuel and Bhingwa raised important concerns that could be part and parcel of AI in cybersecurity and that could be influenced by the world viewpoint of the developers of these systems. Participant

Adrian Makwetla also alluded to this concern in his interview also adding concern about how AI learns:

> *"The information that we are using in the AI technology might be compromised or biased… I have concerns of integrity in terms of what the data produces and how the information was sourced."*

The analysis revealed that artificial intelligence in cybersecurity uses a variety of datasets to train and learn about the environment in which they operate. The training AI includes using machine learning, deep learning, natural language processing, network traffic data, log files, threat intelligence feeds, human in the loop interventions, common vulnerability and exposure (CVE) datasets, phishing data, open datasets, etc. all these combined form part of the algorithms used for anomaly detection. A study by Salem et al. (2024) supports this observation, arguing that the efficiency of AI in cybersecurity is highly dependent of the quality of data it ingests and poor data quality from the training datasets can lead to inaccuracies, they data quality can be affected by data privacy and compliance requirements.

The analysis found data quality can be susceptible to adversarial biases that could result from malicious activities on AI that would affect algorithms. Participant Johan De Bruyn expressed this view in his interview:

> *"They can get hacked, and your data can come up in other people's searches."*

De Bruyn raised an important point about the quality of data should there be malicious manipulation of data models by unauthorised parties, and this could impact the algorithms and the overall efficiency of AI in cybersecurity.

From the data analysis it has been discovered that the quality of AI training datasets is crucial for quickly and accurately detecting anomalies. It has been revealed that AI has its own inherent biases that can affect the outcome of detections, biases that can also enumerate from what has been learnt from the datasets, and these biases have a potential to erroneously detect activities as malicious or exempt malicious activities.

4.4.2 Scalability

Data analysis revealed that AI in cybersecurity scalability is crucial to ensure that it grows and adapts to the organisation's changing dynamics, i.e. User base, infrastructure, evolving threats etc. Participant Alex Bhingwa stated that the scalability of AI cybersecurity is crucial for its success

> *"As the organisation we are growing, we have to ensure that AI is scalable and not static or should there be a need for it to keep the pace, 'does it have capacity'? This should be designed with scalability and flexibility in mind. We need to harness and harvest the strength of our current infrastructure if there is any strength at all, then this can actually ensure the smooth incorporation of into our ideal operations. So we need a robust integration strategy and architecture, and for me this is essential for driving efficiency and effectiveness."*

Bhingwa emphasised that scalability ensures that the AI cybersecurity has flexibility to adjust, and that protection mechanism will not be compromised by the changes in the organisation's size, processes and adoption of new technologies, he further stressed that scalable AI for cybersecurity must be interoperable with existing and future infrastructure. The importance of AI cybersecurity interoperability was also emphasised by participant Jacob Motaung:

> *"They have to be interoperable with existing infrastructure, there must be some kind of an integration of some sort."*

Motaung highlighted the necessity for AI cybersecurity to seamlessly work together with other systems even though they are not operating on the same platform or developed by the same vendor. Participant Rebamang Lekoelea also highlighted the importance of infrastructure as the backbone of interventions.

> *"Organisational challenge could be infrastructure; do we have the required infrastructure to implement such tools?"*

The analysis found that AI cybersecurity can be interoperable to a certain extent, with obsolete systems this could raise a need to upgrade them to the latest technology. Participant Joice Mphahlele indicated that there have been a few cases of non-compatibility

"In the environment, I think I have picked up that there were issues related to lack of compatibility or integration related issues"

Participant Alfred Manyika also agreed that legacy systems are not interoperable with new technologies

"It has been very challenging given that you the existing systems using technology that is legacy technology, that is probably no longer supported by certain vendors and has reached end of sale and the end of life. It has been very challenging and now we've got new technologies that are faster and better and more accurate, So it has led us now to start looking into a holistic road map for which we can terminate all the legacy systems that are within the organisation and replace them with the new systems and newer and faster systems."

Data analysis revealed the importance of AI cybersecurity flexibility and scalability with the changes in the organisation. It has also revealed that it has to be operable with existing systems however also be cognizant of the fact that legacy systems need to be replaced with new systems as they cause compatibility issues.

4.4.3 Automation

The data analysis found that artificial intelligence in cybersecurity automates the prevention, detection, and resolution of threats quicker than humans. Research by Sarker (2023) supports this observation, arguing that artificial intelligence surpasses humans when it comes to detecting and predicting patterns, and analysing big data at a high speed to detect anomalies. Participant Anele Moloto supports the fact that AI enhances the automation of cybersecurity functions.

"A tool we once used made it difficult to detect insider threats, it offered basic data loss prevention capabilities, and it did not offer automated incident response which can lead an organisation exposed to threats… But thanks to AI, which is able to do this for us, to detect any activity and provide security orchestration, and automated response"

Moloto highlighted that automation decreases the load of work for the cybersecurity operations team, giving them more time to focus on issues in the environment. Participant Jacob Motaung supports this statement by saying:

> *"AI actually helps a lot because it doesn't require one to always be manually checking and all that it becomes automated. So by doing so, once you use such tools, it actually alleviate all that hard work or having so many warm body resources to actually do manual work"*

Participant Alfred Manyika agrees with this.

> *"AI becomes a pivotal role player in terms of automation. So once we have managed to automate certain processes, it becomes easier for us to optimize the processes and make sure that we deliver our services quite fast…So the one of the biggest benefits that we have had with the AI part is that it has helped us in identifying areas of weakness where we can actually automate and optimize protecting the perimeter security of the organisation, making sure that it's solid."*

Participant Alex Bhingwa looked at automation from a staffing perspective, highlighting the benefit of automation in an environment with limited budget for cybersecurity staff. He stated that:

> *"AI can help us address the cybersecurity skills gap, by automating some processes… Like now we are saying that we only have one senior person that is dealing with cybersecurity and maybe we need more hands, but we don't have the budget. So, then we can then automate some processes"*

The data analysis revealed that AI in cybersecurity brings about automation that makes manual processes seamless and quick to complete compared to human capacity. The amount of data that is collected by cybersecurity systems needs quick analysis that is availed by AI. Through automation, the team that is responsible for cybersecurity gets time to deal with other pressing issues while allowing AI to work on its own. This makes it feasible for an organisation to hire fewer employees as some of the manual work is taken over by AI and automation.

4.4.4 Efficiency

Data analysis revealed that efficiency is crucial for the success of AI in cybersecurity implementation. Participant Rebamang Lekoelea raised this view, further highlighting that efficiency can mean different things to different people.

> *"What I mean by efficiency is that it is not one is not one-size-fits-all, some people may view it as there aren't any improvements and changes that we see in terms of the organisation and this can be different for the next person"*

Data analysis revealed that the efficiency of AI cybersecurity is how well the solution achieves the intended purpose of mitigating threats, with minimal false positives or negatives. Frequent false positives or negatives have a direct impact on the reliability of the solution as it would indicate that some threats would go unnoticed and later affect the organisation. This was raised by participant Pule Molefe

> *"I do have the concerns because AI is still developing. I think it also needs to be improved. In our research, for instance we found that AI can bring about false negatives or false positives. It needs training and then the data that was used to train this particular technologies might still need to be to be improved"*

Molefe highlighted that since AI is still evolving there are many factors that could contribute to inaccuracies, these include the algorithms, quality of training data, misconfigured rules, network infrastructure changes, adversary attacks to the AI models etc. These false detections can raise incorrect alarms and have an impact on the operations of the organisation. Participant Akani Mathebula also alluded that false positives can impact efficiency

> *"The shortcomings will be sometimes missing the information employees require because of false blacklisting of an e-mail for example."*

Participant John Mokwena argued that these can potentially have an impact on business operations.

"It can increase turnaround time for that specific client, however. Depending on how quickly it was reported and also flagged on our side and then we take care of it and it through the help of ICT, then it gets solved."

The analysed data revealed efficiency can mean different things to different people, but what is common is that there should be fewer disruptions to business operations due to inaccuracies. The inaccuracies that are a result of frequent false positives or negatives have been identified to have a negative impact on the efficiency of AI cybersecurity. The

4.5 SOCIO-TECHNICAL INTERPLAY

The data analysis found that there is a reciprocal relationship between social and technical factors. This shows that these factors are interconnected and that they influence each other in various ways. The data analysis revealed complex and multi-dimensional interplay between change acceptance, top management support, cybersecurity awareness, ethical considerations, human oversight, usability, data quality, scalability, automation and efficiency. The AI cybersecurity is intertwined, there is an interplay between various components with some interplays being social factors only and some between technical factors only. However, this focuses on four key socio-technical interplays namely human oversight and automation, top management support and scalability, usability and efficiency, data quality and ethical considerations.

4.5.1 Human Oversight and Automation

The analysis of data found that, while artificial intelligence-driven automation provides convenience and surpasses humans in some cases, it does not replace humans. Further revealing that there are sophisticated cases of cybersecurity that cannot be properly handled by AI and require human intervention, this makes AI and humans co-exist and complement each other. Participant Alex Bhingwa mentioned the interplay between AI and human intervention in his interview.

"I think one of the things that I still grappling with is how do we find an optimal balance in human and AI interaction, striking the perfect balance between automation and human intervention."

Bhingwa's observation aligns with the observations made in the previous sections that human oversight and AI in cybersecurity automation complement each other and work together in pursuit of achieving a secure environment for the organisation.

The interlink between the automation and human oversight has been evident in from the data analysis. This shows the interrelationship between humans and technology, and it conforms that the two complement each other. With AI automating the repetitive tasks of humans and humans providing oversight to AI.

4.5.2 Top management support and scalability

The data revealed that in securing the organisation, the top management needs to sponsor the procurement of the necessary technology and tools which contribute to the efficiency of cybersecurity. Participant Alex Bhingwa commented on the interplay between top management support, investment on AI cybersecurity and its efficiency

> *"I am very positive and convinced that there is a return on investment, and we shall continue to invest in this area to make sure that our systems are really watertight."*

The data analysis found that technology is the backbone of AI in cybersecurity and investing in technology enhances the capability and scalability of AI cybersecurity. And that the top management have the responsibility to fund these initiatives and invest in technology that is interoperable with the latest solutions. The analysis revealed that the is expectation of returns on investment in AI cybersecurity initiatives.

4.5.3 Usability and Efficiency

The analysis found that usability can be view in different way between the technical and not technical staff, however, usable AI cybersecurity provides comfort and ease of use to both parties. The analysed data revealed that technical efficiency is twofold. Highly efficient and advanced AI can produce outputs that can be overwhelming to users, and this can have a negative effect of the usability of the solution as cause frustration that could lead to resistance, in the same breath a more relaxed solution can allow threats to pass undetected therefore affecting its perceived efficiency and performance.

While artificial intelligence offers advanced automation and analytical capabilities, usability is essential to closing the gap between technology efficiency and social effectiveness in an organisational setup. Negative user experience can lead to resistance.

4.5.4 Data Quality & Ethical Considerations

The analysis found that AI cybersecurity can have inherent biases that are influenced by the developers and training datasets, these biases that be discriminatory in nature and be unethical. AI has to comply to regulations and some regulations can impact data ingestion which may lead to inconsistency in data that would affect the algorithms and the outcomes.

4.6 CHAPTER SUMMARY

This chapter presented and analysed the findings from the data collected through interviews. An overarching theme, themes, sub-themes and codes emerged from the thematic data analysis, and they were presented on the thematic map. These teams indicate the social and technical critical success factors of AI in cybersecurity and their associated interplay. The findings will be discussed into detail in the next chapter.

5 CHAPTER 5. DISCUSSION OF THE FINDINGS

5.1 INTRODUCTION

The findings will be discussed and explained in this chapter, taking into consideration the literature contained in the literature review chapter. Thus, this chapter links back to the literature review and the Socio-Technical Critical Success Factors (STCSF) Model from figure 2, with the additional findings. The findings will be discussed and will highlight how these findings align with or differ from existing literature. The discussion will begin with the social factors, followed by technical factors and lastly the socio-technical interplay.

5.2 SOCIAL FACTORS

The thematic map led to the identification of the social factors that are critical to the success of AI cybersecurity implementation. Social factors can be described as aspects that are not technological and mainly have to do with but are not limited to the people, organisation, environment and organisational processes. Six social factors or subthemes emerged during data analysis.

5.2.1 Change Acceptance

Change acceptance refers to the ability and willingness of organisations, teams or individuals to adapt and embrace the implementation of AI cybersecurity into their workspaces and workflows. It is an important element of change management, in that it ensures smooth, effective and rewarding integration of AI-driven cybersecurity (Lapointe & Rivard, 2007). The analysis revealed that the acceptance of change is critical to the successful implementation of AI cybersecurity. Malatji et al. (2019) argue that implementing a technical change, such as AI cybersecurity, requires careful consideration of multiple socio-technical dimensions due to their interconnected nature. Effective communication and strategic change management are therefore essential to ensure that change is handled in such a manner that would improve the cybersecurity posture. The analysis found that timely, concise and unambiguous communication plays a pivotal role in making everyone aware of what to expect and what their responsibilities are. Chaudhary (2024) came to a similar conclusion; he argued that knowing what information needs to be communicated and how the communication should be packaged is crucial because people interpret information in different ways and end up coming to different conclusions despite having access to the same communication. He further suggests that a

message should be packaged in such a way that it will persuade or convince the reader to perform certain actions.

The use of change champions was highlighted as a key enabler of successful AI cybersecurity adoption. Participants observed that employees are more receptive to change when it is introduced by peers rather than the IT department. Koch & Fortkord (2024) support this view, arguing that change champions facilitate smoother transitions by fostering peer-to-peer learning and engagement within business units.

The analysis found that introducing change is prone to resistance challenges and that resistance can lead to ineffective cybersecurity. In their study Merhi & Ahluwalia (2019) came to the same conclusion arguing that resistance and can lead to cybersecurity breaches. They further recommend following up on the resistance to find the root cause to improve acceptance rate. Cram et al. (2024) argue that resistance can be sometimes due how the individual feels at a given time, and then further advocate for the implementation to consider contextual and individual differences in user acceptance.

Communication has been found to be the cornerstone of change; it is essential that the communication to be relevant at a given point in time and be as concise as possible. We have noted that beyond clear communication, the social influence by the change champions plays an important in AI cybersecurity success. It has been found that change is prone to resistance, and this need to be addressed from the core and alternative strategies to achieve change may need to be introduced.

5.2.2 Top Management Support

Top management support can be described as the assent, advocacy and involvement of senior leadership in AI cybersecurity initiatives (Rajapakse & Thushara, 2023). The study found that the top management at Edu X has the strategic responsibility to ensure that the organisation is secured, the operations part of this is delegated to lower levels. However, their visibility of the operations is helpful in decision-making. This finding aligns with the observation made by Kosutic & Pigni (2022) who emphasised that the strategic involvement of senior leadership in cybersecurity is not just compliance but likened its importance to a competitive advantage. Gale et al. (2022) came to similar conclusions arguing that the top management needs visibility in the form of cybersecurity reports on the state of cybersecurity in the organisation. This visibility helps top management to understand what is

happening at the operational level and puts them in a better position to make data-driven decisions on the allocation of enough strategic direction and budget for cybersecurity spending (Berg, 2024; Li et al., 2021).

This study found that budget allocation is a key determinant of AI cybersecurity success, with participants emphasising that securing financial investment for cybersecurity tools, skilled personnel, and system maintenance is a significant challenge. Several participants noted that while top management acknowledged the importance of AI cybersecurity, budget constraints often delayed or limited investment in necessary resources. This finding partially aligns with Li, Yoo, and Kettinger (2021), who argue that organisations that underfund cybersecurity are at a higher risk of security breaches due to outdated security measures. However, unlike Li et al.'s study, which primarily focuses on corporate prioritisation of cybersecurity spending, our findings suggest that budget allocation is not solely a matter of prioritisation but is often constrained by broader financial pressures within the organisation. Additionally, while Merhi (2023) asserts that top management must take a firm stance on cybersecurity investments to ensure strategic alignment, our findings indicate that some executives lack the technical knowledge to make informed cybersecurity investment decisions, leading to delays in acquiring AI-driven security solutions. This highlights a gap between executive-level decision-making and cybersecurity operational needs, suggesting that financial prioritisation alone is insufficient without technical awareness among decision-makers.

The data analysis revealed that top management plays a crucial role in instilling a cybersecurity culture. Uchendu et al. (2021) made the same observation, highlighting the pivotal role of top management in building a cybersecurity culture, ensuring that cybersecurity is the responsibility of not only the ICT team but everyone in the organisation is involved. While the role of top management is recognized, Wessel et al. (2021) differ from this view and argue that the over-reliance on top management has the potential of rendering cybersecurity culture a tick-box exercise, where there is no genuine involvement and commitment from the users who would only participate for the sake of compliance.

This section revealed that a good cybersecurity posture is a competitive advantage that the top management aims to achieve. The visibility of operations is crucial for attaining this competitive advantage, providing strategic direction and making data-driven decisions. Data is an enabler of

crucial decisions related to budget allocation and investment in cybersecurity, it also reveals areas that need top management intervention. However, overreliance on top management can potentially negatively impact cybersecurity.

5.2.3 Cybersecurity Awareness

The goal of cybersecurity awareness is to influence employees' behaviour and attitudes towards AI in cybersecurity Chaudhary (2024). The study found that an informed user base is essential in ensuring that there is a human line of defence, as an additional layer of security. This aligns with the study by Zwilling et al. (2022) who argued that while artificial intelligence offers advanced automation and analytical capabilities, cybersecurity awareness is essential to closing the gap between technology efficiency and the human counterpart. Li et al. (2021) argue that cybersecurity awareness arises due to end-users being the weakest link to cybersecurity threats. Contrary to this notion Cram et al. (2024) argue that the end-users cannot be singled out as the weakest link because the technology is also supposed to play its part, and it can also arguably be the weakest link leading to end-user vulnerability. Taherdoost (2022) argues that with the advent of AI cybersecurity which actively detects and prevents anomalies the claim that end users are the weakest link is less relevant.

The analysis revealed that cybersecurity awareness gives users an understanding of what cybersecurity is about and how they can play a role in protecting themselves and their organisation. Khando et al. (2021) made similar observations highlighting that cybersecurity awareness helps in fortifying cybersecurity in situations where the technology might not be able to detect an anomaly the cyber-aware users would be able to tell when there is a threat. Taddeo et al. (2023) allude to this and add that, this teaming up of AI and end-user helps foster conducive human-computer interaction leading to inclusive AI-cybersecurity.

Cybersecurity awareness adds a social layer to the defence of cybersecurity, it makes the employees know of the behaviour to avoid and the consequences involved. The awareness bridges the human-computer knowledge for AI cybersecurity and plays an important role in the implementation success. Cyber aware users understand that they have to be vigilant, protect themselves and their organisation.

5.2.4 Human Oversight

The study found that even though AI can learn as time goes by, there are certain cases that will need human decision-making. For example, AI can make mistakes, and it should not have autonomous permission to make decisions that will prevent human rectification. Artificial intelligence is continuously learning and prone to making mistakes at times, the human intelligence then comes into play when they critical thinking or certain decisions that need to be taken. Díaz-Rodríguez et al. (2023) made the same observation arguing that artificial intelligence does not replace human intelligence and that they complement each other and work in collaboration.

The study found that AI cybersecurity sometimes makes errors of wrongly classifying and blocking legitimate communication mistaking it as an anomaly. Quick human intervention in these and similar situations is critical to ensure that the business processes are not affected. In what Sarker et al. (2024) termed Human-AI teaming, this finding aligns with their observation that organisations should leverage and balance both the expertise of AI and the cybersecurity professionals to enhance their security posture through AI training and intervention.

AI and humans do not operate in isolation, their actions and presence complement each other. The human oversight has been found to be critical in situations whereby AI would need to make decisions based on human thinking, or where a human is needed to rectify mistakes committed by AI. The intervention by humans helps AI to learn patterns of resolving future cases.

5.2.5 Ethical Considerations

This study found that the incorporation of AI into cybersecurity raises concerns in relation to privacy, and surveillance which can have an impact on the trust and adoption of AI cybersecurity. This finding aligns with the studies by Bauer et al. (2023) & Galli et al. (2024) who argued that transparency and explainable AI (XAI) cybersecurity leads to an improved understanding by end-users which then influences them to trust the systems. In addressing the issue of trust, Díaz-Rodríguez et al. (2023) advocate for the incorporation of ethics in the development and implementation of AI cybersecurity, they argue that this would lead to trustworthy solutions that are aligned with social values and can be easily adopted.

This study found that it is important to adhere to ethical conduct, laws and regulatory requirements.

This aligns with the study by Moabalobelo et al. (2023) who assert that in the context of South Africa organisations that collect, and process data have to comply with the Protection of Personal Information Act by ensuring that the information is kept safe and to achieve data sovereignty it should be stored within the confines of the country. A study by Radanliev (2024) also emphasises the need to abide by the laws and regulations to ensure that data is protected.

The data analysis revealed in pursuit of compliance and regulating AI, it is important to update and create policies and frameworks to guide AI cybersecurity. Edu X saw a need to adopt new policies so that they would accommodate AI cybersecurity. This aligns with Bruschi & Diomede (2023) who posit that integrating policies that cover AI cybersecurity plays a big role in enabling organisations to ensure that technological improvements are aligned with social values and ethics.

The study found that stakeholders can have ethical concerns that could affect trust and the way they perceive AI in cybersecurity. They can feel that they are being watched and that their privacy is being violated. In pursuit of addressing ethical considerations, this study found the need for compliance with data laws and regulations.

5.2.6 Usability

In a socio-technical ecosystem, usability does not only refer to how a solution or system is technically designed, but it also further goes to how well it integrates with workflows, humans and their day-to-day activities. This aligns with Herrmann & Pfeiffer (2023) who posit that by following a socio-technical approach organisations can implement AI cybersecurity systems that not are not only user-friendly but also align with organisational goals.

The study found that usability would have different meanings to end-users and cybersecurity professionals. To the end-users, it would mean non-intrusive protection that seamlessly enhances their protection without affecting their workflows, and to the cybersecurity professionals, it would mean efficiency of interacting and tweaking the systems to perform the required action or produce required results (Alneyadi & Kassim, 2023; Guerino & Valentim, 2020).

The study revealed that AI cybersecurity should be usable enough not to cause frustration to those on the receiving end. This aligns with the study by Baruwal et al. (2024) who made the same observation

and argued that alert fatigue affects the usability of AI cybersecurity by delaying processes and workflows. They further argued that in detecting and preventing anomalies it is crucial that AI cybersecurity does not counter its mandate of preventing hiccups that would affect business operations, by rather causing more delays related to threat detection fatigue.

Usability has been found to be to be aligned with the ease of use, and integration with the day-to-day activities of the human stakeholders. The study found that AI cybersecurity that is disruptive of the day-day activities can be deemed not user-friendly as it may cause frustration to the users.

5.3 TECHNICAL FACTORS

The second theme that was found in the data analysis is technical factors that are critical for AI cybersecurity implementation. Technical factors can be described as factors that have to do with technology and its underlying software, systems and hardware. Four technical factors or subthemes emerged during data analysis.

5.3.1 Data Quality

The study found that AI cybersecurity has inherent biases that are based on the views of their developers and vendors, these biases can sometimes be discriminatory in nature which leads to inaccuracy in the classification of threats. This is aligned with the study by Kaushik et al. (2024) which found that the inherent biases can lead to unexpected responses and impact the reliability of AI.

The study revealed that the Human-AI teaming allows cybersecurity professionals an opportunity to validate AI insights and improve algorithmic performance. Aminu et al. (2024) made the same observation and advocated for continuous improvement in data collection, consistency, relevance, and that the AI models must be updated with the latest data and threat intelligence to detect emerging threats and reduce bias.

The study found that data integrity is important, data quality and quantity are critical aspects of AI-cybersecurity, and it has a direct influence on the efficiency of AI-Cybersecurity, machine learning models and algorithms, anomaly detection and automation, it can also have an indirect impact on the

acceptance and adoption of the system. This aligns with the study by Jabbarova (2023) which emphasised that AI algorithms are heavily dependent on high-quality data for training and decision-making processes, and that poor quality leads to poor results. Erhan et al. (2021) alluded to this by stating that high-quality data improves the accuracy of machine learning models, improves anomaly detection and prevention, and reduces the number of false positives. Merhi (2023) argues that data quantity is as important as data quality, therefore it is essential to ensure that the AI-cybersecurity systems have access to enough data quantity for better accuracy.

Data is collected from a variety of entities and data sets that AI has accessibility on, and AI can be trained by cybersecurity professionals through using the deep learning techniques to enhance the quality from various quantities of data. The quantity of data is as important as the quality because a large number of datasets increases the probabilities.

5.3.2 Scalability

The study found that AI Cybersecurity scalability is critical for ensuring that it has the flexibility to adapt to the technology, organisational, threats landscape changes. This aligns with the observation by Shamsan Saleh ((2024) who emphasised AI flexibility to the ever evolving technology landscape. In the study, AI cybersecurity was introduced in an environment that already has underlying infrastructure and existing systems. Interoperability with the existing infrastructure and systems is crucial for seamless integration, it is therefore important for an organisation to leverage what it already has and the systems to work in conjunction with each other.

It is important to note that interoperability in this study includes systems that are currently supported by their original equipment manufacturer (OEM), and does not necessarily include legacy and obsolete systems that are a security risk on their own. Okdem & Okdem (2024) argued that interoperability creates an effective and conducive cybersecurity ecosystem. The study found that as organisations evolve with time, the probability of growth or decline in the number of personnel is imminent, it is therefore crucial for AI cybersecurity to be scalable enough to accommodate such changes. By addressing scalability organisations can form a solid cybersecurity posture that can withstand changes and be able to integrate different systems,

5.3.3 Automation

The study found that automation of tasks is an integral part of AI cybersecurity, this is due to its ability to process data faster than humans and being able to be trained on how to tackle situations. Sarker (2024) and Naseer et al. (2024) are of a similar view and they argue that traditional cybersecurity systems are labour-intensive and that AI cybersecurity reduces the fatigue of doing manual work repeatedly. Automation requires little to no human intervention and significantly reduces the workload of the cybersecurity staff, allowing them time to deal with other pressing issues. Gafni & Levy (2024) highlight that AI can enhance cybersecurity by fully automation technical cybersecurity tasks and the automation of management cybersecurity tasks is not yet fully fledged. They further caution that human in loop should be considered to address potential errors. The study also found that from a cost perspective, automation is cost-effective as it significantly reduces the number of warm bodies to execute manual tasks

5.3.4 Efficiency

The study found that the identification and resolution of false positives give an opportunity to train the AI on how to handle future similar occurrences and improve the system's efficiency. This aligns with Aminu et al. (2024) who argue that algorithms can be refined from the misclassified benign incidents. However, the rise in false positives the negatively impacts usability of AI cybersecurity Erhan et al. (2021). The study revealed that efficiency is the goal of implementing AI cybersecurity, efficiency means the competency and ability to do what is intended timely and accurately. To have effective AI cybersecurity organisations must choose the right solutions. Organisations tend to choose solutions that are less expensive regardless of vendors, Merhi (2023) argues that it is important for organisations to select vendors based on various factors including the alignment with their core business rather than based only on price. Striking a balance between the solution costs and what it has to offer is necessary to ensure adequate security without excessively exceeding the budget. Measuring the efficiency of cybersecurity is not a straightforward exercise as there are various metrics used; however, its inefficiency can be glaring.

This section highlights that the efficiency of AI cybersecurity can be affected by a large number of false positives and negatives, this has the potential of wrongly classifying threats thereby delaying business processes and also a potential of allowing through anomalies that may go undetected and

later cause trouble. There are various issues that can contribute to the false positives and negatives. Highly sensitive AI cybersecurity can raise false alarms, while too-relaxed solutions can be dormant.

5.4 SOCIO-TECHNICAL INTERPLAY

5.4.1 Human Oversight and Automation

The study found that there is a multidimensional relationship between the technological systems (AI cybersecurity solutions and tools) and the social systems (Humans, the organisation, the environment, processes and workflows). The interplay is essential in implementing AI cybersecurity solutions that are both technically robust and socially aligned. Implementing AI cybersecurity necessitates introducing change in technology, this change interrelates with acceptance from end-users, it also has the potential to change workflows and at times top management intervention is required to fulfil the change. This aligns with the studies by Brown and Anwar (2021), Gafni & Levy (2024), Nguyen et al. (2024) who posit the AI automates the manual work of cybersecurity personnel allowing them for focus on other issues, and that in return humans provide critical thinking and decision-making to AI when it makes errors.

5.4.2 Top Management Support and Scalability

The study found that top management support has an interplay with the investments in the systems, scalability and funding AI cybersecurity. This aligns with extant literature, Kosutic & Pigni (2022) made the same observation arguing that top management support galvanises through the investments into AI cybersecurity initiatives. The findings by Mtsweni & Thaba (2022) and Pieterse (2021) indicate that the lack of relation between top management and scalability has led to successful cyber-attacks against South African SOEs, this lack of relationship is evident in resource constraints for cybersecurity due to underfunding. Kaur et al. (2023) argues that as technology advances the need for interoperability, scalability, and funding through top management also advances. Kosutic & Pigni (2022) views cybersecurity as a competitive advantage and emphasises the importance of the top management support and scalability interplay to ensure that the organisation is protected

5.4.3 Usability and Efficiency

The study found that that there is a multidimensional intricate relationship between AI cybersecurity efficiency and usability. It has been found that if users resist AI cybersecurity there can be vulnerabilities that would be allowed through thus affecting the efficiency of the solution, in the same breath a solution that is less efficient i.e. has frequent false positives has the high likelihood of being resisted. This aligns with the study by Merhi & Ahluwalia (2019) who argued negative user experience (UX) can lead to change resistance this can negatively impact cybersecurity and open up the organisation to threats that could have been avoided with user-friendly AI cybersecurity. While usability is important Ozkan-Okay et al. (2024) and Erhan et al. (2021) emphasise the importance of AI cybersecurity efficiency, arguing that the algorithmic performance of AI cybersecurity in handling cybersecurity should be evident in its low rate of false positives, detecting of merging threats and shorter mean time to resolve.

The usability and efficiency of interplay highlights an important interrelationship that needs to be nurtured, a break in relationship could spell cybersecurity challenges that would render the solution inefficient and negatively affect the success of AI cybersecurity. This aligns with the

5.4.4 Data Quality & Ethical Considerations

The data analysis found that AI learns from the data it collects and processes from various datasets. The quality of this data determines the performance of AI cybersecurity, however data processing should conform to the regulates in South Africa. This aligns with the studies by Moabalobelo et al. (2023) and Radanliev (2024) which found that the organisation operating in South Africa have to conform to the data regulations and preserve data sovereignty by not allowing the data to be stored outside the country. These regulations are stringent in how certain types of data are processed, more especially personal identifiable data, at times there would arise a need to anonymise data by removing personally identifiable information (PII), and this can affect the data quantity due to regulations limitation and consequentially degrade data accuracy and usability (Moabalobelo et al., 2023).

This study has found that AI has inherent biases that be due to the ignorance of AI developers and this can be at times discriminatory in nature, which is unethical. These biases affect the data quality which shows also impact the output in terms of the performance. This aligns with Kaushik et al. (2024) and Díaz-Rodríguez et al. (2023) who found that AI at times does not generalise well across

different populations, they further advocate for an ethical framework to address these challenges. Witanto et al. (2022) argued that data integrity issues in AI are a cause of ethical concern.

The interplay between data quality and ethical consideration has identified as an important interplay the has a potential to affect the core functionality of AI cybersecurity from within the data it processes. Its has been found that it is crucial to adhere to ethical regulation when implementing AI cybersecurity, the compliance to ethical considerations has also been noted to have the potential to degrade the quality of data.

5.5 CHAPTER SUMMARY

This chapter discussed the findings in detail and linked these findings back to the literature to check alignment or misalignment with the extant literature. The findings discussed were the social and technical critical success factors of AI in cybersecurity together with the socio-technical interplay between these factors. Chapter 6 provides a conclusion and recommendations for practice as well as looking into future research opportunities.

6 CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

6.1 INTRODUCTION

This chapter summarises the findings and provides future studies prospects. Provides a summary of the contributions to theory, methodology and practice. It highlights the limitations of the study, makes recommendations and derives conclusions from the case study with regard to the socio-technical critical success factors of implementing AI in cybersecurity.

6.2 SUMMARY OF FINDINGS

Implementing AI cybersecurity following a socio-technical approach has proven to lead to technically robust and socially aligned cybersecurity (Malatji et al., 2019; Van Haastrecht et al., 2021). Previous research has revealed that IS professionals tend to either focus on technical factor or social factors when implementing IS projects, this imbalance has led to systems that are not easily adopted as they do not seamlessly into the organisations' cybersecurity ecosystem that comprises technological factors, humans, the operation environment, the organisation and its workflows and processes (Baxter & Sommerville, 2011; Malatji et al., 2020; S. Sarker et al., 2019).

6.2.1 Objectives of The Study Reaffirmed

The aim of this study was to determine the socio-technical critical success factors of implementing AI for cybersecurity in an organisation. Numerous authors emphasise the importance of balancing social and technical aspects in implementing technology solutions (Kayworth & Whitten, 2010; S. Sarker et al., 2019). Edu X operates in a socio-technical ecosystem that is influenced by social, technical, organisational and environmental factors, it is therefore essential that all these factors are taken into consideration when implementing AI cybersecurity (Hoong & Rezania, 2024). To determine the critical success of AI cybersecurity implementation it was necessary to examine the whole process of implementation from planning to post-implementation monitoring and improvement. The study involved 17 semi-structured interviews that were conducted and transcribed on Microsoft Teams. The thematic analysis of the collected data led to the development of the thematic map that depicts the findings. Ten critical success factors were identified during data analysis. Due to the interconnectedness of the socio-technical ecosystems that Edu X operates in, the

data analysis also revealed that there is an underlying socio-technical interplay from these success factors.

The qualitative data was gathered and analysed with the aim of drawing meaning and extraction codes from the interviews and combining them to make up subthemes that feed into themes which are branches of the overarching theme. A thematic map was developed to illustrate the findings from this study. The thematic map identifies one overarching theme, three themes and ten subthemes. The thematic map provides insight into the critical success factors of AI for cybersecurity from a socio-technical perspective by the end-users and the ICT team and it also provides an insight into the interplay between social and technical success factors.

This study answered the research question together with the sub questions, for ease of reference the questions are reflected below.

The research question was:

What are the socio-technical critical success factors in the implementation of artificial intelligence in cybersecurity.

The sub-questions were:

i. What are the social and technical challenges encountered while implementing AI cybersecurity in state-owned entities?
ii. What critical factors contribute to the successful implementation of AI-based cybersecurity in South African state-owned entities?
iii. What lessons can be learnt from implementing AI cybersecurity in South African state-owned entities?

These are discussed in the following sections.

6.2.2 Social Factors

This section discusses the social factors that are critical for the implementation of AI in cybersecurity. In the context of this study, socio-technical factors are the non-technical factors that have to do with humans, organisation and the environment if AI cybersecurity implementation. The social factors are presented in Table 4.

Social Factor	**Key Findings**	**Future Research**
Change Acceptance	The study found that stakeholder engagements motivate change adoption. This aligns with Ayat et al. (2020) who emphasise that planning and clear communication ensures alignment between stakeholders.	The future could explore adopting the idiographic approach to AI cybersecurity change management.
Top Management Support	It has been found that top management support is essential in AI cybersecurity implementation. Uchendu et al. (2021) made the same observation and highlighted that due to the authority they have, top management play an important role in implementing cybersecurity culture change. Wessel et al. (2021) caution overly involving top management on cybersecurity culture implementation.	Investigate how culture change can be instilled without top management being overly involved.
Cybersecurity awareness	The study revealed that Cybersecurity awareness influences the employees' behaviour towards AI cybersecurity. Chaudhary (2024) came to the same conclusion and argues that cybersecurity awareness educates users on how to handle cybersecurity matters.	Future studies could explore implementation of AI driven cybersecurity awareness programs for personalised trainings that speak to individual users.

Human Oversight	It has been evident in the study that even though AI provides advanced automation of tasks, human oversight is needed. Díaz-Rodríguez et al. (2023) made the same observation arguing humans are needed to guide AI in novel matters.	Explore how human biases influence the performance of AI.
Ethical Considerations	The study found that AI has inherent biases, that affect trust. This aligns with Kaushik et al. (2024) who argue AI perpetuate biases that are existing in its training data.	Future studies could explore the mitigation of algorithmic bias.
Usability	It has been found that is a key component to AI cybersecurity implementation success. This aligns with Alneyadi & Kassim (2023) who argue that user friendly interfaces improve adoption.	Investigate existing AI cybersecurity solutions to identify usability issues and areas for improvement.

Table 4 Social Factors

6.2.3 Technical Factors

The thematic map also shown technical factors that are critical to the successful implementation of AI cybersecurity. These factors mainly have to do with the technology part of the implementation of AI cybersecurity. Based on the analysis of data four subthemes were discovered, these subthemes are presented in Table 5.

Technical Factor	**Key findings**	**Future Research**
Data Quality	The study found that data integrity is crucial for AI cybersecurity, because it is the determining factor of its	Explore ways to improve data consistency, trustworthiness of data.

	performance. Salem et al. (2024) and Erhan et al. (2021) are aligned with this observation arguing that	
Scalability	The study revealed that it is necessary for AI cybersecurity to be interoperable with existing technologies. Ramos-Cruz and Sarker (2023) came to similar conclusion emphasising the AI cybersecurity should be compatible with the existing systems and also be open to future prospects.	Explore the impact of emerging technologies on existing systems.
Automation	It has been evident from the study that AI brings revolution to cybersecurity by automatics cybersecurity tasks. This finding aligns with Gafni & Levy (2024) who posit that AI automates technical cybersecurity task affording the cybersecurity teams to focus on strategic matters.	Investigate the AI cybersecurity orchestration for fully automated Security Operations Centers (SOCs)
Efficiency	This study found that the efficiency of AI cybersecurity in predicting, detecting and prevention threats is essential for the everchanging threat landscape. Ramos-Cruz (2024) & Erhan et al. (2021)	Explore how adaptive algorithms influence AI cybersecurity efficiency.

	made a similar observation in their studies and emphasised that the algorithmic performance of AI cybersecurity in handling cybersecurity should be evident in its high capability of detecting of merging threats, prompt resolution and a low rate of false positives.	

Table 5 Technical Factors

6.2.4 The Socio-Technical Interplay

Lastly, the thematic map also showed that there is a multidimensional and dynamic interplay between the social and technical factors. This interplay is proof of the interconnectedness of the social and technical factors critical for AI cybersecurity. Based on the analysis of data four subthemes were discovered, these subthemes are presented in table 6.

Interplay	**Key Findings**	**Future Research**
Human Oversight and Automation	The study highlights the interplay between human oversight and AI-driven automation in cybersecurity, emphasising their complementary roles in ensuring robust decision-making. Prior research supports the need for explainable AI to enhance human-AI teaming in cybersecurity (Sarker et al., 2024).	Future research could be on further exploration of the explainable AI for human oversight and automation with the aim of advancing human-AI teaming.
Top management support and scalability	This study found a multidimensional relationship top managers support and scalability. Mtsweni & Thaba (2022) and Pieterse (2021) emphasised the importance	Future research could be on exploring AI data driven decision support systems to aid top

	of this relationship, arguing that lack of top management support and scalability interplay enabled cyberattacks due to underfunding cybersecurity (Mtsweni & Thaba, 2022; Pieterse, 2021).	managers in making informed decisions.
Usability and Efficiency	The usability and efficiency interplay has been evident in this study. Merhi & Ahluwalia and found the relationship between usability and efficiency to be intricate, they argued that inefficiencies like frequent false positives impact usability and perceived efficiency, and that efficient algorithm performance improves user experience and usability.	The future research could on AI for user experience exploring ways to improve the user experience (UX).
Data Quality & Ethical Considerations	This study highlights the interplay between data quality and ethical considerations. Previous studies posit that data integrity is a cause for ethical concern and that advocate for AI data regulation (Díaz-Rodríguez et al., 2023; Witanto et al., 2022)	Future research could be on exploring data integrity through trustworthy AI.

Table 6 Socio-technical interplay

6.3 SUMMARY OF CONTRIBUTIONS

The main purpose of this study was to investigate the socio-technical critical success factors of implementing AI in cybersecurity. This study was motivated by the factor that AI in cybersecurity has emerged as one of the solutions to the ever-evolving cyber landscape, with scholars writing about its impact, and importance (Guembe et al., 2022; Jada & Mayayise, 2023; I. H. Sarker, 2023). However, in the extant literature, there has been a scarcity of literature that addresses the socio-technical implementation of AI in cybersecurity from a socio-technical perspective. The focus on socio-technical implementation of AI in cybersecurity has been deemed important due to the success and failures of information systems implementation (Barker & Frolick, 2003; Chatterjee et al., 2024;

Jeyaraj, 2020; Rajapakse & Thushara, 2023). This study makes theoretical, academic and practitioner contributions.

6.3.1 Theoretical Contributions

The contribution of the study from a theoretical perspective, lies in expanding the body of knowledge on socio-technical cybersecurity implementation by empirically testing effectiveness and applicability of the existing theories in a real-world setup. This study closes the theory assumption and practical application gap by validating these theories in a practical organisational environment. The findings in this study provide an understanding which is backed by evidence, that contributes to best practices and helps both academics and practitioners refine cybersecurity strategies. Furthermore, the existing theories were not sufficient for this study, and this led to the development of a new framework through refining and combining the existing critical success factors and socio-technical systems theories. This study produced a called Socio-Technical Critical Success Factors theory. This theory can be used for future studies and practical IS implementation.

6.3.2 Methodological Contributions

The methodological contribution lies in gathering detailed qualitative data on socio-technical AI cybersecurity implementation through in-depth interviews with the executives, managers, ICT staff and general end-users from different business units within a case study research. The in-depth interviews with various stakeholders provide rich firsthand insights into the socio-technical critical success factors of AI in cybersecurity implementation, challenges and the socio-technical interplay. The triangulation of participants diversified the responses, by gathering data from different user groups to ensure the reliability and validity of the findings by uncovering inconsistencies and minimizing bias. Employing an interpretive case study and the qualitative methodology contributes to the advancement of research in cybersecurity and informs the academic discourse.

6.3.3 Practical Contributions

From a practical standpoint, the contribution lies in translating the lessons learnt from this research into interventions that can be applied to help organisations successfully implement AI-driven cybersecurity. By examining the data on real-world opportunities, challenges and critical success factors, this study provides evidence-based recommendations that can be further tailored for different

sectors, organisations and countries. These interventions will be discussed in the lessons learnt section chapter. The findings from this study fill the gap between theory and practice, and equip decision-makers as well as organisations with practical guidance and strategies on AI-driven cybersecurity implementation.

6.3.4 Lessons Learnt

This study uncovered challenges and lessons in relation to the socio-technical implementation of AI in cybersecurity. Change resistance has been identified to pose a threat to the success of AI in cybersecurity, the resistance to change can be due to various factors including uncertainty, trust and poor communication. This aligns with the findings by Merhi & Ahluwalia (2019) and Rajapakse & Thushara (2023) agree that implementing change disrupts the status quo and can lead to resistance, they posit that is that clear and concise communication is essential in stakeholder engagements. Bauer (2023) and Sarker et al. (2024) posit that trust and privacy concerns can make users resist AI implementation, and they advise that awareness campaigns can make users understand the role of AI cybersecurity in their workflows and make them aware of their responsibilities. Cram et al. (2024) argue that there is various reason that may lead to changes resistance. They further emphasise that it is important to treat resistance cases individually and they encourage using the idiographic approach to address this.

The study found that organisations operate with limited resources, and this can have an impact on the investment in cybersecurity. Lack of technical knowledge from the executives can potentially lead to cybersecurity not being one of the areas prioritised amid resource constraints, however the security of an organisation is one of its competitive advantages and the executives should take a firm stance on cybersecurity investments (Kosutic & Pigni, 2022; Merhi, 2023). Pieterse (2021) and Mtsweni & Thaba (2022) highlight that lack of cybersecurity funding in personnel, awareness, systems etc. has been one of the contributing factors to the recent cyberattacks in South Africa.

The study uncovered that while AI automates manual work and makes more quick analyses than humans, it is prone to making errors. Díaz-Rodríguez et al. (2023) , Kaushik et al. (2024) has inherent biases that can lead to it making errors, these errors can lead to false positives and false negatives which have a negative impact on the adoption, efficiency, usability and operation of AI in cybersecurity. Bécue et al (2021) and Baruwal (2024) advocate for a complementary Human-AI in

that will allow AI to automate at the human in the loop to use human judgement to rectify timeously errors, and AI eventually learns the patterns from the human interventions.

6.4 LIMITATIONS OF THE STUDY

This study had limitations that should be considered, in relation to the findings. The rapidly evolving nature of both AI and cybersecurity means that some of the specific technologies, factors and strategies discussed in this study may become outdated over time. The study was focused on a single organisation, which can arguably lead to doubting the generalizability of the findings of the single case study by citing limited external validity of the results, however Baskerville & Lee (1999) argue that a single case study can claim generalizability when it is conducted and presented properly. The sample was diverse enough to cover all groups of users, however, due to time limitations the interviews conducted were limited to seventeen interviews and this can be viewed as limited data collected. However, Yin (2017) and argues that there are no rules dictating the right sample size for qualitative research because unlike quantitative studies qualitative studies are more analytical than statistical. This study can be extended to include more participants or include multiple organisations for more diversity. The qualitative processes that were followed in this study have the potential of being biased as the researcher can apply his interpretation when deducting codes and themes. The critical success factors of AI cybersecurity implementation alongside the socio-technical interplays identified may not be the only ones that exist. Other factors can potentially be discovered with further research in the future.

6.5 CONCLUSION

This study examined the socio-technical critical success factors of implementing AI in cybersecurity by analysing the existing literature and conducting an analysis of the data collected through in-depth interviews with various stakeholders of Edu X. It evaluated the existing theories and produced a new Socio-Technical Critical Success Factors theory. The research found three factors that are critical for the successful implementation of AI in cybersecurity namely social factors, technical factors and the socio-technical interplay. This study contributes to theory, methodology and practice. provided answers to the research questions question.

Future research could include the investigation of the explainable artificial intelligence (XAI) in cybersecurity. This would include an exploration of AI, efficiency, trust and accountability, thus contributing to the socio-technical AI cybersecurity discourse.

7 REFERENCES

Aanestad, M., & Jensen, T. B. (2016). Collective mindfulness in post-implementation IS adaptation processes. *Information and Organization*, *26*(1), 13–27. https://doi.org/10.1016/j.infoandorg.2016.02.001

Ahmad, A., Desouza, K., Maynard, S., Naseer, H., & Baskerville, R. (2019). How integration of cyber security management and incident response enables organizational learning. *Journal of the Association for Information Science and Technology*, *71*. https://doi.org/10.1002/asi.24311

Alavizadeh, H., Jang-Jaccard, J., Enoch, S. Y., Al-Sahaf, H., Welch, I., Camtepe, S. A., & Kim, D. S. (2021). *A Survey on Threat Situation Awareness Systems: Framework, Techniques, and Insights* (No. arXiv:2110.15747). arXiv. https://doi.org/10.48550/arXiv.2110.15747

AL-Ghamdi, A. S. A.-M., Ragab, M., & Sabir, M. F. S. (2022, August 1). *Enhanced Artificial Intelligence-based Cybersecurity Intrusion Detection for Higher Education Institutions. | EBSCOhost*. https://doi.org/10.32604/cmc.2022.026405

Alneyadi, M., & Kassim, N. (2023). Factors Influencing User's Intention to Adopt AI-Based Cybersecurity Systems in the UAE. *Interdisciplinary Journal of Information, Knowledge, and Management*, *18*, 459–486. https://doi.org/10.28945/5166

Al-Suqri, M. N., & Gillani, M. (2022). A Comparative Analysis of Information and Artificial Intelligence Toward National Security. *IEEE Access*, *10*, 64420–64434. https://doi.org/10.1109/ACCESS.2022.3183642

Aminu, M., Akinsanya, A., Oyedokun, O., Dickson, A., & Dako. (2024). Enhancing Cyber Threat Detection through Real-time Threat Intelligence and Adaptive Defense Mechanisms. *International Journal of Computer Applications Technology and Research*, *13*, 11–27. https://doi.org/10.7753/IJCATR1308.1002

Apruzzese, G., Laskov, P., Oca, E. M. de, Mallouli, W., Rapa, L. B., Grammatopoulos, A. V., & Franco, F. D. (2023). The Role of Machine Learning in Cybersecurity. *Digital Threats: Research and Practice*, *4*(1), 1–38. https://doi.org/10.1145/3545574

Araya, D. (2022). *Cybersecurity and AI* (Artificial Intelligence for Defence and Security, pp. 13–14). Centre for International Governance Innovation. https://www.jstor.org/stable/resrep42557.11

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for Artificial Intelligence and Digital technologies. *International Journal of Information Management*, *62*, 102433. https://doi.org/10.1016/j.ijinfomgt.2021.102433

Ayat, M., Imran, M., Ullah, A., & Kang, C. W. (2020). Current trends analysis and prioritization of success factors: A systematic literature review of ICT projects. *International Journal of Managing Projects in Business*, *14*(3), 652–679. https://doi.org/10.1108/IJMPB-02-2020-0075

Barker, T., & Frolick, M. N. (2003). Erp Implementation Failure: A Case Study. *Information Systems Management*, *20*(4), 43–49. https://doi.org/10.1201/1078/43647.20.4.20030901/77292.7

Baruwal Chhetri, M., Tariq, S., Singh, R., Jalalvand, F., Paris, C., & Nepal, S. (2024). Towards Human-AI Teaming to Mitigate Alert Fatigue in Security Operations Centres. *ACM Trans. Internet Technol.*, *24*(3), 12:1-12:22. https://doi.org/10.1145/3670009

Baskerville, R., & Lee, A. S. (1999). Distinctions among Different Types of Generalizing in Information Systems Research. In O. Ngwenyama, L. D. Introna, M. D. Myers, & J. I. DeGross (Eds.), *New Information Technologies in Organizational Processes: Field Studies and Theoretical Reflections on the Future of Work* (pp. 49–65). Springer US. https://doi.org/10.1007/978-0-387-35566-5_5

Baskerville, R., & Myers, M. D. (2004). Special Issue on Action Research in Information Systems: Making IS Research Relevant to Practice: Foreword. *MIS Quarterly*, *28*(3), 329–335. https://doi.org/10.2307/25148642

Bauer, K., von Zahn, M., & Hinz, O. (2023). Expl(AI)ned: The Impact of Explainable Artificial Intelligence on Users' Information Processing. *Information Systems Research*, *34*(4), 1582–1602. https://doi.org/10.1287/isre.2023.1199

Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. *Interacting with Computers*, *23*(1), 4–17. https://doi.org/10.1016/j.intcom.2010.07.003

Bécue, A., Praça, I., & Gama, J. (2021). Artificial intelligence, cyber-threats and Industry 4.0: Challenges and opportunities. *Artif. Intell. Rev.*, *54*(5), 3849–3886. https://doi.org/10.1007/s10462-020-09942-2

Berg, B. van den. (2024). Dealing with uncertainty in cyberspace. *Computers & Security*, *144*, 103939. https://doi.org/10.1016/j.cose.2024.103939

Bharosa, N. (2022). The rise of GovTech: Trojan horse or blessing in disguise? A research agenda. *Government Information Quarterly*, *39*(3), 101692. https://doi.org/10.1016/j.giq.2022.101692

Bhattacherjee, A. (with Open Textbook Library). (2012). *Social Science Research: Principles, Methods, and Practices*. Global Text Project.

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

Brett Van Niekerk. (2017). An Analysis of Cyber-Incidents in South Africa. *The African Journal of Information and Communication (AJIC)*, *20*. https://doi.org/10.23962/10539/23573

Brown, J., & Anwar, M. (2021). Blacksite: Human-in-the-loop Artificial Immune System for Intrusion Detection in Internet of Things. *Human-Intelligent Systems Integration*, *3*. https://doi.org/10.1007/s42454-020-00017-9

Bruschi, D., & Diomede, N. (2023). A framework for assessing AI ethics with applications to cybersecurity. *AI and Ethics*, *3*(1), 65–72. https://doi.org/10.1007/s43681-022-00162-8

Camilleri, M. A., & Falzon, L. (2020). Understanding motivations to use online streaming services: Integrating the technology acceptance model (TAM) and the uses and gratifications theory (UGT). *Spanish Journal of Marketing - ESIC*, *25*(2), 217–238. https://doi.org/10.1108/SJME-04-2020-0074

Chatterjee, S., Chakraborty, S., Fulk, K., & Lowry, P. B. (2024). *The Role of Dissonant Relational Multiplexity in Information System Implementation Failures* (SSRN Scholarly Paper No. 4715716). Social Science Research Network. https://papers.ssrn.com/abstract=4715716

Chaudhary, S. (2024). Driving behaviour change with cybersecurity awareness. *Computers & Security*, *142*, 103858. https://doi.org/10.1016/j.cose.2024.103858

Chen, W., & Hirschheim, R. (2004). A paradigmatic and methodological examination of information systems research from 1991 to 2001. *Information Systems Journal*, *14*(3), 197–235. https://doi.org/10.1111/j.1365-2575.2004.00173.x

Chilisa, B. (2019). *Indigenous research methodologies*. Sage publications. https://scholar.google.com/scholar?cluster=16226324406274659958&hl=en&inst=5957094470577306745&inst=12570381107082219421&inst=56936736054743433 9&inst=7893438682994895401&oi=scholarr

Chorąś, M., & Woźniak, M. (2022). The double-edged sword of AI: Ethical Adversarial Attacks to counter artificial intelligence for crime. *AI and Ethics*, *2*(4), 631–634. https://doi.org/10.1007/s43681-021-00113-9

Chuang, M. (2024). Pressure to Adopt AI: Towards an Interdisciplinary Framework. *AMCIS 2024 TREOs*. https://aisel.aisnet.org/treos_amcis2024/90

Logicalis (2024). *CIO Report*. Retrieved February 24, 2025, from https://www.za.logicalis.com/cio-report

Cisco (2024). *Cybersecurity_Readiness_Index* Retrieved February 24, 2025, from https://newsroom.cisco.com/c/dam/r/newsroom/en/us/interactive/cybersecurity-readiness-index/documents/Cisco_Cybersecurity_Readiness_Index_FINAL.pdf

Cooper, V. A. (2008). *The Critical Success Factor Method: A review and practical example*.

Cram, W., Arcy, J., & Benlian, A. (2024). Time Will Tell: The Case for an Idiographic Approach to Behavioral Cybersecurity Research. *MIS Quarterly*, *48*, 95–136. https://doi.org/10.25300/MISQ/2023/17707

Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE.

Creswell, J. W., & Clark, V. L. P. (2011). *Designing and Conducting Mixed Methods Research*. SAGE.

Creswell, J. W., & Creswell, J. D. (2017). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications.

Dagada, R. (2024). The Advancement of 4IR Technologies and Increasing Cyberattacks in South Africa. *Southern African Journal of Security*, *2*, 27 pages-27 pages. https://doi.org/10.25159/3005-4222/15157

Delone, W., & McLean, E. (2003). The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *J. of Management Information Systems*, *19*, 9–30. https://doi.org/10.1080/07421222.2003.11045748

Díaz-Rodríguez, N., Del Ser, J., Coeckelbergh, M., López de Prado, M., Herrera-Viedma, E., & Herrera, F. (2023). Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. *Information Fusion*, *99*, 101896. https://doi.org/10.1016/j.inffus.2023.101896

Diesch, R., & Krcmar, H. (2020). SoK: Linking information security metrics to management success factors. *Proceedings of the 15th International Conference on Availability, Reliability and Security*, 1–10. https://doi.org/10.1145/3407023.3407059

Ebert, N., Schaltegger, T., Ambuehl, B., Schöni, L., Zimmermann, V., & Knieps, M. (2023). Learning from safety science: A way forward for studying cybersecurity incidents in organizations. *Computers & Security*, *134*, 103435. https://doi.org/10.1016/j.cose.2023.103435

Erhan, L., Ndubuaku, M., Di Mauro, M., Song, W., Chen, M., Fortino, G., Bagdasar, O., & Liotta, A. (2021). Smart anomaly detection in sensor systems: A multi-perspective review. *Information Fusion*, *67*, 64–79. https://doi.org/10.1016/j.inffus.2020.10.001

Fonseca, R. S., & van Wyk, J.-A. (2021). Cybersecurity in South Africa: Status, governance, and prospects. In *Routledge Companion to Global Cyber-Security Strategy* (pp. 591–607). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9780429399718-56/cybersecurity-south-africa-raymond-steenkamp-fonseca-jo-ansie-van-wyk

Fowler, J. J., & Horan, P. (2007). Are Information Systems' Success and Failure Factors Related? An Exploratory Study. *Journal of Organizational and End User Computing (JOEUC)*, *19*(2), 1–22. https://doi.org/10.4018/joeuc.2007040101

Freund, Y. P. (1988). Critical success factors. *Planning Review*, *16*(4), 20–23. https://doi.org/10.1108/eb054225

Fujs, D., Mihelič, A., & Vrhovec, S. L. R. (2019). The power of interpretation: Qualitative methods in cybersecurity research. *Proceedings of the 14th International Conference on Availability, Reliability and Security*, 1–10. https://doi.org/10.1145/3339252.3341479

Gafni, R., & Levy, Y. (2024). The role of artificial intelligence (AI) in improving technical and managerial cybersecurity tasks' efficiency. *Information & Computer Security*, *32*(5), 711–728. https://doi.org/10.1108/ICS-04-2024-0102

Gale, M., Bongiovanni, I., & Slapnicar, S. (2022). Governing cybersecurity from the boardroom: Challenges, drivers, and ways ahead. *Comput. Secur.*, *121*(C). https://doi.org/10.1016/j.cose.2022.102840

Galli, A., La Gatta, V., Moscato, V., Postiglione, M., & Sperlì, G. (2024). *Explainability in AI-based behavioral malware detection systems*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85189932116&doi=10.1016%2fj.cose.2024.103842&partnerID=40&md5=9014c6e268be81ae6750297603fef669

Goldkuhl, G. (2012). Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems*, *21*(2), 135–146. https://doi.org/10.1057/ejis.2011.54

Goodrick, D., & Rogers, P. J. (2015). Qualitative Data Analysis. In *Handbook of Practical Program Evaluation* (pp. 561–595). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119171386.ch22

Gordon, D., Stavrakakis, I., Gibson, J. P., Tierney, B., Becevel, A., Curley, A., Collins, M., O'Mahony, W., & O'Sullivan, D. (2021). Perspectives on computing ethics: A multi-stakeholder analysis. *Journal of Information, Communication and Ethics in Society*, *20*(1), 72–90. https://doi.org/10.1108/JICES-12-2020-0127

Grobler, M., Gaire, R., & Nepal, S. (2021). User, Usage and Usability: Redefining Human Centric Cyber Security. *Frontiers in Big Data*, *4*. https://doi.org/10.3389/fdata.2021.583723

Guembe, B., Azeta, A., Misra, S., Osamor, V. C., Fernandez-Sanz, L., & Pospelova, V. (2022). The Emerging Threat of Ai-driven Cyber Attacks: A Review. *Applied Artificial Intelligence*, *36*(1), 2037254. https://doi.org/10.1080/08839514.2022.2037254

Guerino, G. C., & Valentim, N. M. C. (2020). Usability and user experience evaluation of natural user interfaces: A systematic mapping study. *IET Software*, *14*(5), 451–467. https://doi.org/10.1049/iet-sen.2020.0051

Gumede, V., Bila, S., Biyase, M., Chauke, S., & Arogundade, S. (2024). Government and the Market: State-Owned Enterprises in South Africa. In V. Gumede, S. Bila, M. Biyase, S. Chauke, & S. Arogundade (Eds.), *South African Economy: Trails and Possibilities* (pp. 123–141). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-54180-3_5

Haag, S., Siponen, M., & Liu, F. (2021). Protection Motivation Theory in Information Systems Security Research: A Review of the Past and a Road Map for the Future. *SIGMIS Database*, *52*(2), 25–67. https://doi.org/10.1145/3462766.3462770

Happa, J., Agrafiotis, I., Helmhout, M., Bashford-Rogers, T., Goldsmith, M., & Creese, S. (2021). Assessing a Decision Support Tool for SOC Analysts. *Digital Threats*, *2*(3), 22:1-22:35. https://doi.org/10.1145/3430753

Hassan, N. R., Mingers, J., & Stahl, B. (2018). Philosophy and information systems: Where are we and where should we go? *European Journal of Information Systems*, *27*(3), 263–277. https://doi.org/10.1080/0960085X.2018.1470776

He, Y., Xin, T., & Luo, C. (2024). Cybersecurity Investments Metrics using FAIR-ROSI. *UK Academy for Information Systems Conference Proceedings 2024*. https://aisel.aisnet.org/ukais2024/6

Heeks, R. (2002). Information Systems and Developing Countries: Failure, Success, and Local Improvisations. *The Information Society*, *18*(2), 101–112. https://doi.org/10.1080/01972240290075039

Herrmann, T., & Pfeiffer, S. (2023). Keeping the organization in the loop: A socio-technical extension of human-centered artificial intelligence. *AI & SOCIETY*, *38*(4), 1523–1542. https://doi.org/10.1007/s00146-022-01391-5

Hoong, Y., & Rezania, D. (2024). Navigating Cybersecurity Governance: The influence of opportunity structures in socio-technical transitions for small and medium enterprises. *Computers & Security*, *142*, 103852. https://doi.org/10.1016/j.cose.2024.103852

Intelligence, M. T. (2024, February 14). *Staying ahead of threat actors in the age of AI*. Microsoft Security Blog. https://www.microsoft.com/en-us/security/blog/2024/02/14/staying-ahead-of-threat-actors-in-the-age-of-ai/

Jabbarova, K., & Jafarov, S. (2023). AI and Cybersecurity—New Threats and Opportunities. *Pakistan Journal of Life and Social Sciences (PJLSS)*, *22*. https://doi.org/10.57239/PJLSS-2024-22.2.00752

Jada, I., & Mayayise, T. O. (2023). The impact of artificial intelligence on organisational cyber security: An outcome of a systematic literature review. *Data and Information Management*, 100063. https://doi.org/10.1016/j.dim.2023.100063

Jeyaraj, A. (2020). DeLone & McLean models of information system success: Critical meta-review and research directions. *International Journal of Information Management*, *54*, 102139. https://doi.org/10.1016/j.ijinfomgt.2020.102139

Kaloudi, N., & Li, J. (2020). The AI-Based Cyber Threat Landscape: A Survey. *ACM Comput. Surv.*, *53*(1), 20:1-20:34. https://doi.org/10.1145/3372823

Kaur, R., Gabrijelčič, D., & Klobučar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion*, *97*, 101804. https://doi.org/10.1016/j.inffus.2023.101804

Kaushik, K., Khan, A., Kumari, A., Sharma, I., & Dubey, R. (2024). Ethical Considerations in AI-Based Cybersecurity. In K. Kaushik & I. Sharma (Eds.), *Next-Generation Cybersecurity: AI, ML, and Blockchain* (pp. 437–470). Springer Nature. https://doi.org/10.1007/978-981-97-1249-6_19

Kayworth, T., & Whitten, G. (2010). Effective Information Security Requires a Balance of Social and Technology Factors. *MIS Quarterly Executive*, *9*.

Khan, N. F., Ikram, N., Murtaza, H., & Javed, M. (2023). Evaluating protection motivation based cybersecurity awareness training on Kirkpatrick's Model. *Computers & Security*, *125*, 103049. https://doi.org/10.1016/j.cose.2022.103049

Khan, Z. C., & Mkuzangwe, N. (2022). Advancing cybersecurity capabilities for South African organisations through R&D. *International Conference on Cyber Warfare and Security*, *17*(1), Article 1. https://doi.org/10.34190/iccws.17.1.34

Khando, K., Gao, S., Islam, S. M., & Salman, A. (2021). Enhancing employees information security awareness in private and public organisations: A systematic literature review. *Computers & Security*, *106*, 102267. https://doi.org/10.1016/j.cose.2021.102267

Kim, H.-W., & Kankanhalli, A. (2009). Investigating User Resistance to Information Systems Implementation: A Status Quo Bias Perspective. *MIS Quarterly*, *33*, 567–582. https://doi.org/10.2307/20650309

Koch, T., & Fortkord, C. (2024). Opinion Leaders in Organizational Change Processes: Insights from a Large-Scale Quantitative Survey to Identify, Characterize, and Classify Influential

Employees within a Multinational Corporation. *Journal of Change Management*, *24*(4), 325–345. https://doi.org/10.1080/14697017.2024.2426789

Kosutic, D., & Pigni, F. (2022). Cybersecurity: Investing for competitive outcomes. *Journal of Business Strategy*, *43*(1), 28–36.

Lapointe, L., & Rivard, S. (2007). A Triple Take on Information System Implementation. *Organization Science*. https://doi.org/10.1287/orsc.1060.0225

Lauriola, I., Lavelli, A., & Aiolli, F. (2022). An introduction to Deep Learning in Natural Language Processing: Models, techniques, and tools. *Neurocomputing*, *470*, 443–456. https://doi.org/10.1016/j.neucom.2021.05.103

Lee, A. S. (1989). A Scientific Methodology for MIS Case Studies. *MIS Quarterly*, *13*(1), 33–50. https://doi.org/10.2307/248698

Lee, I. (2021). Cybersecurity: Risk management framework and investment cost analysis. *Business Horizons*, *64*(5), 659–671. https://doi.org/10.1016/j.bushor.2021.02.022

Li, E. Y. (1997). Perceived importance of information system success factors: A meta analysis of group differences. *Information & Management*, *32*(1), 15–28. https://doi.org/10.1016/S0378-7206(97)00005-0

Li, H., Yoo, S., & Kettinger, W. J. (2021). The Roles of IT Strategies and Security Investments in Reducing Organizational Security Breaches. *Journal of Management Information Systems*, *38*(1), 222–245. https://doi.org/10.1080/07421222.2021.1870390

Li, Y., & Liu, Q. (2021). A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments. *Energy Reports*, *7*, 8176–8186. https://doi.org/10.1016/j.egyr.2021.08.126

Lim, E. (2022). B2b Artificial Intelligence Transactions: A Framework for Assessing Commercial Liability. *Singapore Journal of Legal Studies*, 46–74.

Liu, J., Wang, Y., Cao, J., Dong, Y., & Xie, X. (2021). Secure Adaptive-Event-Triggered Filter Design With Input Constraint and Hybrid Cyber Attack. *IEEE Transactions on Cybernetics*, *51*(8), 4000–4010. https://doi.org/10.1109/TCYB.2020.3003752

Malatji, M., Marnewick, A., & Von Solms, S. (2020). Validation of a socio-technical management process for optimising cybersecurity practices. *Computers & Security*, *95*, 101846. https://doi.org/10.1016/j.cose.2020.101846

Malatji, M., & Tolah, A. (2024). Artificial intelligence (AI) cybersecurity dimensions: A comprehensive framework for understanding adversarial and offensive AI. *AI and Ethics*. https://doi.org/10.1007/s43681-024-00427-4

Malatji, M., Von Solms, S., & Marnewick, A. (2019). Socio-technical systems cybersecurity framework. *Information & Computer Security*, *27*(2), 233–272. https://doi.org/10.1108/ICS-03-2018-0031

Malik, S., Malik, P., & Naim, A. (2024). *Opportunities and Challenges in New Generation Cyber Security Applications Using Artificial Intelligence, Machine Learning and Block Chain* (pp. 23–37). https://doi.org/10.1007/978-981-97-1249-6_2

Martin, T. (2022). On the Need for Collaborative Intelligence in Cybersecurity. *Electronics*, *11*(13), Article 13. https://doi.org/10.3390/electronics11132067

Meredith, J. R. (1981). The implementation of computer based systems. *Journal of Operations Management*, *2*(1), 11–21. https://doi.org/10.1016/0272-6963(81)90032-2

Merhi, M. I. (2023). An evaluation of the critical success factors impacting artificial intelligence implementation. *International Journal of Information Management*, *69*, 102545. https://doi.org/10.1016/j.ijinfomgt.2022.102545

Merhi, M. I., & Ahluwalia, P. (2019). Examining the impact of deterrence factors and norms on resistance to Information Systems Security. *Computers in Human Behavior*, *92*, 37–46. https://doi.org/10.1016/j.chb.2018.10.031

Mingers, J. (2001). Combining IS Research Methods: Towards a Pluralist Methodology. *Information Systems Research*, *12*(3), 240–259. https://doi.org/10.1287/isre.12.3.240.9709

Moabalobelo, T., Ngobeni, S., Molema, B., Pantsi, P., Dlamini, M., & Nelufule, N. (2023). Towards a Privacy Compliance Assessment Toolkit. *2023 IST-Africa Conference (IST-Africa)*, 1–8. https://doi.org/10.23919/IST-Africa60249.2023.10187837

Mohamed, N. (2023). Current trends in AI and ML for cybersecurity: A state-of-the-art survey. *Cogent Engineering*, *10*(2), 2272358. https://doi.org/10.1080/23311916.2023.2272358

Morland, J., Feagin, J., Orum, A., & Sjoberg, G. (1992). A Case for the Case Study. *Social Forces*, *71*, 240. https://doi.org/10.2307/2579984

Mtsweni, J., & Thaba, M. (2022). Building an Integrated Cyber Defence Capability for African Missions. *Journal of Information Warfare*, *21*(1), 17–34.

Mumford, E. (2006). The story of socio-technical design: Reflections on its successes, failures and potential. *Information Systems Journal*, *16*(4), 317–342. https://doi.org/10.1111/j.1365-2575.2006.00221.x

Mwim, E. N., & Mtsweni, J. (2022). Systematic Review of Factors that Influence the Cybersecurity Culture. In N. Clarke & S. Furnell (Eds.), *Human Aspects of Information Security and Assurance* (pp. 147–172). Springer International Publishing. https://doi.org/10.1007/978-3-031-12172-2_12

Myers, M. D. (2013). *Qualitative Research in Business and Management*. SAGE Publications.

Myers, M. D., & Klein, H. K. (2011). A Set of Principles for Conducting Critical Research in Information Systems. *MIS Quarterly*, *35*(1), 17–36. https://doi.org/10.2307/23043487

Naidoo, R., & Jacobs, C. (2023). Cyber Warfare and Cyber Terrorism Threats Targeting Critical Infrastructure: A HCPS-based Threat Modelling Intelligence Framework. *European Conference on Cyber Warfare and Security*, *22*(1), 311–318. https://doi.org/10.34190/eccws.22.1.1443

Naik, B., Mehta, A., Yagnik, H., & Shah, M. (2022). The impacts of artificial intelligence techniques in augmentation of cybersecurity: A comprehensive review. *Complex & Intelligent Systems*, *8*(2), 1763–1780. https://doi.org/10.1007/s40747-021-00494-8

Naseer, H., Desouza, K., Maynard, S. B., & Ahmad, A. (2024). Enabling cybersecurity incident response agility through dynamic capabilities: The role of real-time analytics. *European Journal of Information Systems*, *33*(2), 200–220. https://doi.org/10.1080/0960085X.2023.2257168

Nguyen, M.-D., Mallouli, W., Cavalli, A. R., & Montes de Oca, E. (2024). AI4SOAR: A Security Intelligence Tool for Automated Incident Response. *Proceedings of the 19th International Conference on Availability, Reliability and Security*, 1–8. https://doi.org/10.1145/3664476.3670450

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic Analysis: Striving to Meet the Trustworthiness Criteria. *International Journal of Qualitative Methods*, *16*(1), 1609406917733847. https://doi.org/10.1177/1609406917733847

Nyakasoka, L., & Naidoo, R. (2024). Understanding the inertial forces impeding dynamic cybersecurity learning capabilities: The case of a South African healthcare software services firm. *South African Computer Journal*, *36*(1). https://doi.org/10.18489/sacj.v36i1.18877

Oates, B. (2005). *Researching Information Systems and Computing*.

Okdem, S., & Okdem, S. (2024). Artificial Intelligence in Cybersecurity: A Review and a Case Study. *Applied Sciences*, *14*(22), Article 22. https://doi.org/10.3390/app142210487

Ozkan-Okay, M., Akin, E., Aslan, Ö., Kosunalp, S., Iliev, T., Stoyanov, I., & Beloev, I. (2024). A Comprehensive Survey: Evaluating the Efficiency of Artificial Intelligence and Machine Learning Techniques on Cyber Security Solutions. *IEEE Access*, *12*, 12229–12256. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3355547

Pan, G., Hackney, R., & Pan, S. L. (2008). Information Systems implementation failure: Insights from prism. *International Journal of Information Management*, *28*(4), 259–269. https://doi.org/10.1016/j.ijinfomgt.2007.07.001

Paré, G., & Elam, J. J. (1997). Using Case Study Research to Build Theories of IT Implementation. In A. S. Lee, J. Liebenau, & J. I. DeGross (Eds.), *Information Systems and Qualitative Research: Proceedings of the IFIP TC8 WG 8.2 International Conference on Information Systems and Qualitative Research, 31st May–3rd June 1997, Philadelphia, Pennsylvania, USA* (pp. 542–568). Springer US. https://doi.org/10.1007/978-0-387-35309-8_27

Petter, S., DeLone, W., & McLean, E. (2008). Measuring information systems success: Models, dimensions, measures, and interrelationships. *European Journal of Information Systems*, *17*(3), 236–263. https://doi.org/10.1057/ejis.2008.15

Phahlamohlaka, J., Theron, J., & Aschmann, M. (2024). *National Cybersecurity Implementation in South Africa*.

Pieterse, H. (2021). The Cyber Threat Landscape in South Africa: A 10-Year Review. *The African Journal of Information and Communication*, *28*. https://doi.org/10.23962/10539/32213

Poehlmann, N., Caramancion, K. M., Tatar, I., Li, Y., Barati, M., & Merz, T. (2021). *The Organizational Cybersecurity Success Factors: An Exhaustive Literature Review* (pp. 377–395). https://doi.org/10.1007/978-3-030-71017-0_27

Pollini, A., Callari, T. C., Tedeschi, A., Ruscio, D., Save, L., Chiarugi, F., & Guerri, D. (2022). Leveraging human factors in cybersecurity: An integrated methodological approach.

Cognition, Technology & Work, *24*(2), 371–390. https://doi.org/10.1007/s10111-021-00683-y

Proudfoot, J. G., Cram, W. A., & Madnick, S. (2023). Weathering the storm: Examining how organisations navigate the sea of cybersecurity regulations. *European Journal of Information Systems*, *0*(0), 1–24. https://doi.org/10.1080/0960085X.2024.2345867

Radanliev, P. (2024). Cyber diplomacy: Defining the opportunities for cybersecurity and risks from Artificial Intelligence, IoT, Blockchains, and Quantum Computing. *Journal of Cyber Security Technology*, 1–51. https://doi.org/10.1080/23742917.2024.2312671

Rahi, S. (2017). Research Design and Methods: A Systematic Review of Research Paradigms, Sampling Issues and Instruments Development. *International Journal of Economics & Management Sciences*, *06*(02). https://doi.org/10.4172/2162-6359.1000403

Rajapakse, D. P. P. K., & Thushara, S. C. (2023). Critical Failure Factors in ERP Implementation: A Systematic Literature Review. *Journal of Business and Technology*, *7*(1). https://doi.org/10.4038/jbt.v7i1.109

Ramos-Cruz, B., Andreu-Perez, J., & Martínez, L. (2024). The cybersecurity mesh: A comprehensive survey of involved artificial intelligence methods, cryptographic protocols and challenges for future research. *Neurocomputing*, *581*, 127427. https://doi.org/10.1016/j.neucom.2024.127427

Roba Abbas, K. M., Pitt, J., Vogel, K. M., & Zaferirakopoulos, M. (2023). Artificial Intelligence (AI) in Cybersecurity: A Socio-Technical Research Roadmap. *The Alan Turing Insitute*. https://www.turing.ac.uk/sites/default/files/2023-11/ai_in_cybersecurity.pdf

Sadok, M., Welch, C., & Bednar, P. (2020). A socio-technical perspective to counter cyber-enabled industrial espionage. *Security Journal*, *33*(1), 27–42. https://doi.org/10.1057/s41284-019-00198-2

Salem, A. H., Azzam, S. M., Emam, O. E., & Abohany, A. A. (2024). Advancing cybersecurity: A comprehensive review of AI-driven detection techniques. *Journal of Big Data*, *11*(1), 105. https://doi.org/10.1186/s40537-024-00957-y

Sarker, I. H. (2023). Machine Learning for Intelligent Data Analysis and Automation in Cybersecurity: Current and Future Prospects. *Annals of Data Science*, *10*(6), 1473–1498. https://doi.org/10.1007/s40745-022-00444-2

Sarker, I. H. (2024). Detecting Anomalies and Multi-attacks Through Cyber Learning: An Experimental Analysis. In I. H. Sarker (Ed.), *AI-Driven Cybersecurity and Threat Intelligence: Cyber Automation, Intelligent Decision-Making and Explainability* (pp. 61–77). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-54497-2_4

Sarker, I. H., Janicke, H., Mohammad, N., Watters, P., & Nepal, S. (2024). AI Potentiality and Awareness: A Position Paper from the Perspective of Human-AI Teaming in Cybersecurity. In P. Vasant, V. Panchenko, E. Munapo, G.-W. Weber, J. J. Thomas, R. Intan, & M. Shamsul Arefin (Eds.), *Intelligent Computing and Optimization* (pp. 140–149). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-50887-5_14

Sarker, I. H., Janicke, H., Mohsin, A., Gill, A., & Maglaras, L. (2024). Explainable AI for cybersecurity automation, intelligence and trustworthiness in digital twin: Methods, taxonomy, challenges and prospects. *ICT Express*, *10*(4), 935–958. https://doi.org/10.1016/j.icte.2024.05.007

Sarker, S., Chatterjee, S., Xiao, X., & Elbanna, A. (2019). The sociotechnical axis of cohesion for the IS discipline: Its historical legacy and its continued relevance. *MIS Q.*, *43*(3), 695–720. https://doi.org/10.25300/MISQ/2019/13747

Saunders, M., Thornhill, A., & Lewis, P. (2019). *Research methods for business students* (Eigth Edition). Pearson.

Seddon, P. B. (1997). A Respecification and Extension of the DeLone and McLean Model of IS Success. *Information Systems Research*. https://doi.org/10.1287/isre.8.3.240

Shakir, M. (2002). *The selection of case studies: Strategies and their applications to IS implementation case studies*. http://hdl.handle.net/10179/4373

Shamsan Saleh, A. M. (2024). Blockchain for secure and decentralized artificial intelligence in cybersecurity: A comprehensive review. *Blockchain: Research and Applications*, *5*(3), 100193. https://doi.org/10.1016/j.bcra.2024.100193

Shneiderman, B. (2022). *Human-Centered AI*. Oxford University Press.

Siphambili, N., Mahlasela, O., Baloyi, E., & Mukondeleli, E. (2024). A Review of the South African Public Sector's Capability in Combating Ransomware. *2024 4th International Multidisciplinary Information Technology and Engineering Conference (IMITEC)*, 493–499. https://doi.org/10.1109/IMITEC60221.2024.10850969

Sony, M., & Naik, S. (2020). Critical factors for the successful implementation of Industry 4.0: A review and future research direction. *Production Planning & Control*, *31*(10), 799–815. https://doi.org/10.1080/09537287.2019.1691278

Stahl, B. (2012). Morality, Ethics, and Reflection: A Categorization of Normative IS Research. *Journal of the Association for Information Systems*, *13*(8). https://doi.org/10.17705/1jais.00304

Taddeo, M., Jones, P., Abbas, R., Vogel, K., & Michael, K. (2023). Socio-Technical Ecosystem Considerations: An Emergent Research Agenda for AI in Cybersecurity. *IEEE Transactions on Technology and Society*, *4*(2), 112–118. https://doi.org/10.1109/TTS.2023.3278908

Taherdoost, H. (2022). Understanding Cybersecurity Frameworks and Information Security Standards—A Review and Comprehensive Overview. *Electronics*, *11*(14), 2181. https://doi.org/10.3390/electronics11142181

Tam, C., Balau, M., & Oliveira, T. (2024). What Influences People's Adoption of Cognitive Cybersecurity? *International Journal of Human–Computer Interaction*, *40*(23), 8295–8312. https://doi.org/10.1080/10447318.2023.2279411

Teichmann, F., Boticiu, S., & Sergi, B. S. (2023). The risk of abuse of arbitration proceedings in jurisdictions where corruption is pervasive. *Journal of Economic Criminology*, *2*, 100032. https://doi.org/10.1016/j.jeconc.2023.100032

Teo, Z. L., Quek, C. W. N., Wong, J. L. Y., & Ting, D. S. W. (2024). Cybersecurity in the generative artificial intelligence era. *Asia-Pacific Journal of Ophthalmology*, *13*(4), 100091. https://doi.org/10.1016/j.apjo.2024.100091

Timcke, S., Gaffley, M., & Rens, A. (2023). The centrality of cybersecurity to socioeconomic development policy: A case study of cyber-vulnerability at South Africa's Transnet. *The African Journal of Information and Communication (AJIC)*, *32*, 1–28. https://doi.org/10.23962/ajic.i32.16949

Uchendu, B., Nurse, J. R. C., Bada, M., & Furnell, S. (2021). Developing a cyber security culture: Current practices and future needs. *Computers & Security*, *109*, 102387. https://doi.org/10.1016/j.cose.2021.102387

Van Haastrecht, M., Yigit Ozkan, B., Brinkhuis, M., & Spruit, M. (2021). Respite for SMEs: A Systematic Review of Socio-Technical Cybersecurity Metrics. *Applied Sciences*, *11*(15), 6909. https://doi.org/10.3390/app11156909

van Niekerk, B., Ramluckan, T., & Collard, A. (2023). *A South African Perspective on Cybercrime During the Pandemic* (pp. 177–209).

Wada, I. U., Izibili, G. O., Babayemi, T., Abdulkareem, A., Macaulay, O. M., & Emadoye, A. (2025). AI-driven cybersecurity in higher education: A systematic review and model evaluation for

enhanced threat detection and incident response. *World Journal of Advanced Research and Reviews*, *25*(3), 2233–2245. https://doi.org/10.30574/wjarr.2025.25.3.0989

Wade, M., & Hulland, J. (2004). Review: The Resource-Based View and Information Systems Research: Review, Extension, and Suggestions for Future Research. *MIS Quarterly*, *28*(1), 107–142. https://doi.org/10.2307/25148626

Walker, G. H., Stanton, N. A., Salmon, P. M., & Jenkins, D. P. (2008). A review of sociotechnical systems theory: A classic concept for new command and control paradigms. *Theoretical Issues in Ergonomics Science*, *9*(6), 479–499. https://doi.org/10.1080/14639220701635470

Walsham, G. (1995). The Emergence of Interpretivism in IS Research. *Information Systems Research*. https://doi.org/10.1287/isre.6.4.376

Walsham, G. (2006). Doing interpretive research. *European Journal of Information Systems*, *15*(3), 320–330. https://doi.org/10.1057/palgrave.ejis.3000589

Walsham, G., & Waema, T. (1994). Information systems strategy and implementation: A case study of a building society. *ACM Trans. Inf. Syst.*, *12*(2), 150–173. https://doi.org/10.1145/196734.196744

Wessel, L., Baiyere, A., Ologeanu-Taddei, R., Cha, J., & Jensen, T. B. (2021). Unpacking the Difference Between Digital Transformation and IT-Enabled Organizational Transformation. *Journal of the Association for Information Systems*, *22*(1). https://doi.org/10.17705/1jais.00655

William H. Delone, & and, E. R. M. (2003). The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of Management Information Systems*, *19*(4), 9–30. https://doi.org/10.1080/07421222.2003.11045748

Williams, J., & Ramaprasad, A. (1996). A taxonomy of critical success factors. *European Journal of Information Systems*, *5*(4), 250–260. https://doi.org/10.1057/ejis.1996.30

Witanto, E. N., Oktian, Y. E., & Lee, S.-G. (2022). Toward Data Integrity Architecture for Cloud-Based AI Systems. *Symmetry*, *14*(2), Article 2. https://doi.org/10.3390/sym14020273

Wu, J.-H., & Wang, Y.-M. (2006). Measuring KMS success: A respecification of the DeLone and McLean's model. *Information & Management*, *43*(6), 728–739. https://doi.org/10.1016/j.im.2006.05.002

Yeoh, W., Wang, S., Popovič, A., & Chowdhury, N. H. (2022). A systematic synthesis of critical success factors for cybersecurity. *Computers & Security*, *118*, 102724. https://doi.org/10.1016/j.cose.2022.102724

Yin, R. K. (2017). *Case Study Research and Applications: Design and Methods*. SAGE Publications.

Yuan, H., & Li, S. (2022). *Cyber Security Risks of Net Zero Technologies* (p. 11). https://doi.org/10.1109/DSC54232.2022.9888883

Zhang, J., Pan, L., Han, Q.-L., Chen, C., Wen, S., & Xiang, Y. (2022). Deep Learning Based Attack Detection for Cyber-Physical System Cybersecurity: A Survey. *IEEE/CAA Journal of Automatica Sinica*, *9*(3), 377–391. IEEE/CAA Journal of Automatica Sinica. https://doi.org/10.1109/JAS.2021.1004261

Zibak, A., Sauerwein, C., & Simpson, A. (2021). A success model for cyber threat intelligence management platforms. *Computers & Security*, *111*, 102466. https://doi.org/10.1016/j.cose.2021.102466

Zwilling, M., Klien, G., Lesjak, D., Wiechetek, Ł., Cetin, F., & Basim, H. N. (2022). Cyber Security Awareness, Knowledge and Behavior: A Comparative Study. *Journal of Computer Information Systems*, *62*(1), 82–97. https://doi.org/10.1080/08874417.2020.1712269

8 APPENDIX A: INTERVIEW PROTOCOL

Introduction and Work Background:

1. What is your role at Edu X?
2. How many years have you been employed with Edu X?
3. How many years of experience do you have working with cybersecurity?

Social factors

4. How would you describe your experience of working with AI-driven cybersecurity tools?
5. What are the organizational challenges that you came across when implementing AI cybersecurity?
6. How has AI cybersecurity influenced the way you work? Please elaborate
7. How did regulatory requirements impact the implementation of AI cybersecurity? Please elaborate
8. How were users prepared for the change that came with AI-cybersecurity implementation? Who was involved, how was it done.
9. What ethical considerations, if any, were you mindful of before implementing AI-cybersecurity?
10. What are the challenges encountered with individuals, teams or other divisions while implementing AI cybersecurity?
11. How did users react to the implementation of AI-cybersecurity? Did they willingly accept, did they resist?
12. How was the support/involvement top management in the AI-cybersecurity initiative?

Technical Factors

13. What are the benefits of AI cybersecurity from a technical point of view? Your perception
14. What technical concerns do you have about AI cybersecurity?
15. What technical challenges/barriers did you come across?
16. What are the shortcomings of AI cybersecurity from a technical point view?

17. How effective has AI cybersecurity been in detecting and mitigating threats?
18. How effective are the AI cybersecurity systems to continuously learn and adapt to emerging cybersecurity threats and changing attack vectors? Any issues on false positives and negatives
19. How has the integration of AI-cybersecurity with existing systems been?

Critical success factors

20. What were your goals in implementing AI cybersecurity?
21. What lessons can be learnt from implementing AI cybersecurity?
22. Have you noticed any improvement in cybersecurity since the implementation of AI cybersecurity?
23. What strategies did you find most effective in implementing AI cybersecurity?
24. What has been the success of this AI-cybersecurity initiative? And what were the shortcomings?

Outcomes

25. Any other comments?
26. Who else do you think I should talk to?

Note 1: Some interviews questions may involve fewer questions to accommodate the participant's role and time constraints.
Note 2: Some questions will be adjusted to better fit the participant's role.
Note 3: Some questions will be further explored with appropriate probes.

9 APPENDIX B: PARTICIPANT INFORMATION SHEET (PIS)



University of the Witwatersrand | Faculty of Commerce, Law and Management |School of Business Sciences Commerce| Law & MNGT Building (CLM)| West Campus| 09/09/2024

Research participant
South African Qualifications Authority
SAQA House, 1067 Arcadia St,
Hatfield,
Pretoria, 0028

Dear Sir / Madam

Participant Information Sheet (PIS)

My name is Awonke Mamane, I am a Master of Commerce Information systems student at the University of the Witwatersrand, Johannesburg. My supervisor is Prof. Rennie Naidoo, I am conducting a research study about the implementation of AI cybersecurity. The study title is CRITICAL SUCCESS FACTORS FOR AI CYBERSECURITY IMPLEMENTATION IN A SOUTH AFRICAN STATE-OWNED ENTITY: A SOCIO-TECHNICAL APPROACH.

I am inviting you to take part in an interview. If you decide to take part, your participation in this research study will last about 30 minutes. The interview activity will take place at [redacted] at your most suitable time that we will agree upon your giving consent to participate in this study.

With your permission, I would like to audio record the interview. This data will be stored in a password encrypted folder and will be deleted after one years. Only the researcher will have access to the data. During the research activity, I will need to ask for some personal information about you, including your role in the organisation.

The interview will be confidential and anonymous. When I share the results of the research study, I will not include your name or anything else that could identify you. With your permission, other researchers may use the data collected from this research study, but your name and any personal information will not be used or passed on.

If you decide to take part in the research study, it should be because you want to volunteer. You do not have to take part. You can stop being in the study at any time. You do not have to answer any questions if you do not want to. You will not get any direct benefits if you choose to join the research study. You will not lose any services, benefits or rights you would normally have if you decide not to join. Taking part in the research study will not cost you anything. You will not be paid for being in this research study.

The risks for this research study are no more than what happens in everyday work life, some of the questions asked may make you feel sad or upset. If this happens, I will stop the interview and continue another time.

This research study will be written up as a research report. The report will be available on the university library website. If you would like to receive a summary of this report, I will be happy to send it to you.

If you have any questions during or afterwards about this research study, feel free to contact me on the details listed below. If you have any concerns or complaints about the ethical procedures of this research study, you are welcome to contact the University Human Research Ethics Committee (Non-Medical), telephone +27(0) 11 717 1408, email hrecnon-medical@wits.ac.za.

Yours sincerely,



Researcher:
Awonke Mamane, 2138857@students.wits.ac.za, +27(0) 732794077

Supervisor:
Rennie Naidoo, rennie.naidoo@wits.ac.za , +27(0) 11 717 8141

10 APPENDIX C: CONSENT FORM

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

University of the Witwatersrand | Faculty of Commerce, Law and Management |School of Business Sciences Commerce| Law & MNGT Building (CLM)| West Campus| 09/09/2024

RESEARCH CONSENT FORM

Title of the Study:

CRITICAL SUCCESS FACTORS FOR AI CYBERSECURITY IMPLEMENTATION IN A SOUTH AFRICAN STATE-OWNED ENTITY: A SOCIO-TECHNICAL APPROACH

Principal Investigator:

Awonke Mamane| University of the Witwatersrand| Faculty of Commerce, Law and Management |School of Business Sciences Commerce|0732794077

Purpose of the Study:

You are invited to participate in a research study conducted by Awonke Mamane at the South African Qualifications Authority. The purpose of this study is to seek to understand the critical success factors in the implementation of artificial intelligence cybersecurity according to end-users and ICT professionals.

Risks:

There are no known risks associated with participating in this study. However, if you feel uncomfortable at any point during the research, you may stop participating at any time.

Consent:

I, [redacted], by signing this form acknowledge that I have read and understood the information provided above, and that I agree to participate in this study.

- I understand that my participation is voluntary and that I may withdraw at any time without penalty.
- I understand that there will be no compensation for participating in this study.
- I understand the procedures and any risks involved in the study.
- I agree to the terms outlined in this consent form.
- The research study was explained to me and I understand what this study is about.
- I agree that the interview may be audio recorded.
- I agree that direct quotations from my interview may be used by the researcher.

Participant's Name (Printed): [redacted]

Participant's Signature: [redacted] **Date**: 05/11/2024

Researcher's Name (Printed): Awonke Mamane

Researcher's Signature: [redacted] **Date**: 05 November 2024

Page **1** of **1**

11 APPENDIX D: ETHICS CLEARANCE

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

SCHOOL OF BUSINESS SCIENCES ETHICS COMMITTEE
CONSTITUTED UNDER THE UNIVERSITY HUMAN RESEARCH ETHICS COMMITTEE (NON-MEDICAL)

CLEARANCE CERTIFICATE **PROTOCOL NUMBER: CBUSE2287**

PROJECT TITLE	Critical success factors for AI cybersecurity implementation in a South African state-owned entity: A socio-technical approach.
INVESTIGATOR	Awonke Mamane
SCHOOL/DEPARTMENT OF INVESTIGATOR	School of Business Sciences
DATE CONSIDERED	18 October 2024
DECISION OF THE COMMITTEE	Approved Unconditionally
RISK LEVEL	Low Risk
EXPIRY DATE	31 December 2027

ISSUE DATE OF CERTIFICATE 29 October 2024 **CHAIRPERSON** _____
(Neetu Ramsaroop)

cc: Supervisor: Rennie Naidoo

DECLARATION OF INVESTIGATOR

To be completed in duplicate and **ONE COPY** returned to the Chairperson of the School/Department ethics committee.

I fully understand the conditions under which I am are authorized to carry out the abovementioned research and I guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee.

_____ 30 / 10 / 2024
Signature **Date**

PLEASE QUOTE THE PROTOCOL NUMBER ON ALL ENQUIRIES

12 APPENDIX E: PERMISSION TO CONDUCT THE CASE STUDY

[redacted]

5 September 2024

Mr. Awonke Mamane
[redacted]

Email: amamane@[redacted]

Dear Mr Awonke Mamane

AUTHORIZATION TO CONDUCT ACADEMIC RESEARCH

I am pleased to inform you that you have been granted authorization to conduct your academic research within the [redacted] for your Master of Commerce (Information Systems) case study research titled: **CRITICAL SUCCESS FACTORS FOR AI CYBERSECURITY IMPLEMENTATION IN A SOUTH AFRICAN STATE-OWNED ENTITY: A SOCIO-TECHNICAL APPROACH**. This approval reflects our support for your academic endeavors and commitment to contributing valuable insights through your research.

You will have access to the necessary resources, data, and personnel required for the successful completion of your study. We trust that you will conduct your research with the highest level of professionalism, respecting the privacy and confidentiality of all involved, and adhering to ethical standards.

Please ensure that you coordinate with the relevant departments and personnel to facilitate your research activities. If you encounter any challenges or require additional support, do not hesitate to reach out to the appropriate department heads.

I wish you success in your research and look forward to the valuable contributions it will bring to our institution and the broader academic community.

Kind Regards

[redacted]

CHIEF EXECUTIVE OFFICER

[redacted]

13 APPENDIX E: TURNITIN REPORT

