

https://dspace.org/

Ethical AI Readiness within Utilities in South Africa

Masipa, Isaiah Mmakhudu

2024
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/49227
Masipa, Isaiah Mmakhudu. (2024). Ethical AI Readiness within Utilities in South Africa [Master`s dissertation, University of the Witwatersrand, Johannesburg]. WIReDSpace. https://hdl.handle.net/10539/49227

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

Research Report

Master of Management in the field of Digital Business

Ethical AI Readiness within Utilities in South Africa

Isaiah Mmakhudu Masipa

Student No: 744853

Email: 744853@students.wits.ac.za : Mobile Phone: +27 76 941 4885

MMDB 2024: BUSA 7480A

Supervisor: Professor Tendani Mawela

Tendani.mawela@up.ac.za | 012 420 6484

A research report submitted to the Faculty of Commerce, Law and Management, University of the Witwatersrand, in partial fulfilment of the requirements for the degree of Master of Management in the field of Digital Business

Johannesburg, 2024

DECLARATION

I, Isaiah Mmakhudu Masipa (Student No: 744853), declare that this research report is my own work except as indicated in the references and acknowledgements. It is submitted in partial fulfilment of the requirements for the degree of Master of Management in the field of Digital Business at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in this or any other university.

Name: Isaiah M. Masipa Signature: 

Signed at …Germiston……

On the ……19th…….. Day of …February…… 2025

ACKNOWLEDGEMENTS

This research journey has been a transformative experience, and I am deeply grateful to the many individuals who supported and guided me along the way.

First, my heartfelt thanks to my supervisor, Professor Tendani Mawela, for your expert guidance, patience, and unwavering belief in my work. Your insights challenged me to push boundaries and think critically, making this study possible.

I extend my gratitude to the faculty members of the University of the Witwatersrand, especially those who offered their expertise, encouragement, and inspiration throughout my academic journey. To the respondents who participated in this study, thank you for sharing your time and insights, which formed the backbone of this research.

To my friends, colleagues, and peers, thank you for your camaraderie and encouragement throughout this process. Finally, to all who contributed directly or indirectly to this research, your support and contributions will forever be appreciated.

This work reflects the invaluable contributions of many, and I extend my deepest gratitude to each individual involved.

DEDICATION

This work is lovingly dedicated to my family, whose unwavering support and encouragement have made this journey possible.

To my two incredible children, Mohau and Lubanzi: You are my greatest motivation. Your laughter, curiosity, and boundless energy remind me every day of the beauty and potential in the world.

To my partner, Noxolo: Your love and belief in me have been my anchor. Thank you for your sacrifices, your patience, and for walking this path with me. You inspire me to keep going, even when the road feels steep.

To my mother, Catherine Masipa: Your wisdom and prayers have been my guiding light. You taught me the value of hard work and the importance of kindness. This is as much a reflection of your sacrifices as it is of my efforts.

TABLE OF CONTENTS

LIST OF TABLES viii

LIST OF FIGURES ix

LIST OF ACRONYMS x

CHAPTER 1. INTRODUCTION 1

1.1 STATEMENT OF PURPOSE 1
1.2 BACKGROUND OF THE STUDY 1
1.3 RESEARCH PROBLEM 3
1.4 RESEARCH OBJECTIVES 4
1.5 RATIONALE 4
1.6 DELIMITATIONS OF THE STUDY 5
1.7 DEFINITION OF TERMS 7
1.8 ASSUMPTIONS 8
1.9 CHAPTER OUTLINE 9

CHAPTER 2. LITERATURE REVIEW AND THEORETICAL FRAMEWORK 11

2.1 INTRODUCTION 11
2.2 DEFINITION OF TOPIC OR BACKGROUND DISCUSSION 13
2.3 RESEARCH OBJECTIVE ..... **ERROR! BOOKMARK NOT DEFINED.**
2.3.1 TO DETERMINE WHICH RELEVANT ETHICAL GUIDELINES AND FRAMEWORKS ARE CURRENTLY ADOPTED BY SOUTH AFRICAN ENERGY UTILITY PROVIDERS FOR AI IMPLEMENTATION 14
2.3.2 TO DETERMINE IF NECESSARY MEASURES ARE IN PLACE TO IDENTIFY AND MITIGATE BIASES IN AI SYSTEMS USED BY SOUTH AFRICAN ENERGY UTILITY PROVIDERS. 15
2.3.3 TO ASSESS HOW ENERGY UTILITY PROVIDERS ENSURE FAIRNESS IN AI-DRIVEN DECISIONS THAT AFFECT CUSTOMERS AND STAKEHOLDERS. 16
2.3.4 TO EVALUATE IF THERE ARE MECHANISMS TO ENSURE ACCOUNTABILITY FOR AI-DRIVEN DECISIONS AND THEIR CONSEQUENCES. 16
2.3.5 PROPOSITION 17
2.5 ANALYTICAL FRAMEWORK 18
2.5.1 THEORETICAL FRAMEWORK 18
2.5.2 CONCEPTUAL FRAMEWORK 21
2.6 CONCLUSION OF LITERATURE REVIEW **ERROR! BOOKMARK NOT DEFINED.**
2.6.1 PROPOSITION ..... **ERROR! BOOKMARK NOT DEFINED.**

CHAPTER 3. RESEARCH METHODOLOGY 28

3.1 RESEARCH APPROACH 28
3.2 RESEARCH DESIGN 30

3.3 DATA COLLECTION METHODS 31
3.4 POPULATION AND SAMPLE 32
3.4.1 POPULATION 32
3.4.2 SAMPLE 32
3.4.3 SAMPLING METHOD 33
3.5 THE RESEARCH INSTRUMENT 34
3.6 PROCEDURE FOR DATA COLLECTION 36
3.7 DATA ANALYSIS STRATEGIES AND INTERPRETATION 36
3.8 POSSIBLE LIMITATIONS AND CHALLENGES OF THE STUDY 38
3.9 QUALITY ASSURANCE 40
3.9.1 TRANSFERABILITY 40
3.9.2 CREDIBILITY 40
3.9.3 DEPENDABILITY 41
3.10 ETHICAL CONSIDERATIONS 41

CHAPTER 4. PRESENTATION OF FINDINGS 43

4.1 INTRODUCTION 43
4.2 RESULTS PERTAINING TO PROPOSITION 1 44
4.3 RESULTS PERTAINING TO PROPOSITION 2 47
4.4 SUMMARY OF THE RESULTS/FINDINGS 53

CHAPTER 5. DISCUSSION OF THE RESULTS OR FINDINGS . 57

5.1 INTRODUCTION 57
5.2 DEMOGRAPHIC PROFILE OF RESPONDENTS 57
5.3 DISCUSSION PERTAINING TO PROPOSITIONS 58
5.4 DISCUSSION PERTAINING TO SECONDARY PROPOSITION 59
5.5 CONCLUSION 64

CHAPTER 6. CONCLUSIONS & RECOMMENDATIONS 65

6.1 INTRODUCTION 65
6.2 CONCLUSIONS REGARDING MAIN RESEARCH OBJECTIVE 65
6.3 CONCLUSIONS REGARDING RESEARCH SUB-QUESTIONS 66
6.4 RECOMMENDATIONS 67
6.4.1 ENVIRONMENTAL-RELATED RECOMMENDATIONS 67
6.4.2 ORGANIZATIONAL-RELATED RECOMMENDATIONS 68
6.4.3 TECHNOLOGY-RELATED RECOMMENDATIONS 69
6.5 SUGGESTIONS FOR FURTHER RESEARCH 72

REFERENCES **79**

APPENDIX A **83**

APPENDIX B **84**

APPENDIX C **85**

APPENDIX D **101**

APPENDIX E **103**

APPENDIX F **104**

LIST OF TABLES

Table 1: Definition of Terms……. **Error! Bookmark not defined.**

Table 2: Profile of respondents……..... **Error! Bookmark not defined.**

Table 3: THEMATIC ANALYSIS PROCESS ……..……..37

Table 4: Distribution of Respondents across categories…… **Error! Bookmark not defined.**

Table 5: Ethical Guideline and Frameworks Adoption…… 46

Table 6: Transparency and Accountability…… 47

Table 7: Best Practice to Ensure Transparency…….. **Error! Bookmark not defined.**

Table 8: Challenges to Transparency…….. **Error! Bookmark not defined.**

Table 9: Measure to Mitigate Bias……. **Error! Bookmark not defined.**

Table 10: Effectiveness of Measures……... **Error! Bookmark not defined.**

Table 11: Ethical Concern……...**Error! Bookmark not defined.**

Table 12: Future practices……..51

Table 13: Improvement Areas……**Error! Bookmark not defined.**

Table 14: Consistency Table …….....78

LIST OF FIGURES

Figure 1 : TOE Framework.....**Error! Bookmark not defined.**

Figure 2 : Ethical Decision –Making Model 22

Figure 3 : Amalgamation of TOE and Ethical decision making**Error! Bookmark not defined.**

Figure 4 - Ethical Frameworks Commonly Used .**Error! Bookmark not defined.**

Figure 5 : Transparency level for AI -driven Decision-Making**Error! Bookmark not defined.**

Figure 6 : Proposed improvement for Stakeholder Engagement .....**Error! Bookmark not defined.**

Figure 7 : Proposed Improvement on Accountability in AI**Error! Bookmark not defined.**

Figure 8 : Frequency of themes across categories 101

Figure 9 : Sentiment Pie Chart - About the current AI adoption in utilities..... 102

Figure 10 : Accountability Improvement in AI – Word-cloud..... 102

Figure 4 - Ethical Frameworks Commonly Used ……. **Error! Bookmark not defined.**

Figure 5 : Transparency level for AI -driven Decision-Making ……...**Error! Bookmark not defined.**

Figure 6 : Proposed improvement for Stakeholder Engagement …….**Error! Bookmark not defined.**

Figure 8: Frequency of themes across categories ……101

Figure 9 : Sentiment Pie Chart - About the current AI adoption in utilities ……. 102

Figure 10 : Accountability Improvement in AI – Word-cloud ……102

Figure 10: Accountability Improvement in AI – Word-cloud ……102

LIST OF ABBREVIATIONS

ACRONYM	MEANING
4IR	Fourth Industrial Revolution
AI	Artificial Intelligence
AIML	Artificial Intelligence and Machine Learning
BCM	Business Continuity Management
CEO	Chief Executive Officer
CIO	Chief Information Officer
CSP	Concentrated Solar Power
CSR	Corporate Social Responsibility
CTO	Chief Technology Officer
GDPR	General Data Protection Regulation
ICT	Information and Communication Technology
IoT	Internet of Things
IR	Industrial Revolution
IT	Information Technology
ML	Machine Learning
NGO	Non-Governmental Organization
NLP	Natural Language Processing
OT	Operational Technology
PV	Photovoltaic
R&D	Research and Development
RES	Renewable Energy Sources

ABSTRACT

This study investigates the readiness of South African energy utility providers to implement artificial intelligence (AI) in ways that uphold ethical standards. While AI offers clear operational advantages—such as improved efficiency and data-driven decision-making—it also raises ethical concerns related to algorithmic bias, decision transparency, and accountability for outcomes.

A qualitative research design was adopted to gain in-depth, context-rich insights into the ethical dimensions of AI adoption. Data were collected through structured, open-ended questionnaires and semi-structured interviews conducted with a purposive sample of stakeholders, including utility executives, IT professionals, ethics officers, regulatory representatives, and customers. This approach enabled the study to explore stakeholders' lived experiences, perceptions, and interpretations of ethical AI readiness within their organizations.

Thematic analysis was used to identify recurring patterns and key concerns across stakeholder responses. The findings reveal that while ethical AI is acknowledged as a strategic priority, implementation efforts are inconsistent and often lack formal policy direction. Most organizations operate without standardized ethical guidelines or robust governance frameworks, resulting in fragmented practices and limited oversight. Respondents highlighted specific challenges, such as the absence of tools to detect algorithmic bias and inadequate mechanisms for ensuring transparent, accountable decision-making.

The study concludes that meaningful ethical AI integration requires sector-wide regulatory alignment, investment in targeted training, and multi-stakeholder collaboration. A unified ethical framework, grounded in local and international best practices, is essential to ensure that AI systems used by utility providers are fair, transparent, and accountable.

KEYWORDS:

Ethical AI, Artificial Intelligence, 4IR and AI, Utility Providers, Energy Sector, Machine Learning in Utilities, and Fairness in AI.

CHAPTER 1. INTRODUCTION

1.1 Statement of purpose

The purpose of this study is to assess the ethical readiness of South African energy utility providers in adopting artificial intelligence (AI) technologies. It aims to examine the extent to which ethical principles—such as fairness, transparency, accountability, and bias mitigation—are integrated into AI systems across selected providers. By exploring organizational practices, governance mechanisms, and stakeholder perspectives, the study seeks to offer practical recommendations for the responsible and ethical deployment of AI in the energy sector.

1.2 Background of the study

The energy utilities sector in South Africa is fundamental to the country's industrial base, economic growth, and public welfare (Kruger & Steyn, 2020; Papagiannidis, Harris, & Morton, 2023). It encompasses more than electricity production and distribution; it also involves governance, policy, infrastructure resilience, and service delivery critical to sustaining human life and industrial competitiveness. Eskom, a state-owned utility responsible for generating approximately 90% of the country’s electricity, dominates South Africa’s electricity supply industry. Eskom operates alongside municipal electricity distributors and a growing number of independent power producers (IPPs), particularly in the renewable energy space (Department of Mineral Resources and Energy [DMRE], 2023; Eberhard & Naude, 2021).

Historically, the sector has been heavily regulated due to its strategic importance to the economy and society. The National Energy Regulator of South Africa (NERSA) oversees economic regulation and ensures affordability, reliability, and equitable access to electricity services (CIDB, 2017; De Silva & Alahakoon, 2022). However, the sector faces mounting pressure to modernize in response to global technological trends and local infrastructure failures. These pressures

have made digital transformation imperative for utility providers to remain viable and sustainable.

Technological advances such as Artificial Intelligence (AI), the Internet of Things (IoT), Big Data analytics, and autonomous systems are increasingly being adopted in the energy sector to enhance grid efficiency, predictive maintenance, customer service, and energy distribution (Aguilar, Díaz-Rodríguez, & Prieto, 2021; Clark & Valente, 2020). In South Africa, Eskom and some municipalities have begun piloting or implementing AI-based systems for demand forecasting, fault detection, and customer engagement (Naude, 2022; Diallo, Moyo, & Ntsimane, 2024).

While these innovations offer operational benefits, they also introduce complex ethical dilemmas related to algorithmic bias, data privacy, decision transparency, and institutional accountability (Floridi & Cowls, 2019; Jobin, Ienca, & Vayena, 2019). For instance, AI algorithms used to manage load shedding or prioritize maintenance could inadvertently favor some communities over others, raising questions of equity and fairness. In a country like South Africa, where socio-economic disparities are deeply entrenched, the ethical deployment of AI is not merely a technical issue but a socio-political imperative.

The onset of the Fourth Industrial Revolution has intensified these challenges, prompting the urgent need for governance frameworks that safeguard human rights while facilitating innovation (Schwab, 2017). As utility providers in South Africa strive to adopt AI tools, it becomes crucial to assess their readiness to navigate these ethical imperatives, particularly given the sector's fragile public trust and strategic national importance.

Therefore, this study aims to investigate the extent to which leading South African energy utility providers are prepared to adopt AI technologies ethically. It explores whether appropriate ethical frameworks, governance mechanisms, and accountability systems are in place to mitigate risks and promote fair, transparent, and responsible AI deployment.

1.3 Research problem

The rapid advancements of the Fourth Industrial Revolution (4IR) have introduced transformative technologies, including Artificial Intelligence (AI), into critical sectors such as energy utility provision. While these technologies promise enhanced efficiency and innovation, they also present significant ethical challenges. A major concern is the minimal or complete absence of human oversight in AI decision-making processes, which heightens the risk of unintentional unethical outcomes (De Silva & Alahakoon, 2022). Current ethical frameworks, though recommended, often lack enforceable mechanisms to address contraventions effectively, relying instead on soft deterrents such as reputational damage or professional restrictions. These measures have proven inadequate in ensuring accountability and adherence to ethical principles (Hagendorff, 2022).

Furthermore, every stage of the AI lifecycle—including data collection, algorithm design, optimization, and ongoing oversight—requires stringent measures to detect and address ethical risks, such as bias and inequitable outcomes (Aguilar, Díaz-Rodríguez, & Prieto, 2021).

Currently, most South African energy utility providers lack standardized ethical frameworks or enforceable mechanisms to guide the responsible use of artificial intelligence (AI). For instance, while Eskom has initiated digital transformation efforts through smart grid technologies and predictive maintenance systems, there is limited evidence of an overarching ethical governance structure guiding the use of AI in decision-making (Naude, 2022). The absence of such frameworks increases the risk of unintentional harm, particularly in contexts like automated load management, where AI-based decisions may inadvertently reinforce existing inequalities in electricity distribution. Furthermore, stakeholder mistrust can be amplified in the absence of transparent AI systems—especially when affected communities are excluded from decision-making processes or lack access to recourse mechanisms. Although ethical AI principles have been widely advocated in global discourse, there remains limited empirical research examining how South African energy utilities integrate these principles in

practice, or how organizational, technological, and regulatory conditions shape such implementation.

Therefore, the central problem this study addresses is the absence of coherent, enforceable, and context-specific ethical AI governance within South African energy utilities—an omission that hinders responsible innovation and exposes both institutions and citizens to operational, legal, and ethical vulnerabilities.

1.4 Research Objective

“To identify and analyse if South African energy utility providers are ready for ethical AI implementation”

Sub- Objectives:

1. To determine if relevant ethical guidelines and frameworks are currently adopted by South African energy utility providers for AI implementation.

2. To determine if measures are in place to detect and address biases present in AI systems utilized by South African energy utility providers.

3. To assess how energy utility providers ensure fairness in AI-driven decisions that affect customers and stakeholders.

4. To evaluate if mechanisms exist to ensure accountability for AI-driven decisions and their consequences.

1.5 Rationale

The incentive for this investigation stems from the increasing reliance on technology and automation, alongside the transformative influence of the Fourth Industrial Revolution (4IR), is significantly altering South Africa's industrial and economic landscape. Energy utility providers are central to this transformation, leveraging AI technologies to enhance operational efficiency, optimize resources, and improve service delivery (Schwab, 2017). Nevertheless, this shift presents

critical challenges related to ethical considerations and systemic readiness, which demand careful investigation.

For instance, several factors shape the integration of AI within the energy utility sector including cost, regulation, and the need for skilled personnel. Implementing and maintaining AI systems can be financially demanding, especially for smaller companies with limited resources. Additionally, the regulatory environment plays a significant role in AI adoption, where supportive regulations can facilitate implementation, while stringent ones may hinder progress. Moreover, the availability of skilled personnel to develop and manage AI systems is crucial for successful adoption (Slate et al., 2024).

Despite these advancements, AI adoption in South Africa is considered low, with experts suggesting that large organizations are not AI-ready. However, research indicates that organizations are AI-ready, and the low adoption is mainly because AI use cases have not been evident (Seymour, 2023).

Ethical challenges also emerge as a prominent issue in AI adoption. Incorporating AI into existing systems necessitates a thorough assessment of ethical considerations. The ethical deployment of AI presents distinct challenges that must be addressed. As advancements in AI research and adoption continue, establishing clear guidelines is essential to ensure that AI development adheres to ethical principles (Mhlambi, 2020).

1.6 Delimitations of the study

The scope of this study is intentionally narrowed to ensure a focused and manageable examination of AI readiness in major South African energy utility providers. This delimitation addresses the following areas:

a. **Geographic Scope**: - The study is confined to energy utility providers operating in South Africa, allowing for a more detailed analysis of its AI's ethical considerations and practices.

b. **Technology Focus**: - This study examines the incorporation of artificial intelligence (AI) within the domains of IT, OT, and cyberspace. Other areas

of AI application, such as customer service or automation in other sectors, are outside the scope of this study.

c. **Time Frame**: Data collection and analysis are be based on recent developments, specifically within the last two years, reflecting the most current practices and frameworks for AI implementation.

d. **Research Methods**: The study uses a qualitative approach, incorporating questionnaire with relevant professional stakeholders within the selected energy utility providers.

e. **Ethical Focus**: This study examines the readiness specifically related to AI implementation in the energy utility sector. Broader legal issues, such as those pertaining to ethical AI in healthcare or other industries, are not addressed in this research.

1.7 Definition of terms

Table 1 : Definition of Terms

Term	Definition	Citation
Artificial Intelligence (AI)	A branch of computer science focused on creating systems capable of performing tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making, and language translation.	(Russell & Norvig, 2021)
Ethical AI	The process of creating, implementing, and managing AI systems in a way that upholds fairness, accountability, transparency, and safeguards user privacy and rights.	(Floridi & Cowls, 2019)
Bias in AI	Systematic and unfair discrimination in AI systems that can occur when the data used to train the model is unbalanced or reflects historical inequities, leading to outcomes that favor certain groups over others.	(Hagendorff, 2022)
Fairness	Ensuring that AI systems operate impartially, ensuring no undue advantage or disadvantage to any individual or group unfairly, and that decisions made by these systems are just and equitable.	(De Silva & Alahakoon, 2022)
Transparency	The quality of AI systems being understandable and interpretable by stakeholders, allowing them to see how decisions are formulated and the degree to which algorithms impact those choices.	(Floridi & Cowls, 2019)
Accountability	The obligation of organizations and ensuring that AI systems are utilized responsibly by individuals, including mechanisms for identifying and addressing any adverse effects that may arise from their use.	(Jobin, Ienca, & Vayena, 2019)
Ethical Guidelines	A set of principles and standards designed to guide the ethical development and deployment of AI technologies, ensuring they align with societal values and legal requirements.	(De Silva, T. M., & Alahakoon, 2022)

Regulatory Compliance	Adherence to laws, regulations, and guidelines that govern the deployment and use of AI systems, ensuring they meet established standards and protect public interests.	(Donnelly, 2022)
Human Oversight	The involvement of humans in supervising AI systems, particularly in critical areas such as decision-making, to prevent or mitigate harmful outcomes and ensure ethical use.	(Floridi & Cowls, 2019)
Stakeholders	Individuals, groups, or organizations that have an interest or stake in the outcomes of AI system implementations, including customers, employees, regulators, and the broader community.	(Floridi & Cowls, 2019)

1.8 Assumptions

In conducting this research, certain assumptions are established to frame the study and guide its methodology. These assumptions clarify the study's boundaries and ensure that the analysis remains relevant to the research objectives and background provided.

1. **Consistency in IT Practices:** The first assumption is that the IT practices observed at the major South African energy utility providers are consistent over time and representative of standard operations within the organization. This assumption allows the research to draw insights from current practices and identify trends in ethical AI implementation.

2. **Participant Honesty and Transparency:** This study assumes that the participants as stakeholders within the energy utility providers has offered honest and transparent responses during questionnaire. This assumption is critical to ensure the validity and reliability of the collected data.

3. **Existing Ethical Frameworks Apply to Energy Utility Providers:** It is assumed that the current ethical frameworks designed for AI implementation are relevant to the energy utility sector and are intended

to guide responsible AI adoption. This assumption enables the study to evaluate the adequacy of these frameworks in the context of a major energy utility provider.

4. **Ongoing or Imminent AI Implementation:** The study presumes that the energy utility providers are either already implementing AI or have plans to do so in the near future.

5. **Systems Concerns:** Given the critical nature of the services provided by the energy utility providers, this study assumes that systems and infrastructure are similar and essential to its operations. This assumption is crucial to focus on how AI adoption might affect the security and reliability of IT systems.

1.9 Chapter Outline

The following outline provides an overview of the structure and content of the thesis, allowing the reader to understand how the research has unfolded logically and systematically.

Chapter 2: Literature Review – The literature review critically examines existing research on ethical AI frameworks, transparency, accountability, bias, and fairness, providing a theoretical foundation for the study. This chapter also identifies gaps in the literature that the research aims to address.

Chapter 3: Research Methodology – This chapter details the research design, including the methods approach, data collection methods (e.g. Questionnaires), and analytical techniques used to answer the research questions.

Chapter 4: Results and Discussion – In this chapter, the findings from the data analysis are presented. The results are organized by research

question and proposition, with a focus on the adoption of ethical AI frameworks, transparency, accountability, and fairness. The results are discussed within the context of existing literature, highlighting similarities, differences, and new insights.

Chapter 5: Discussion – This chapter provides a detailed interpretation of the results, aligning the findings with existing literature while examining how they address the research questions. Additionally, the study evaluates the broader implications of the results. for ethical AI practices in South African utilities and beyond.

Chapter 6: Conclusion and Recommendations – The final chapter summarizes the key findings of the study, provides actionable recommendations for the utility sector, and suggests directions for future research in ethical AI.

CHAPTER 2. LITERATURE REVIEW AND THEORETICAL FRAMEWORK

2.1 Introduction

This chapter explores theoretical and conceptual frameworks, case studies, and challenges in AI ethics. This research utilizes the Technology-Organization-Environment (TOE) framework alongside the Ethical Decision-Making Model (EDM) to evaluate the preparedness of South African energy utilities for the ethical adoption of AI.

2.1.1 *Overview of AI in the Energy Sector*

Artificial Intelligence (AI) is increasingly becoming a cornerstone of the energy sector, enhancing operational efficiencies, grid optimization, and predictive maintenance (Aguilar, Díaz-Rodríguez, & Prieto, 2021; IEA, 2021). In applications such as smart metering, fault detection, and load forecasting, AI-driven systems are helping utilities reduce energy losses and improve reliability (Wang et al., 2020). However, as these technologies are integrated into critical infrastructure, ethical challenges—including transparency in decision-making, fairness in resource allocation, and institutional accountability—continue to persist (Floridi & Cowls, 2019; Jobin, Ienca, & Vayena, 2019; Xue et al., 2023). These concerns are particularly relevant in contexts with socio-economic disparities, where the unintended consequences of algorithmic decisions may disproportionately affect vulnerable populations. As such, robust and context-sensitive ethical frameworks are essential to govern the design, deployment, and oversight of AI applications in energy utilities.

As AI becomes increasingly integrated into the energy sector, it is likely that the conventional breach-specific model of preventing and handling privacy prejudices and situations will come under increased strains and may not be entirely ideal to the toolbox as AI commits cross-border data sharing and automatic judgments upon them (Korobenko et al., 2024). Moreover, many proposals are made that include or implicitly require AI to fully realize their potential, for example, see, e.g.

Gans .et al. (2019) for the potential of artificial intelligent (AI) in Canadian health sector. Given the scale and reach of potential AI impacts for public interest, it is little surprising that ethics in AI has become a research domain and main focus for many organizations. Ethical AI implies innovation and use of AI that are subject to human values and humane principles. Like with many other technological developments before (e.g. nuclear technology) choices about the acceptability and control of AI are now shaped by normative reflection.

It also has the prospective to deal more effectively with the respective onset of renewable energy resources, lead contrivance or group status diagnosis and enhance the stability and security of the facilities (Donnelly, 2022). A study on the use of AI and AI gauged IT repairs into Skaungwintern and Jarrowi is indicative of a potential fall in climate pollution and a significantly more equitable division of social profits (Isbanner et al., 2022). As AI gets rolled out into the energy sector, a sense of AI ethics should accompany the process.

Ethical guidelines and frameworks provides essential guidance for responsible AI implementation, while ensuring technologies continues aligning with societal values and provisioned human rights. In South Africa, energy utility providers often draw from both local and international standards to address the intricate ethical challenges associated with AI. According to Floridi and Cowls (2019), key principles in these frameworks include fairness, transparency, accountability, and privacy. These principles are crucial for preventing harm and promoting trust in AI systems. De Silva and Alahakoon (2022) emphasize that adherence to these guidelines is vital for compliance with both ethical norms and regulatory requirements. Furthermore, Jobin, Ienca, and Vayena (2019) highlight that adopting comprehensive ethical frameworks aids in reducing AI-related risks, including bias and privacy breaches, to promote the responsible application of AI technologies in the energy sector.

AI system biases have the potential to produce inequitable results, particularly in sectors like energy, where decisions can significantly impact consumers. Techniques such as bias detection algorithms, the use of diverse datasets, and continuous monitoring are employed to minimize biases (Clark & Valente, 2020). Isbanner, Möller, and Heinemann (2022) emphasizing the need for strong bias

detection and mitigation measures to enhance public confidence and prevent discriminatory effects.

Ensuring fairness in AI-driven decisions is critical, particularly in utilities where decisions can affect service delivery and pricing. According to Hagendorff (2022), fairness can be maintained through transparent decision-making processes, regular audits of AI systems, and stakeholder engagement. These practices help identify and rectify any biases or unfair treatment that may arise from AI applications. Diallo, Moyo, and Ntsimane (2024) emphasize the role of clear guidelines and accountability measures in maintaining fairness.

Ensuring accountability in AI is essential for maintaining transparency in decision-making and holding relevant entities responsible. Jobin, Ienca, and Vayena (2019) highlight that robust accountability mechanisms involve defining clear responsibilities, implementing audit trails, and establishing oversight committees. These approaches help enhance transparency in AI-driven decisions and facilitate prompt resolution of any negative consequences.

2.2 Definition of topic

AI systems in the energy sector play a role in predictive analytics, smart grid management, and energy distribution (McKinsey & Company, 2023). However, ethical challenges include:

- Bias in AI decision-making leading to unequal access to electricity.
- Lack of transparency in AI-powered pricing models (IEA, 2023).
- Concerns over accountability in AI-driven operational failures (Columbia University, 2024).

Effectively tackling these issues necessitates compliance with both global and region-specific ethical AI frameworks, ensuring equitable and accountable AI governance within the energy sector.

The integration of artificial intelligence (AI) within the energy sector represents a significant transformation in how utilities operate, optimize, and manage resources. As AI technology advances, it introduces new complexities and

challenges, particularly concerning ethical considerations. Korobenko et al. (2024) highlight that the conventional breach-specific models for handling privacy issues may become inadequate as AI facilitates cross-border data sharing and automated decision-making processes. This necessitates a re-evaluation of existing governance frameworks to ensure they address these emerging challenges.

AI technologies are becoming abundant across various industries, including energy, where they are leveraged for renewable energy optimization, carbon sequestration, nuclear power management, and waste reduction (Clark & Valente, 2020). The adoption of AI in these areas underscores its potential to revolutionize the energy sector by enhancing efficiency, predicting trends, and maintaining system balance, which are crucial for economic stability and environmental sustainability (Donnelly, 2022).

2.3 Factors Influencing Readiness for Ethical AI

The factors influencing the readiness of ethical AI implementation are explored, demonstrating how adherence to ethical frameworks, transparency and accountability, and measures to address bias and fairness influence the readiness of South African energy utility providers for ethical AI implementation. Understanding these determinations, is essential for ensuring that AI adoption not only enhances operational efficiency but also aligns with ethical standards. **The main objective for this research is:**

> "To identify and analyse if South African energy utility providers are ready for ethical AI implementation".

2.3.1 *Ethical AI Guidelines and Frameworks*

Sub- Objective 1: To determine which relevant ethical guidelines and frameworks are currently adopted by South African energy utility providers for AI implementation.

Ethical AI frameworks serves as guidelines and are essential for guaranteeing the ethical and responsible deployment of AI technologies. Frameworks such as the European Commission's Ethics Guidelines for Trustworthy AI and the OECD AI Principles highlight the significance of openness, equity, and responsibility in AI governance (OECD, 2021). However, literature highlights that South African energy utility providers lack a unified ethical AI framework, leading to fragmented adoption of ethical principles (Mhlambi, 2020). These frameworks typically include principles such as fairness, transparency, accountability, and privacy (Floridi & Cowls, 2019). In South African energy utility providers may integrate both domestic and global ethical frameworks to manage the challenges associated with AI adoption (De Silva & Alahakoon, 2022). The alignment with such frameworks ensures that AI systems are developed and deployed in a manner that respects human rights and societal values (Jobin, Ienca, & Vayena, 2019).

2.3.2 *Bias in AI Systems*

Sub Objective 2: To determine if necessary measures are in place to identify and mitigate biases in AI systems used by South African energy utility providers**.**

Bias in AI is a critical issue that can result in discriminatory outcomes, particularly in industries like energy, where AI is used for decision-making processes such as pricing, demand forecasting, and resource allocation (Hagendorff, 2022). Algorithmic bias often stems from skewed training data, inadequate representational models, and flawed decision-making logic (Mehrabi et al., 2021). South African energy utility providers face challenges in identifying and mitigating AI bias, as most existing bias detection tools are developed in Western contexts and may not effectively address local socio-economic disparities (Diallo, Moyo, & Ntsimane, 2024). To ensure fairness, industry experts recommend the integration of continuous bias audits, the use of diverse datasets, and regulatory oversight mechanisms (Binns, 2018).

Identifying and mitigating biases in AI systems which ensures fairness and equity is essential. Techniques such as bias detection algorithms, diverse data sets, and continuous monitoring are employed to reduce biases (Clark & Valente, 2020). South African energy utility providers must implement robust bias detection and mitigation strategies to foster trust and avoid discriminatory outcomes (Isbanner, Möller, & Heinemann, 2022).

2.3.3 Fairness in AI Driven Decisions

Sub Objective 3: To assess how energy utility providers ensure fairness in AI-driven decisions that affect customers and stakeholders.

Fairness in AI-driven decisions involves ensuring that AI outcomes do not disproportionately impact any group of customers or stakeholders. This can be achieved through transparent decision-making processes, stakeholder engagement, and regular audits of AI systems (Hagendorff, 2022). Energy utility providers must establish clear guidelines and accountability measures to uphold fairness in their AI applications (Diallo, Moyo, & Ntsimane, 2024).

2.3.4 Mechanisms for Accountability

Sub Objective 4: To evaluate if there are mechanisms to ensure accountability for AI-driven decisions and their consequences***.***

Accountability is essential to ethical AI implementation, ensuring that organizations take responsibility for AI-driven outcomes. Accountability mechanisms include establishing clear lines of responsibility, implementing audit trails, and setting up oversight committees to review AI decisions and their impacts (Jobin, Ienca, & Vayena, 2019). Energy utility providers need to create robust governance frameworks to ensure that AI-driven decisions are accountable and that any adverse outcomes are promptly addressed (Cherryholmes, 1992).

2.3.5 *Proposition*

"Prominent South African energy utility providers might be inadequately prepared for AI implementation with regard to ethical considerations due to insufficient adherence to ethical frameworks, lack of transparency and accountability and inadequate measures to address bias and fairness."

2.4 Examining Case Studies and Leading Practices in Ethical AI Deployment.

a) **International case studies**

AI in European Energy Management**:** Despite efficiency gains, ethical concerns arose over unequal energy pricing models (World Economic Forum, 2025).

b) **South African Case Study**

Eskom's AI Initiatives**:** AI-powered grid optimization improves reliability, but concerns over algorithmic transparency and bias persist (Centre for Trustworthy Technology, 2024).

2.5 Challenges pertaining Implementation Ethical AI in Energy Sector

a) **Bias and Fairness Issues**

AI systems may perpetuate socio-economic disparities in energy allocation due to biased training data (Numalis, 2024).

b) **Transparency and Accountability Gaps**

Many AI-driven decisions in energy pricing and allocation function as “black-box” systems, making accountability difficult (Public Power Association, 2024).

c) **Regulatory and Compliance Gaps**

South Africa lacks a standardized AI ethics governance policy, leading to inconsistent adoption of ethical AI principles (U.S. Department of Energy, 2024).

d) Data Privacy and cybersecurity Risks

Cybersecurity risks pose a significant threat to AI-powered energy management systems (Centre for Trustworthy Technology, 2024).

Research Gaps and Justification for the Study

Despite global advances in AI governance, South Africa's energy sector lacks localized AI ethics research. Identified gaps include:

- Limited research on AI ethics in South African utilities (Xue et al., 2023).
- No standardized AI governance framework (Centre for Trustworthy Technology, 2024).
- Lack of empirical studies on AI fairness and bias in energy pricing models (Columbia University, 2024).

This study fills these gaps by providing a policy-focused assessment of AI ethics readiness in South Africa's energy utilities.

2.5 ANALYTICAL FRAMEWORK

2.5.1 *Theoretical Framework*

In order to provide a comprehensive understanding of the dominant factors influencing the readiness for ethical AI implementation, this study have considered two theoretical frameworks: Technology-Organization-Environment (TOE) framework and Ethical Decision-Making models. Each framework offers unique insights into the interplay between technology adoption, organizational behaviour, and ethical considerations.

1. Technology-Organization-Environment (TOE) Framework: The TOE framework emphasizes the role of technological characteristics, organizational resources and structure, and environmental factors in technology adoption. It highlights the critical role of regulatory mandates and industry benchmarks in ensuring AI adoption adheres to ethical standards (Tornatzky &

Fleischer, 1990). Furthermore, Benjamins (2020) introduced organizational directives for the ethical deployment of AI, stressing the importance of integrating ethical principles and methodologies into business operations.

2. Ethical Decision-Making Model: Provides a framework for understanding how ethical decisions are made within organizations. They emphasize the role of individual characteristics, such as personal values and moral development, and situational factors, such as organizational culture and ethical climate, in shaping ethical decision-making processes. For instance, a 2024 review highlights the profound impact of ethical culture on decision-making, suggesting that organizational culture significantly shapes ethical behaviour among managers and employees (Kaptein, 2024)

The two above frameworks, TOE and Ethical decision making models are value adding to the study as a whole. As a results this study focused only on the combination of this two frameworks and their application. The combination of Ethical Decision-Making Models and the Technology-Organization-Environment (TOE) Framework is justified by their complementary perspectives, which enable a comprehensive analysis of the readiness of South African energy utility providers to adopt ethical AI systems. The integration ensures that both the structural, contextual factors influencing AI implementation and the ethical considerations underpinning decision-making are adequately addressed.

2.5.2 **Theoretical Framework Adopted for the Study**

The following theoretical frameworks informed the development of the conceptual framework of this study”

a. Technology-Organization-Environment (TOE) Framework

The TOE framework provides a structured perspective on AI adoption, considering:

- Technological factors (AI capabilities, cybersecurity, and data privacy).
- Organizational factors (leadership, workforce readiness, and ethical policies).
- Environmental factors (regulatory compliance, customer expectations) (Tornatzky & Fleischer, 1990).

This framework helps assess how AI ethics are integrated within South African energy providers. It also provides a systemic understanding of the macro-level factors that affect AI adoption. The technological dimension assesses AI's compatibility and complexity; the organizational dimension evaluates cultural and resource readiness; and the environmental dimension explores external pressures like regulatory requirements (Tornatzky & Fleischer, 1990). However, it does not address the human and ethical nuances of decision-making processes.

b. Ethical Decision-Making (EDM) Model

- The EDM model (Rest, 1986) focuses on ethical challenges in AI adoption:
- Moral Awareness (recognizing AI bias and fairness issues).
- Moral Judgment (ensuring AI-driven decisions align with ethical principles).
- Moral Intent & Behavior (establishing AI accountability mechanisms).

The **TOE** and **EDM** models together provide a dual-layered assessment of ethical AI readiness. They delve into the micro-level ethical considerations of AI adoption, focusing on how decisions are influenced by individual and organizational values, ethical cultures, and the situational context (Trevino, 1986). However, they lack the structural context provided by the TOE framework. Also, Hochstein (2024) proposes a theoretical framework to guide AI ethical decision-making, emphasizing the need for principles that establish boundaries for algorithmic intelligence. This model parallels individual decision-making processes emphasizing the necessity of ensuring AI systems adhere to human ethical principles.

Additionally, Mokander et al. (2021) discuss ethics-based auditing of automated decision-making systems, addressing the ethical challenges posed by delegating tasks to such systems. They emphasize the need for new governance mechanisms to ensure that organizations design and deploy AI in ways that are ethical, thereby enabling society to reap the full benefits of automation.

2.5.3 *Conceptual Framework*

Figure 1 : TOE Framework

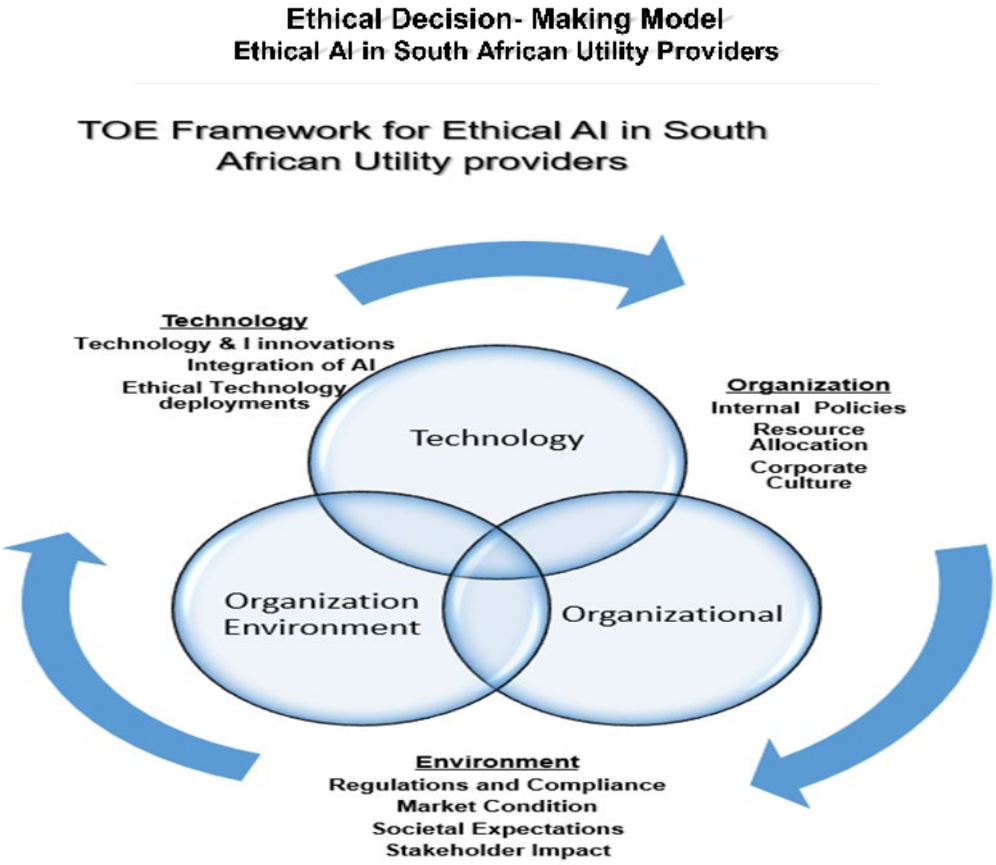
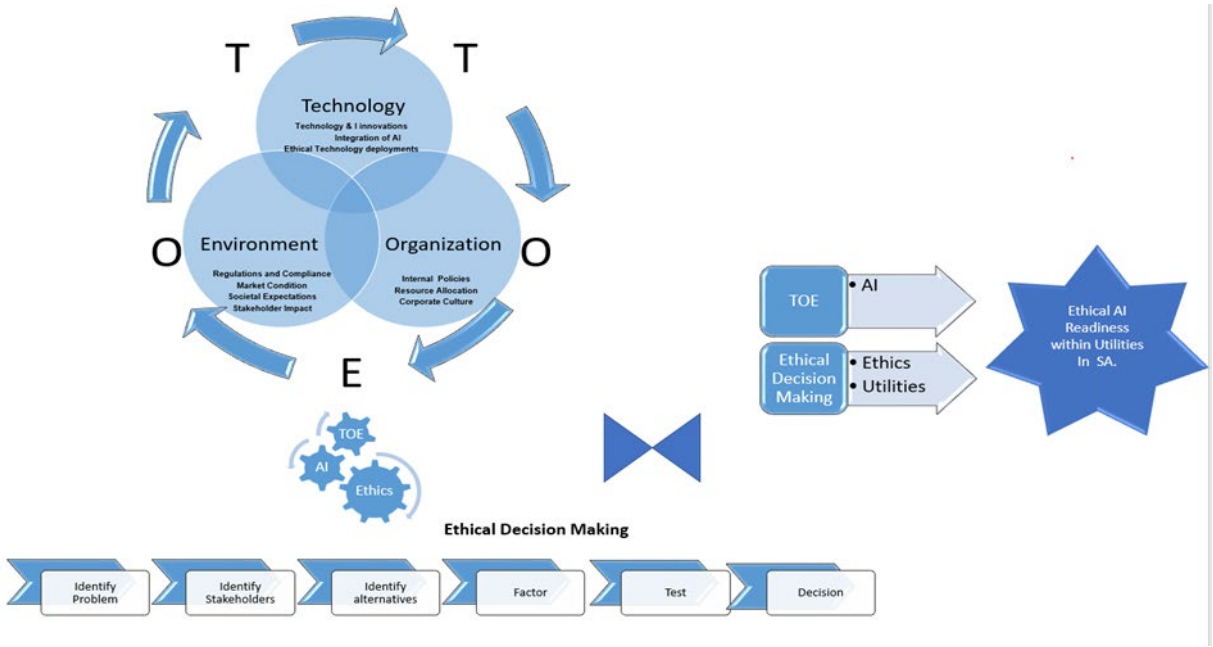


Figure 2 : Ethical Decision –Making Model



Figure 3 : Amalgamation of TOE and Ethical decision making



Explanation of the ultimate amalgamation of the two Frameworks

The Technology, Organization and Environment model provides a systemic understanding of the macro-level factors that affect AI adoption (*See figure 1*). The technological aspect evaluates AI's integration feasibility and complexity; the

organizational dimension evaluates cultural and resource readiness; and the environmental dimension explores external pressures like regulatory requirements. This ensure that relevant AI ethical frameworks which could be from industries, national or /and international are adopted, and are applied throughout the three dimension (TOE). Should there be a need to either amend or develop/generate an AI ethical decision, then the Ethical Decision Making frameworks (*See Figure 2*) will be followed and be used together with TOE dimensions in order to get ethically sound AI outcomes for the energy utility providers. So, the energy provider entities needs to have a capability to develop and implement ethical decisions as ethical requirements evolves with time.

For South African energy utility providers to effectively integrate AI technologies, a thorough grasp of the associated ethical considerations is essential. As depicted on above on figure 3, The Technology-Organization-Environment (TOE) framework and Ethical Decision-Making Models offer essential insights into the key elements influencing the sector's readiness for ethical AI integration. The TOE framework emphasizes the role of technological, organizational, and environmental factors in shaping AI adoption (Tornatzky & Fleischer, 1990). Ethical Decision-Making Models focus on the processes through which ethical decisions are made, highlighting the importance of individual and organizational values (Rest, 1986).

Upon considering the relevant literature, and identifying most relevant frameworks which are the TOE, DOI, Institutional Theory, Ethical Decision-Making Models, and Corporate Social Responsibility (CSR), it has become apparent that there is a need for incorporating multiple frameworks. By integrating these frameworks, the study bridges the gap between systemic readiness and ethical implementation, facilitating a comprehensive grasp of the elements influencing ethical AI implementation.

The following two are instrumental in achieving the objectives of this study. The chosen two are:

1. **Technology-Organization-Environment (TOE) Framework (*See Figure 1*)**

2. Ethical Decision-Making Models (*See Figure 2*)

The combination of the TOE framework and Ethical Decision-Making Models offers a comprehensive approach to understanding and addressing the factors influencing the readiness of South African energy utility providers for ethical AI implementation. This contributes to each research objectives discussed. Recent studies emphasize. The TOE framework's utility in analysing organizational readiness for AI. For instance, Satyro et al. (2024) expand the framework to include sustainability, demonstrating its adaptability for analysing complex, multidimensional phenomena like ethical AI adoption. It presents a broad perspective on technological, organizational, and environmental influences, while Ethical Decision-Making Models deliver a detailed analysis of the ethical frameworks and decision-making principles governing AI deployment.

It further underscores that both global and local standards serve as a foundation for ethical AI adoption, enabling energy utility providers to align their practices with industry best standards (Jobin, Ienca, & Vayena, 2019; Floridi & Cowls, 2019). Addressing biases in AI systems is essential to uphold fairness and prevent discriminatory consequences. Effective bias identification and mitigation techniques, including algorithmic audits and the use of diverse datasets, play a crucial role in ensuring equitable AI outcomes (Mehrabi et al., 2021; Hagendorff, 2022).

Achieving fairness in AI-based decision-making requires the development of transparent and accountable systems that prioritize stakeholder interests. Mechanisms such as fairness audits and ethical reviews help in maintaining equity in AI operations (Binns, 2018). Accountability mechanisms, including transparent reporting and independent oversight, are essential for monitoring and addressing the consequences of AI-driven decisions (Floridi & Cowls, 2019). Together, these frameworks and measures provide a comprehensive approach to understanding and addressing the factors influencing the readiness of South African energy utility providers for ethical AI implementation.

Together, the above frameworks enables a thorough exploration of the research objectives and Secondary objectives, providing a solid foundation for identifying

the dominant factors influencing ethical AI adoption and ensuring that the study's objectives are met.

The TOE and Ethical decision making model theoretical frameworks collectively argue that South African energy utility providers face significant challenges in ethical AI implementation due to gaps in ethical guidelines, issues with bias and fairness and lack of transparency and accountability. Addressing these challenges is essential for ensuring ethical AI adoption in the energy utility sector.

The research objectives investigates the readiness of prominent South African energy utility providers for AI implementation with regard to ethical considerations, amidst escalating AI adoption. The proposition suggests that these providers are inadequately prepared due to several factors, which can be understood through various theoretical frameworks.

1. **Technology-Organisation-Environment (TOE) Framework**

The Relevance to Research objectives:

- **Main Objective:** The TOE framework helps identify and asses' energy utility provider's readiness for AI implementation by examining technological characteristics, organizational capabilities, and environmental factors such as regulatory pressures and industry standards. This holistic approach ensures that all critical aspects impacting readiness are considered.
- **Sub-Objective 1 (Ethical Guidelines and Frameworks):** TOE's focus on environmental factors, including regulatory and industry standards, directly addresses the ethical guidelines and frameworks adopted by energy utility providers.
- **Sub-Objective 2 (Mitigating Biases):** TOE examines technological characteristics and organizational resources, offering an understanding of the strategies implemented to detect and address biases in AI systems.

- **Sub-Objective 3 (Ensuring Fairness**): TOE's consideration of organizational factors and environmental pressures aids in understanding how fairness is ensured in AI-driven decisions.

- **Sub-Objective 4 (Accountability Mechanisms):** TOE's focus on organizational processes and environmental regulations helps identify mechanisms ensuring accountability for AI-driven decisions.

2. **Ethical Decision-Making Model:**

Relevance to Research Objectives:

- **Main Objectives:** Ethical Decision-Making Models focus on the processes through which ethical decisions are made within organizations, directly addressing the readiness for ethical AI implementation.

- **Sub-Objective 1 (Ethical Guidelines and Frameworks):** These models highlight the importance of ethical frameworks and individual and organizational values in guiding AI implementation, ensuring alignment with adopted guidelines.

- **Sub-Objective 2 (Mitigating Biases):** Ethical Decision-Making Models provide a basis for understanding how biases are identified and mitigated by emphasizing ethical analysis and decision processes.

- **Sub-Objective 3 (Ensuring Fairness):** The focus on ethical principles and organizational culture within these models explains how fairness is integrated into AI-driven decisions.

- **Sub-Objective 4 (Accountability Mechanisms):** These models underscore the importance of accountability mechanisms by stressing ethical responsibility and the consequences of decisions, ensuring transparency and accountability.

The blend of the two above adopted frameworks enable a good study to understand the gabs in the factors influencing the readiness of South African energy utility providers for ethical AI implementation. The qualitative research design enables provision of in-depth insights in to issues of interest, capturing the perspective and experiences of relevant stakeholders.

2.6 Conclusion of Literature Review

The reviewed literature underscores a growing global and local interest in leveraging AI within the energy sector, yet highlights a persistent gap in addressing ethical implications within implementation frameworks. While previous studies have identified core ethical concerns—such as fairness, accountability, and transparency—there remains limited empirical focus on how these are operationalized in South Africa's energy utilities context. This gap necessitates a context-specific investigation into ethical AI readiness.

Grounded in established theories of AI ethics, the theoretical framework for this study supports a structured evaluation of ethical preparedness, while the conceptual framework delineates the key dimensions—bias mitigation, transparency, fairness, and accountability—relevant to utility providers. This study thus contributes by offering evidence-based insights into the ethical integration of AI in South African energy utilities, with the aim of informing both governance strategies and practical implementation.

CHAPTER 3. RESEARCH METHODOLOGY

3.1 Research approach

The research paradigm guiding this study is Interpretivism, which is rooted in the idea that reality is socially constructed and best understood through the subjective experiences and meanings individuals ascribe to their world (Bryman, 2016). Interpretivism emphasizes understanding the complex world of lived experience from the perspective of those who live it, making it particularly suitable for exploring human-centred phenomena such as the adoption of ethical AI systems. This paradigm allows researchers to delve into the diverse and nuanced experiences, beliefs, and values of stakeholders, offering insights into how they perceive and navigate the difficulties and opportunities connected to AI implementations (Creswell & Poth, 2018).

In the context of this study, Interpretivism provides a framework to examine the readiness of South African energy utility providers for AI adoption, particularly in relation to ethical considerations. By prioritizing the subjective perspectives of key stakeholders—such as executives, managers, IT professionals, ethics officers, and potent customers. This approach enables a deeper understanding of how organizational culture, regulatory pressures, and technological complexities shape their attitudes and preparedness for ethical AI deployment. This paradigm is especially relevant for addressing issues that require understanding multiple viewpoints, as the ethical implementation of AI is inherently influenced by varying individual, organizational, and societal values (Patton, 2015).

Furthermore, Interpretivism aligns with the qualitative methods employed in this study, such as semi-structured interviews and thematic analysis, which aim to capture the rich, context-specific insights needed to address the research objectives. This paradigm's emphasis on context and meaning is vital for identifying the barriers and enablers of ethical AI readiness within energy utility providers, thus contributing to actionable and socially responsible recommendations (Flick, 2020). (Bryman, 2016).

Qualitative Methods Approach

The research approach for this study is a qualitative method approach. This approach allows for an exploration of the research objectives, leveraging the advantages of qualitative data to gain deeper insights into the ethical preparedness of energy utility providers for AI adoption (Creswell & Plano Clark, 2018).

Within the scope of the research objective, “To identify and analyse if South African energy utility providers are ready for ethical AI implementations." focusing solely on qualitative methods can be both appropriate and beneficial. Highlighted below is the justification and alignment to research objectives:

- The ethical readiness for AI implementation involves complex, nuanced, and context-specific issues. Qualitative methods allow for a deeper exploration of these complexities by facilitating an in-depth understanding of participants' experiences and perceptions (Evans et al., 2023).
- Participants can offer in-depth, descriptive insights into their experiences, perspectives, and the ethical challenges they encounter, which quantitative methods may not capture adequately. Open-ended questions, for example, enable respondents to express their thoughts in their own words, offering insights that structured questions might miss (Amberscript, 2023).
- The qualitative method, particularly through open-ended questionnaires, allows for flexibility in data collection. This flexibility is advantageous when exploring emerging themes or unanticipated issues that may arise during the study, capturing the complexity of human experiences and opinions (Restack, 2023).
- Qualitative research can capture the cultural and organizational context within which AI implementation occurs, including specific practices, norms, and values that influence ethical considerations. Understanding these contexts is crucial for addressing the ethical implications of AI-driven systems in healthcare and other sectors (Evans et al., 2024).

3.2 Research design

The research design is based on a case study design, involving multiple case studies of prominent South African energy utility providers. This design allows for an in-depth, contextual analysis of provider's AI implementation and ethical practices.

Justification for qualitative method design

A qualitative design is best suited to capture the complex and context-specific nature of ethical decision-making processes within utility providers. Given the limited empirical research on this topic in South Africa, a qualitative inquiry enables the researcher to generate rich, descriptive data that reflects real-world practices and stakeholder concerns. This approach supports the study's aim of developing a grounded understanding of current readiness levels and informing ethically aligned AI governance strategies.

Advantages and disadvantages of adopting a qualitative method design

Table 2: Advantage and disadvantage of qualitative method

Advantage	Disadvantage
In-depth Exploration: The case study approach allows for a comprehensive examination of each energy utility provider's unique context, strategies, and challenges related to AI ethics. This method is particularly effective for understanding complex, context-specific phenomena (Yin, 2018).	**Time and Resource Intensive:** Conducting comprehensive qualitative case studies can be time-consuming and require significant resources, potentially limiting the research's scope and feasibility (Merriam & Tisdell, 2016).

Flexibility: The qualitative case study approach is adaptable, enabling researchers to modify their focus based on emerging findings and delve deeper into specific areas of interest, which enhances the richness of the data collected (Stake, 1995).	**Subjectivity and Researcher Bias:** Qualitative research is inherently subjective, and there is a risk of researcher bias influencing data collection and analysis, which may impact the validity and reliability of findings (Maxwell, 2013).
Rich Data Collection: Qualitative methods facilitate the collection of detailed data that provide insights into participants' perspectives, experiences, and decision-making processes, capturing the nuanced complexity of human behavior (Creswell & Poth, 2018).	**Limited Generalizability:** Findings from qualitative case studies may have limited applicability to other contexts or populations due to the specific focus on certain organizations and stakeholders, restricting the broader application of results (Yin, 2018).

3.3 Data collection methods

Including a questionnaire in the data collection enhances the research by providing a dimension that complements and validates the qualitative insights.

3.4 Questionnaire

This approach ensures a balanced, and robust exploration of the ethical readiness of South African energy utility providers for AI implementation. The qualitative methods facilitate a comprehensive exploration of the research objectives from diverse viewpoints (Kvale & Brinkmann, 2009). Structured questionnaires was distributed to a sample of stakeholders across different organisation’s levels, including energy utility provider’s employees, executives, and IT regulatory bodies, etc. for qualitative insights.

3.5 Population and sample

Questionnaires serve as an efficient tool for gathering data from diverse samples, enabling the collection of valuable insights and offering a comprehensive overview of the research area. In this study, questionnaire enabled the collection of data on AI readiness indicators, such as stakeholders' perceptions on AI ethical framework, attitudes, and experiences regarding AI implementations, biases and fairness handling, and transparency and accountability considerations.

3.5.1 *Population*

Questionnaire:

Sample: The participants of interest for the questionnaire are individuals working within the South African energy utilities, including executives, management, IT professionals, and ethics officers and customer.

Table 3 : Profiles of respondents

Respondent Category	Position	Context
Executives	CEO, CTO, CIO	Strategic decision-making, AI policy development, oversight
Manager	IT Manager, Operations Managers	Implementation of AI projects, compliance with ethical guidelines
Employees	Engineers, Data Scientists	Day-to-day use and development of AI Technologies
AI Ethics Experts	Academic Researchers, Consultants	Expert insights on AI ethics, best practices, and frameworks
Customers	Residential Users, Business Clients	End-users affected by AI applications in utility services

3.5.2 *Sample*

The study utilizes purposive sampling, a non-probabilistic sampling technique designed to deliberately select approximately 15 participants who are most knowledgeable and relevant to the research objectives. This approach aligns with

qualitative research methodology and the Interpretivism paradigm adopted in the study.

3.5.3 *Sampling method*

The study targeted about 16 participants within South African energy utility providers. The sample includes individuals across various organizational levels, such as executives, management, IT professionals, ethics officers and customers. This range ensures a comprehensive understanding of ethical AI readiness from multiple stakeholder perspectives.

Justification of Sampling Method

1. **Alignment with Research Objectives**: The research seeks to explore nuanced, context-specific insights into the ethical readiness of energy utility providers for AI implementation. Purposive sampling allows the inclusion of participants directly involved with AI initiatives, ensuring their responses are rich and relevant.
2. **Expertise Focus:** Selecting participants with direct experience in AI decision-making, policy formulation, or ethical governance ensures the data collected is meaningful and directly applicable to the research objectives.
3. **Feasibility and Depth:** A targeted sample size allows for detailed, in-depth data collection, appropriate for qualitative research. This approach supports the thematic analysis method planned for this study.

Demographic Profile of the Sample

- Industry: Utilities within South Africa.
- Roles *(See Table 3)*:
    - Senior executives (e.g., CEOs, CTOs) to provide strategic insights on AI implementation.
    - Middle management to discuss operationalization and adherence to ethical guidelines.

- IT professionals to explore technical aspects, including bias mitigation and transparency mechanisms.
- Ethics officers or compliance professionals to focus on governance and regulatory perspectives.
- Customer's representation or individual's involved in or knowledgeable about AI integration processes within their organizations.

3.6 The research instrument

Questionnaire Instrument: The questionnaire instrument is designed to evaluate AI readiness in energy utility providers, focusing on ethical considerations such as fairness, transparency, and accountability. To ensure comprehensive data collection, the questionnaire included a mix of Likert-scale and open-ended questions, allowing for both structured responses and in-depth qualitative insights.

Recent studies have emphasized the effectiveness of Likert-scale questions in assessing ethical perceptions of AI systems, particularly in sectors where transparency and accountability are critical (Dardiry et al., 2024). Furthermore, open-ended questions play a vital role in capturing nuanced insights into ethical concerns and AI governance frameworks (Naeem et al., 2023). This approach aligns with thematic analysis methodologies, which are widely used in AI ethics research to identify key themes and patterns in qualitative data (Naeem et al., 2023).

Additionally, the questionnaire was structured to align with the research objectives, literature, and theoretical frameworks, particularly addressing core elements such as ethical AI frameworks, bias mitigation strategies, fairness in AI-driven decision-making, and mechanisms for ensuring transparency and accountability. This approach ensures that the instrument remains both methodologically sound and contextually relevant to the study of ethical AI adoption in South Africa's energy utilities.

Alignment with Research Objectives, Literature, and Theoretical Framework

1. **Ethical Guidelines and Frameworks**
   Relevant objectives and questions are designed to evaluate the current adoption of ethical guidelines by energy utility providers and their alignment with global standards. This addresses the research objective on the sufficiency of ethical frameworks, informed by literature emphasizing the importance of guidelines for fairness, accountability, and transparency (Floridi & Cowls, 2019).
   Example Objective: "To determine which ethical guidelines and frameworks are currently used in your organization for AI implementation? How effective are they?"

2. **Bias and Fairness**
   This section includes questions to explore measures for identifying and mitigating biases, ensuring alignment with the ethical decision-making model. This directly addresses the objective on mitigating bias (Hagendorff, 2022).
   Example Question: "What processes or tools does your organization use to identify and mitigate biases in AI systems?"

3. **Transparency and Accountability**
   The questions evaluate the measures in place to uphold accountability and transparency in AI-driven decision-making, aligning with the TOE framework's focus on organizational and environmental influences (Tornatzky & Fleischer, 1990).
   Example Question: "How does your organization ensure transparency in AI-driven decisions that impact stakeholders?"

4. **Fairness in AI Decision-Making**
   Questions target fairness in decision-making processes, consistent with the literature on ethical AI practices ensuring equitable treatment (Jobin, Ienca, & Vayena, 2019).
   Example Question: "Can you provide examples of how fairness in AI decision-making is evaluated in your organization?"

3.7 Procedure for data collection

Questionnaire: Data collection was conducted through the electronic administration of a structured questionnaire, distributed via a secure online link to the selected sample of participants. Participants were invited to complete the questionnaire at their convenience through a digital platform designed to facilitate ease of access and response. In accordance with the ethical standards prescribed by the Wits Business School (WBS), an informed consent section was embedded within the questionnaire. This section clearly outlined the purpose of the study, confidentiality assurances, and participant rights, ensuring compliance with institutional ethical guidelines.

3.8 Data analysis strategies and interpretation

Questionnaire Data

This study utilizes a qualitative research design with data collected via semi-structured questionnaires distributed to relevant stakeholders in South African energy utility providers. The data analysis employed thematic analysis, a robust method for identifying, analysing, and reporting patterns (themes) within qualitative data.

Thematic analysis was conducted following Braun and Clarke's (2006) six-phase framework which includes: Familiarisation with the data, generating initial codes, searching the theme, reviewing the themes, defining and naming the themes, and writing up. *Also see Table 4 below by Naeem et al. (2023)*

Thematic Analysis Process

Table 4 : THEMATIC ANALYSIS PROCESS Source: Adapted from Naeem et al. (2023)

Phases	Description
Step 1: Data Transcription and Initial Review	In this phase, collected data is transcribed, and researchers thoroughly examine it to gain familiarity. Key segments are identified for further analysis.
Step 2: Identifying Key Terms and Patterns	The data is carefully assessed, whether from interviews, focus groups, or other qualitative sources, to detect frequently occurring terms and emerging themes.
Step 3: Assigning Codes to Data	The information is systematically labelled with short descriptive codes that capture the essential meaning of each data segment, ensuring structured analysis.
Step 4: Categorization and Theme Identification	The identified codes are organized into broader categories, allowing researchers to recognize common patterns and key insights relevant to the study.
Step 5: Interpretation and Concept Development	Researchers further analyse the established categories, refining their meaning and extracting theoretical connections between concepts.
Step 6: Constructing a Conceptual Model	In the final phase, the insights gained through thematic analysis are synthesized into a structured model that reflects the relationships between key themes, guided by relevant theories.

Data

The Interpretation of the data involves making sense of the themes in relation to the research objective:" To identify and analyse if South African energy utility providers' are ready for ethical AI implementation".

1. Ethical Guidelines and Frameworks
    - Analysis: Identify themes related to the adoption and implementation of ethical guidelines and frameworks among energy utility providers.
    - Interpretation: Evaluate how these guidelines compare to international standards and their effectiveness in guiding ethical AI practices.

2. Bias and Fairness

- Analysis: Recognize patterns associated with strategies used to detect and minimize biases in AI systems.
- Interpretation: Assess effectiveness of these measures and how they ensure fairness in AI-driven decisions affecting customers and stakeholders.

3. Transparency and Accountability
   - Analysis: Identify themes related to the transparency of AI processes and the mechanisms for ensuring accountability.
   - Interpretation: Evaluate the adequacy of these mechanisms and their role in ensuring responsible AI deployment.

Literature

The literature is synthesized and analysed thematically to extract key findings and insights relevant to the research objectives. Themes and patterns identified in the literature were compared and contrasted with the findings from the questionnaire to enrich the analysis.

Justification

The chosen populations and sampling methods are justified based on their alignment with the research objectives. This approach did offer a qualitative insights into ethical considerations, and the literature review has contextualize the findings within existing scholarship. By employing this method and sampling techniques, the study aims to ensure the validity and reliability of the research outcomes.

3.9 Possible limitations and challenges of the study

The limitations which identifies and anticipates potential weaknesses in this study are:

- Sampling Bias: There may be a risk of sampling bias, as participation is voluntary and may attract individuals with specific perspectives or experiences.

- Generalizability: The findings of the study may not be generalizable to all energy utilities in South Africa or other countries, given the focus on a single case study and the specific context of the research.

- Time and Resource Constraints: Limited time and resources may restrict the scope and depth of data collection and analysis, potentially limiting the comprehensiveness of the study.

- Ethical Considerations: Upholding ethical standards throughout the research process by safeguarding informed consent, maintaining confidentiality, and ensuring data protection, may present challenges, particularly in sensitive areas such as AI ethics and legal frameworks.

- Response Bias: There's a possibility of response bias in the questionnaire, where participants may provide socially desirable responses rather than their true opinions or experiences, affecting the validity of the data.

- Data Quality: The quality of data collected through questionnaire may vary depending on respondents' understanding of the questions or their willingness to provide accurate information, impacting the reliability of the findings.

- Sample Representativeness: While efforts are made to ensure a diverse sample, there's a possibility that the selected participants may not fully represent the broader population of energy utility stakeholders, potentially limiting the generalizability of the findings.

- Interpretation Challenges: Analysing qualitative data, May involves subjective interpretation, which may lead to different conclusions drawn by different researchers, affecting the reliability of the qualitative findings.

- Limited Access to Data: Access to proprietary or sensitive data within the energy utilities sector may be restricted, limiting the researcher's ability to fully explore certain aspects of AI ethical readiness or legal frameworks.

- Cross-Cultural Considerations: Given the multicultural nature of South Africa, cultural differences among participants may influence their

perceptions and attitudes towards AI ethical considerations, complicating the interpretation of the findings.

- Temporal Constraints: The study's focus on a specific time frame may limit its ability to capture long-term trends or changes in AI adoption and ethical readiness within the energy utilities sector.

3.10 Quality Assurance

In order to ensure quality assurance in the research on the readiness of South African energy utility providers for ethical AI implementation, the following measures considered for external validity, internal validity, confirmability, and reliability:

3.10.1 ***Transferability***

External validity refers to the extent to which the results of a study can be generalized to other settings, populations, and times (Shenton, 2004).

A clear description of South African energy utility providers, including their size, scope of operations, and the AI technologies they are adopting. This detailed context helps readers understand the setting and assess whether the findings can be applied to other similar settings.

3.10.2 ***Credibility***

Internal validity refers to the extent to which the research findings accurately represent the phenomenon being studied (Shenton, 2004).

The data collection method as questionnaires from different stakeholders enables the cross-check the information obtained as different population groups.

3.10.3 *Dependability*

Dependability refers to the consistency of the research findings if the study were to be repeated under similar conditions (Shenton, 2004).

The use of standardized procedures for data collection and analysis. Ensure that questionnaire instrument is used consistently across all participants from different levels.

3.11 Ethical considerations

Ensuring the ethical conduct of research is paramount to uphold the rights and well-being of participants and maintain the integrity of the study. The following measures are implemented to ensure that the research adheres to ethical standards:

- **Informed Consent**: Prior to participation in the study, all participants were provided with detailed information about the research objectives, procedures, potential risks, and benefits. Informed consent forms were distributed, and participants were given adequate time to review the information and ask questions before providing their consent.
- **Confidentiality:** Confidentiality of participant information was strictly maintained throughout the research process. All collected data, including survey responses, interview transcripts, and related documents, are anonymized and securely stored. Access to identifiable information is restricted to authorized research team members, ensuring that the data is used exclusively for research purposes.
- **Voluntary Participation:** Participation in the study was entirely voluntary, and participants had the right to withdraw at any time without consequence. No coercion or undue influence was exerted to encourage participation, and participants were assured that their decision to participate or withdraw would not affect their relationship with the university or any other entity.
- **Data Security:** Measures were implemented to ensure the security and integrity of data collected during the study. Electronic data is stored on password-protected devices and encrypted where applicable. Hard copies of

data are stored in locked filing cabinets accessible only to authorized personnel.

- **Minimization of Harm:** Efforts are made to minimize any potential harm or discomfort to participants. This includes ensuring that survey questions and interview protocols are sensitive and non-invasive, providing adequate support resources for participants who may experience distress, and swiftly responding to any issues or complaints brought forward by participants.
- **Ethical Review:** The research protocol, including details of the study design, data collection methods, and ethical considerations, were submitted to the Wits University Post-Graduate Committee (PGC) for ethical review and approval. Any modifications to the research protocol was reported to the PGC for review and approval.
- **Researcher Integrity:** Researchers involved in the study has adhere to the highest standards of professional conduct and integrity. They do respect the autonomy and dignity of participants, maintain objectivity in data collection and analysis, and avoid conflicts of interest that could compromise the validity and credibility of the research.

By incorporating these ethical considerations, the study seeks to uphold the principles of respect, beneficence, justice, and integrity in every aspect of the research process.

CHAPTER 4. PRESENTATION OF FINDINGS

4.1 Introduction

This chapter presents the key findings of the study on the ethical readiness of South African utility providers for artificial intelligence (AI) implementation. In alignment with the research objectives, the analysis draws on responses from a diverse group of participants—including utility managers, IT professionals, legal and ethics officers, and customers—providing insight into how ethical principles such as fairness, transparency, accountability, and bias mitigation are currently understood and operationalised within the sector. Given the increasing influence of AI technologies across the energy domain, understanding the ethical preparedness of these providers is essential for ensuring responsible and equitable innovation (Floridi & Cowls, 2019).

The findings are organised according to the study's core propositions: (1) the adoption of ethical frameworks and guidelines, (2) transparency and accountability mechanisms, (3) bias and fairness considerations, and (4) ethical perspectives shaping future directions. Each thematic section includes supporting evidence from the qualitative data, complemented by tables, figures, and illustrative participant quotations. This structure allows for a comprehensive yet focused analysis of the utility sector's current practices, challenges, and gaps, thereby contributing to a broader understanding of ethical AI governance in South Africa's energy landscape.

4.2 Results pertaining to Proposition 1

Summarized presentation of results based on the study on Ethical AI implementation readiness within South African utilities, organized into relevant categories. The study used tables to represent the data, and also included descriptions of each to make it clear on how they connect to the propositions in the research.

Use and context of Verbatim Responses

Verbatim quotations are presented throughout this chapter to support and illustrate the key themes that emerged from the qualitative analysis. The selected responses were purposefully chosen based on their representativeness of the broader sentiments expressed across the participant pool. In each case, the quotations reflect the dominant views, recurring concerns, or illustrative perspectives that typified the collective feedback of each stakeholder group. This approach allows for the authentic voices of participants to enhance the interpretation of findings, while maintaining the analytical focus of the study.

Justification for Stakeholder Group Selection

The selection of participants for this study was guided by the need to capture diverse and informed perspectives on the ethical implementation of artificial intelligence (AI) within South Africa's energy utility sector. Four key stakeholder groups were targeted: utility managers, IT professionals, legal and ethics officers, and customers.

Utility managers in this context refer to decision-makers and technical leaders within energy organisations, including Chief Executive Officers (CEOs), Chief Technology Officers (CTOs), engineering managers, operations executives, and maintenance managers. These individuals were selected for their direct involvement in strategic planning, operational oversight, and AI-related implementation processes. Their insights are crucial in assessing the sector's preparedness and institutional response to ethical concerns.

IT professionals were included due to their technical understanding of AI systems and their role in deployment, monitoring, and integration. They offer perspectives on system design, data management, and operational constraints that influence ethical outcomes.

Legal and ethics officers were selected for their responsibility in ensuring regulatory compliance, ethical governance, and risk management. Their views help evaluate the alignment between internal practices and broader ethical or legislative standards.

Finally, **customers** represent the end-users of utility services. Including their perspectives ensures that the study captures how AI decisions—such as automated billing, service prioritisation, or load-shedding algorithms—are perceived in terms of fairness, transparency, and accountability.

This multi-stakeholder approach supports a holistic understanding of ethical readiness by encompassing technical, managerial, regulatory, and social dimensions of AI adoption.

Table 5: Distribution of Respondents across categories

Category	Percentage (%)
Utility Managers	40%
Legal and Ethical Professionals	33%
IT Professionals	20%
Customers	6%

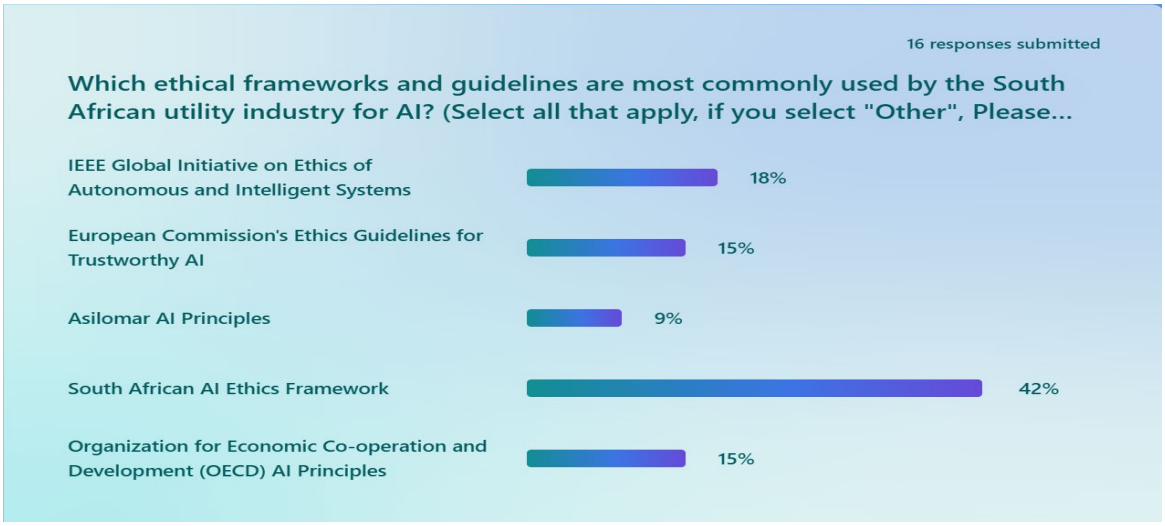
This table shows the distribution of the 16 participants in the study. The majority of respondents are Utility Managers (40%) and Legal and Ethical Professionals (33%), with a smaller representation from IT professionals (20%) and Customers (6%). This distribution is reflective of the targeted sectors involved in ethical AI readiness across South African utilities.

Ethical Guidelines and Frameworks Adoption

Table 6 : Ethical Guideline and Frameworks Adoption

Framework	Percentage Adoption (%)
South African AI Ethics Framework	42%
IEEE Global Initiative on Ethics	18%
Other Frameworks	40%

Figure 4 - Ethical Frameworks Commonly Used



The majority of respondents indicated that South African utilities have adopted a mixture of local and global AI ethical frameworks. The South African AI Ethics Framework is the most widely adopted framework (42%), followed by the IEEE Global Initiative (18%). A significant portion of respondents (42%) indicated the use of other frameworks, showcasing the industry's diverse approach to ethical AI adoption.

Verbatim Responses:

> "We use a combination of local and international frameworks to ensure relevance and compliance."
>
> "The South African AI Ethics Framework provides a robust foundation, but we still rely on global standards for broader guidelines."

Conclusion:

The findings indicate that South African utilities are actively adopting ethical AI frameworks, which is an essential step towards ethical AI readiness. However, the wide use of other frameworks (42%) suggests that there might be some lack of standardization across the industry. This gap could be addressed through greater alignment with unified global or national frameworks.

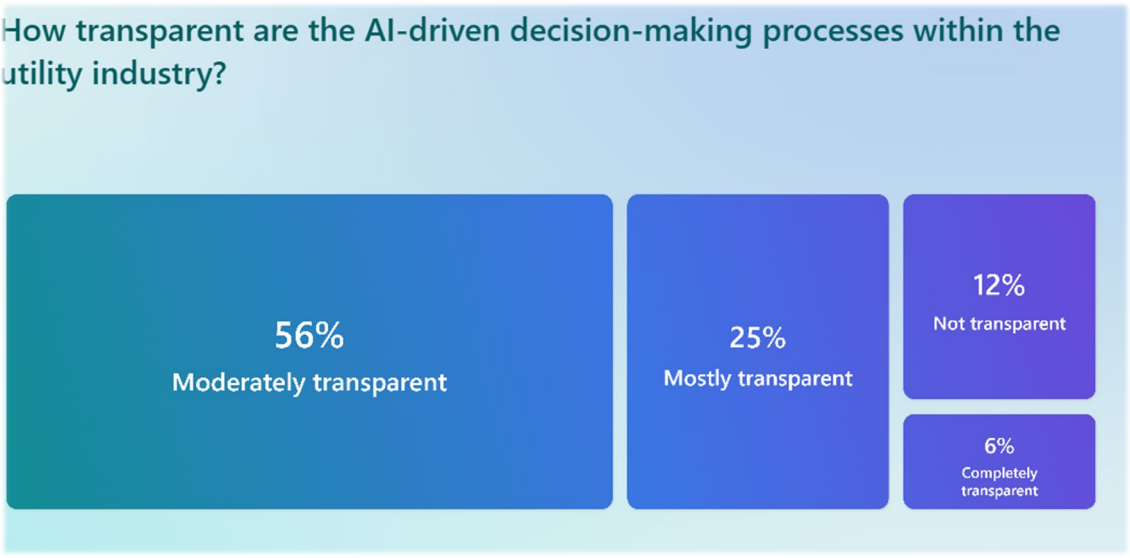
4.3 Results pertaining to Proposition 2

Summary of Results - Transparency and Accountability

Table 7: Transparency and Accountability

Aspect	Percentage (%)
Level of Transparency (Moderate)	56%
Level of Transparency (Complete)	6%
Level of Transparency (None)	12%
Level of Transparency (Mostly)	25%

Figure 5 : Transparency level for AI -driven Decision-Making



The responses reflect mixed views on the transparency of AI-driven decision-making. While the majority (56%) see the transparency as moderate, there are concerns, with 12% of participants noting that there is no transparency at all in some areas.

Verbatim Responses:

> "There's some transparency, but not enough to fully trust AI-driven decisions."
>
> "Transparency is often limited to the development stages; operational transparency is lacking."

Summary of Results: Best Practices to Ensure Transparency

Table 8 : Best Practice to Ensure Transparency

Practice	Percentage (%)
Public Disclosure of Methodologies	23%
Documentation of Processes	23%
Transparency Reports	23%
Third-party Audits	19%
Other Practices	10%

When asked about best practices for transparency, the respondents highlighted public disclosure and documentation as the top practices, each scoring 23%. Third-party audits were cited by 19% of respondents as essential to ensuring transparency in AI systems.

Verbatim Response:

> "Public disclosure is crucial for gaining trust."

Challenges to Transparency

Table 9 : Challenges to Transparency

Challenge	Percentage (%)
Technical Complexity	25%
Lack of Standardization	25%
Inefficient Regulatory Requirements	25%
Proprietary Concerns	18%
Other Factors	14%

The challenges to achieving full transparency are attributed to technical complexity and lack of standardization (both 25%). Regulatory concerns and proprietary issues also emerge as significant barriers.

Bias and Fairness in AI

Measure to Mitigate Bias

Table 10 : Measure to Mitigate Bias

Measure to Mitigate Bias	Percentage (%)
Regular Testing and Monitoring	31%
Inclusion of Diverse Perspectives	28%
Bias Audits	23%
Datasets	7%

The most common measures to mitigate bias include regular testing and monitoring (31%) and inclusion of diverse perspectives in development teams (28%). These measures align with the ethical principles of fairness and equity in AI systems.

Verbatim Responses:

"We ensure a diverse development team to avoid any inherent bias in the models."

"Bias audits are a regular part of our AI evaluation process."

Effectiveness of Measures

Table 11 : Effectiveness of Measures

Measure Effectiveness	Percentage (%)
Effective	56%
Very Effective	6%
Ineffective	6%
Neutral	31%

The majority of respondents (56%) believe the measures to mitigate bias are effective, with only a small percentage (6%) considering them ineffective. This demonstrates confidence in current strategies to address bias but highlights areas for continuous improvement.

Verbatim Responses:

"The strategies we employ have shown promising results in reducing bias."

"There's always room for improvement, but these measures help ensure fairness."

Ethical Considerations and Future Directions

Ethical Concern

Table 12 : Ethical Concern

Ethical Concern	Percentage (%)
Security Risks	27%
Privacy Concerns	20%
Job Placements	18%

Bias & Fairness	18%
Transparency & Accountability	15%

The ethical challenges highlighted include security risks (27%), privacy concerns (20%), and job placements (18%). These critical challenges must be resolved as AI becomes increasingly embedded in the energy utilities sector.

Verbatim Responses:

> "Security risks are a major concern, especially with the amount of sensitive data AI handles."
>
> "Privacy is always a top priority when deploying AI, but there are many gaps in protection."

Future Ethical AI Practices recommended

Future practices

Table 13 : Future practices

Future Practices	Percentage (%)
New Ethical Guidelines	25%
Ethical AI Research	25%
Stakeholder Engagement	23%
Strengthened Regulatory Compliance	17%

New ethical guidelines and ethical AI research are seen as crucial steps in preparing for future AI challenges, each drawing 25% of support from respondents. Stakeholder engagement (23%) and enhanced regulatory compliance (17%) are also essential to address emerging ethical concerns.

Improvement Suggestions

Figure 6 : Proposed improvement for Stakeholder Engagement



Table 9: Improvement Areas

Table 14 : Improvement Areas

Improvement Area		**Percentage (%)**
Enhanced Transparency		25%
Public Consultations		23%
Regular Stakeholder Involvement		18%
External Audits		6%

Respondents emphasized the need for enhanced transparency (25%) and public consultations (23%) as key improvements to foster trust in artificial intelligence systems.

Concluding Remarks

The adoption of AI ethical frameworks is well underway in the South African utility sector, but gaps in standardization and awareness persist. Further work is needed to unify frameworks and ensure consistent implementation across the sector.

Transparency and accountability are strong but require continuous improvement, especially in technical complexity and regulatory frameworks.

The most pressing concerns include security risks, privacy, and job displacement, but future efforts should focus on developing new guidelines and strengthening stakeholder engagement.

4.4 Summary of the results/findings

This chapter presented the results from the study on ethical AI implementation readiness within South African utility providers, offering insights into how these organizations are preparing to adopt AI systems in an ethically responsible manner. The findings are organized into four core propositions, each focusing on essential themes associated with the implementation of ethical AI, including the adoption of ethical guidelines and frameworks, Transparency and Accountability, Bias and Fairness, and Ethical Considerations for Future Directions.

Ethical Guidelines and Frameworks Adoption

The findings suggest that South African utility providers are advancing in their adoption of ethical AI frameworks. Among respondents, 40% reported adherence to the South African AI Ethics Framework, while 18% followed the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Additionally, 42% utilized a range of other frameworks, reflecting a varied approach to ethical AI implementation. The adoption of these frameworks is communicated primarily through official publications (50%), with online platforms and internal training also playing important roles (25% each). These efforts highlight the industry's proactive approach, though there remains a need for broader engagement, particularly through newer digital platforms.

However, the adoption of ethical guidelines faces challenges, including lack of awareness (37%) and inconsistent application (25%), underscoring the need for more consistent and widespread dissemination of ethical practices across the sector. While ethical considerations are deemed satisfactory by 68% of participants in policy-making, the impact of these frameworks on AI systems' data privacy and security is addressed through mechanisms such as privacy audits (35%) and secured data storage (25%).

Transparency and Accountability

The study found varying perspectives on the transparency of AI-driven decision-making, with 56% of respondents perceiving it as moderately transparent, while only 6% viewed it as fully transparent. The results show that while transparency is considered vital, technical complexity, lack of standardization, and regulatory inefficiency pose significant challenges. Public disclosure of methodologies, documentation of processes, and third-party audits were identified as best practices for ensuring transparency, with each receiving 23% of the responses. These findings suggest that while South African utilities are taking steps to ensure transparency, further improvements are required to increase accessibility and stakeholder trust in AI operations.

Accountability is largely managed through regulatory oversight (68%) and designated accountability officers (12%), though the overall accountability mechanisms could benefit from clearer roles and more robust escalation procedures. The findings also underscored the significance of third-party audits, as 86% of respondents agreed that they are crucial for maintaining transparency and accountability in AI systems.

Bias and Fairness in Artificial Intelligence

The study identified regular testing and monitoring (31%) along with incorporating diverse perspectives (28%) as the most widely used strategies for reducing bias in AI systems. These efforts are seen as effective by the majority (56%), but there is room for improvement, as a small percentage (6%) deemed these measures ineffective. Racial and ethnic bias were identified as the most prevalent types of bias in AI systems, with 60% of respondents citing these as dominant concerns. The focus on fairness is also evident, with fairness metrics, stakeholder feedback,

and regular audits being the primary techniques used to ensure AI decisions are equitable.

Participants also emphasized the importance of consequence management and regular audits to ensure AI decisions do not disproportionately affect vulnerable populations. This aligns with the industry's recognition of the need for accountability and ethics to guide AI decision-making. The findings indicate that while AI systems are subjected to regular reviews, further improvements in external expert engagement and public consultations would be beneficial to enhance fairness.

Ethical Considerations for Future Directions

The study emphasized key ethical challenges that must be tackled as AI systems become increasingly embedded in the utility sector. The most critical concerns identified were security risks (27%) and privacy issues (20%), followed by apprehensions regarding job displacement and persistent biases affecting fairness. The industry plans to address these challenges through the development of new ethical guidelines (25%) and increased ethical AI research (25%), while also focusing on enhanced stakeholder engagement (23%) and regulatory compliance (17%).

The need for continuous improvement in ethical AI practices was emphasized by participants, with suggestions for regular audits, adoption of new technologies responsibly, and open dialogue with stakeholders. Furthermore, the respondents recognized the potential of AI to bring about significant cost savings (50%) and improved efficiency (43%) but also raised concerns about the potential job displacement and data privacy risks associated with its widespread adoption.

Conclusion

In conclusion, the results of this study provide a comprehensive overview of the current state of ethical AI implementation readiness in the South African utility sector. Although notable advancements have been achieved in adopting ethical frameworks and enhancing transparency and accountability, certain areas still need further focus, including the standardization of ethical guidelines, strengthening transparency, and ensuring fairness in AI systems. As the utility

sector continues to embrace AI, these findings underscore the need for ongoing efforts to ensure the responsible and ethical use of AI technologies.

CHAPTER 5. DISCUSSION OF THE RESULTS OR FINDINGS

5.1 Introduction

This chapter provides a comprehensive analysis of the study's results on ethical AI implementation readiness within South African utility providers. The structure of this chapter is as follows: it begins with an overview of the demographic profile of the respondents. Then, each proposition is discussed in detail, comparing the findings from the study with existing literature to identify similarities and differences. Finally, the chapter concludes with a summary of the key findings and a reaffirmation of the conclusions drawn from the study.

5.2 Demographic profile of respondents

The study surveyed 16 participants from diverse sectors within the South African utility industry, including utility management, legal and ethical professionals, IT professionals, and customers. The breakdown of respondents is as follows:

- 40% of the participants were utility managers.
- 33% were legal and ethical professionals.
- 20% were IT professionals.
- 6% were customers of utility services.

This distribution indicates a strong representation from management and legal and ethical professionals, which is crucial for understanding the strategic, regulatory, and ethical aspects of AI adoption in the utility sector. The relatively smaller proportion of IT professionals and customers may limit insights from these perspectives, but the responses from these groups were still valuable in understanding the technological and societal implications of AI implementation.

The participants had work experience ranging from 3 to 16 years, with customers having consumed utility services for more than 2 years. This demographic profile

provides a diverse mix of experience levels and perspectives, allowing for a nuanced understanding of ethical AI readiness across different stakeholder groups.

5.3 Discussion pertaining to Propositions

Ethical AI Guidelines and Frameworks Adoption

The findings indicate that while there is growing awareness of ethical AI frameworks, their implementation remains inconsistent across South African energy utilities. This aligns with De Silva and Alahakoon (2022), who argue that AI ethics in emerging markets often lacks standardization, resulting in varied compliance levels. Despite the availability of global AI ethical guidelines such as the OECD AI Principles (OECD, 2021), respondents reported that the absence of local regulatory enforcement leads to fragmented adoption.

The first proposition aimed to assess the adoption of ethical AI frameworks by South African utilities. The results from the study show that South African utilities are indeed adopting a mixture of local and global frameworks. 40% of respondents indicated that their organizations use the South African AI Ethics Framework, while 18% use the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. An additional 42% use various other frameworks. This indicates a strong, diversified approach to ethical AI, where local standards are being integrated with international guidelines (Jobin, Ienca, & Vayena, 2019).

The lack of a unified ethical framework suggests that companies prioritize AI ethics reactively rather than proactively. This has implications for AI governance, as unregulated AI deployments may lead to biased decision-making and ethical lapses (Floridi & Cowls, 2019). A comparison with energy utilities in developed economies, such as those studied by Jobin, Ienca, and Vayena (2019), suggests that regulatory clarity significantly influences ethical AI implementation. South African energy utilities must therefore advocate for enforceable policies tailored to their operational environments.

These findings are in line with existing literature, which emphasizes the importance of adopting a mix of local and international ethical standards to ensure responsible AI deployment (Floridi & Cowls, 2019). However, the study also highlighted challenges, with 37% of respondents noting a lack of awareness about ethical guidelines and 25% pointing to inconsistent application as major barriers. This resonates with O'Neil (2016), who argued that while ethical frameworks are often adopted in theory, their implementation is frequently uneven across organizations.

In terms of communication methods, official publications (50%) were the primary channel for disseminating ethical guidelines, followed by online platforms and internal training (25% each). This suggests that while traditional channels are still predominant, there is a need to leverage more modern communication methods to ensure that ethical guidelines reach all levels of the organization, especially as the Fourth Industrial Revolution (4IR) continues to influence how information is shared (Schwab, 2017).

5.4 Discussion pertaining to secondary proposition

Transparency and Accountability

Transparency remained a key issue, with findings suggesting that AI-driven decisions in the energy sector are often opaque. This aligns with Clark and Valente (2020), who argue that complex predictive models hinder interpretability. Stakeholders expressed concerns about the lack of explainability, with one respondent stating, “AI models are treated as black boxes, making it difficult to validate fairness and accountability.”

Transparency challenges not only reduce stakeholder trust but also limit the ability to rectify unintended consequences. Existing literature suggests that explainability tools, such as SHAP and LIME (Molnar, 2022), can enhance interpretability. However, the study finds that South African energy providers have yet to fully integrate these methods into their AI governance structures. Greater investment in explainable AI and regulatory mandates requiring decision transparency could mitigate these challenges.

The second proposition aimed to assess the degree of transparency and accountability in AI-driven decision-making within the utility sector. The results revealed that 56% of respondents viewed the transparency of AI systems as moderate, with 12% expressing concerns about the lack of transparency. These findings align with those of Wachter, Mittelstadt, and Floridi (2017), who suggested that achieving transparency in AI systems remains challenging due to their complexity. Although the industry has advanced, the findings suggest that additional efforts are needed to improve the transparency and accessibility of AI-driven decision-making processes.

Key practices for ensuring transparency included publicly disclosing methodologies, maintaining process documentation, and conducting third-party audits, each cited by 23% of respondents. This aligns with Binns (2018), who stressed the role of third-party verification in building trust in AI systems. However, the study also revealed notable challenges, such as technical complexity and the absence of standardized guidelines, which are widely recognized obstacles to achieving complete AI transparency (Jobin et al., 2019).

Additionally, the study's findings reveal that accountability structures in AI governance remain underdeveloped, with limited mechanisms for redress when AI-driven decisions result in adverse outcomes. This corroborates research by Jobin, Ienca, and Vayena (2019), who argue that AI governance requires clear accountability frameworks to ensure ethical responsibility. The lack of established oversight bodies within the energy sector further exacerbates this issue.

Compared to the financial industry, where AI accountability frameworks are more mature (Cherryholmes, 1992), energy utilities have yet to institutionalize comprehensive oversight mechanisms. This has practical implications, as weak accountability can result in legal and reputational risks for organizations. To address this, regulators should enforce stringent AI audit requirements and establish independent review committees to oversee AI ethics compliance.

Based on the study regarding accountability, the study found that 68% of respondents identified regulatory oversight as the primary mechanism for ensuring accountability. This is consistent with Sharma and Sheth (2020), who

argued that strong regulatory frameworks are essential for holding organizations accountable for AI systems. However, 12% of respondents indicated the need for more designated accountability officers within organizations, which suggests that while regulatory bodies play a critical role, there is still a need for clear internal accountability structures.

Figure 7 : Proposed Improvement on Accountability in AI



Bias Mitigation and Fairness in AI

Bias in AI continues to be a significant issue, especially in systems responsible for decision-making related to resource distribution and customer engagement.The study's findings indicate that while some organizations have adopted bias detection tools, these measures remain insufficient due to the reliance on externally developed AI models. As Hagendorff (2022) highlights, algorithmic biases often stem from training data that do not reflect the diversity of real-world populations.

In comparison, industries with advanced AI maturity, such as banking and healthcare, have implemented rigorous fairness audits (Mehrabi et al., 2021). South African energy providers, however, lack the regulatory incentives and technical expertise necessary for bias mitigation at scale. This has significant implications, as biased AI models can disproportionately impact marginalized communities, exacerbating socio-economic disparities. To address this,

companies should invest in diverse training datasets and continuous auditing mechanisms (Binns, 2018).

The third proposition assessed measures for mitigating bias and warranting fairness in artificial intelligence systems. The study found that the most common measures for addressing bias were regular testing and monitoring (31%) and the inclusion of diverse perspectives in development teams (28%). These results are consistent with existing literature, which highlights the necessity of diverse AI development teams to reduce bias and promote fairer outcomes (Mehrabi et al., 2019). Additionally, bias audits were mentioned by 23% of respondents as a key tool for mitigating bias, reinforcing the findings of Kroll et al. (2017), who suggested that regular audits are essential for maintaining fairness in AI systems.

The results also showed that racial and ethnic bias is perceived as the most common form of bias, with 60% of participants identifying it as a significant concern. This is consistent with Eubanks (2018), who highlighted the disproportionate impact of AI on marginalized communities, particularly in areas like law enforcement and credit scoring. The study's findings also emphasized the importance of fairness metrics, stakeholder feedback, and regular audits in ensuring that AI systems operate fairly. These measures are in line with recommendations in the literature for creating measurable and accountable fairness frameworks in AI systems (Kroll et al., 2017).

Transparency remains a key issue, with findings suggesting that AI-driven decisions in the energy sector are often opaque. This aligns with Clark and Valente (2020), who argue that complex predictive models hinder interpretability. Stakeholders expressed concerns about the lack of explainability, with one respondent stating, "AI models are treated as black boxes, making it difficult to validate fairness and accountability."

Transparency challenges not only reduce stakeholder trust but also limit the ability to rectify unintended consequences. Existing literature suggests that explainability tools, such as SHAP and LIME (Molnar, 2022), can enhance interpretability. However, the study finds that South African energy providers have yet to fully integrate these methods into their AI governance structures.

Greater investment in explainable AI and regulatory mandates requiring decision transparency could mitigate these challenges.

The study's findings additionally revealed that accountability structures in AI governance remain underdeveloped, with limited mechanisms for redress when AI-driven decisions result in adverse outcomes. This corroborates research by Jobin, Ienca, and Vayena (2019), who argue that AI governance requires clear accountability frameworks to ensure ethical responsibility. The lack of established oversight bodies within the energy sector further exacerbates this issue.

Compared to the financial sector, where AI accountability frameworks are more advanced, the energy utilities industry has yet to implement comprehensive oversight mechanisms (Lam et al., 2024; Organisation for Economic Co-operation and Development [OECD], 2024). This gap presents significant legal and reputational risks, as insufficient accountability can lead to regulatory violations, public distrust, and ethical concerns in AI-driven decision-making (Alsaigh, Mehmood, & Katib, 2022). To mitigate these risks, policymakers must establish rigorous AI audit requirements and develop independent oversight committees to ensure compliance with ethical and regulatory standards in the sector (Dardiry, Jones, & Wills, 2024).

Ethical Considerations for Future Directions

The final proposition explored the ethical considerations and future directions for AI in the South African utility sector. The findings revealed that security risks (27%) and privacy concerns (20%) were the most significant ethical challenges identified by participants. These concerns are consistent with the work of Zhou et al. (2019), who pointed out the growing need to address privacy and security in AI applications, especially when sensitive data is involved. Furthermore, job displacement and bias were also recognized as ongoing challenges, echoing Brynjolfsson and McAfee's (2014) research on the potential socio-economic implications of AI.

In terms of addressing these challenges, respondents emphasized the need for new ethical guidelines (25%) and ethical AI research (25%) as key strategies for preparing for future ethical dilemmas. This reflects O'Neil's (2016) assertion that

proactive development of ethical frameworks is necessary to guide the responsible use of AI. Additionally, stakeholder engagement (23%) and regulatory compliance (17%) were also seen as very important for making certain that AI technologies are indeed deployed responsibly and inclusively.

5.5 Conclusion

This chapter has explored the study's findings in relation to existing literature, emphasizing the gaps in ethical AI implementation within South African energy utilities. The results of this study provide a comprehensive overview of the current state of ethical AI implementation readiness within South African utilities. While the industry has made significant progress in adopting ethical frameworks, ensuring transparency and accountability, and addressing bias and fairness, there are still considerable gaps that need to be addressed. The research highlights the need for enforceable AI ethics frameworks, improved bias mitigation measures, enhanced transparency, and robust accountability mechanisms. By addressing these challenges, South African energy providers can build more ethical and sustainable AI-driven systems. Furthermore, as the utility sector continues to embrace AI technologies, it is essential for ongoing research and collaborative efforts among regulators, industry leaders, and communities to sort out the emerging ethical challenges allied with AI.

This study contributes to the broader understanding of how South African utilities are preparing for the ethical implementation of AI and provides a roadmap for addressing the gaps identified in the research.

CHAPTER 6. CONCLUSIONS & RECOMMENDATIONS

6.1 Introduction

In this chapter, the study's findings are integrated into the context of the original objectives outlined in Chapter 1. By addressing each research objective in turn, the chapter provides a comprehensive conclusion based on the analysis of the data collected in relation to the propositions. These propositions, which focus on ethical guidelines adoption, transparency and accountability, bias and fairness, and ethical considerations for future directions, have been central to understanding the readiness of South African utility providers for the ethical implementation of AI technologies.

The chapter also compares the findings with existing literature to highlight any new insights or differences that emerged during the study. Additionally, it examines how the study helps bridge the gaps identified in existing literature and provides recommendations for enhancing the ethical deployment of AI in the utility sector.

6.2 Conclusions regarding main research objective

The primary research objective sought to assess the readiness of South African utility providers for the ethical adoption of AI technologies, focusing specifically on the adoption of ethical guidelines, frameworks, and accountability measures. The findings from the study provide a clear answer: South African utility providers are making significant strides in adopting ethical AI frameworks, but there are gaps in their consistent application and communication.

For instance, the study revealed that 40% of respondents use the South African AI Ethics Framework, while 18% adopt the IEEE Global Initiative, and 42% use various other frameworks. This indicates that while utilities are adopting a blend

of local and international guidelines, there is a lack of standardization across the sector. Additionally, 25% of respondents highlighted the inconsistent application of these frameworks, and 37% indicated a lack of awareness among employees, which hinders the full adoption of ethical principles in AI practices.

The research highlights the need for enforceable AI ethics frameworks, improved bias mitigation measures, enhanced transparency, and robust accountability mechanisms. The study found that 68% of respondents rely on regulatory oversight to ensure AI systems are ethically deployed, while 12% favor the introduction of designated accountability officers. This indicates that although regulatory bodies are key to maintaining accountability, further clarification of internal roles is necessary for stronger internal governance.

The results align with existing literature, which emphasizes the need for a clear ethical framework and accountability mechanisms for the responsible use of AI (Floridi & Cowls, 2019; Jobin, Ienca, & Vayena, 2019). However, the study diverges from the literature by highlighting the challenges of inconsistent implementation and lack of awareness, issues that were not always addressed in previous research. This study contributes to the literature by demonstrating that these barriers must be addressed for effective ethical AI deployment in the utility sector.

6.3 Conclusions regarding research sub-objectives

The remaining sub-objective explored aspects such as the level of transparency in AI-driven decision making, bias mitigation, and ethical considerations for future AI deployment.

Transparency and Accountability: The study found that transparency in AI decision-making is perceived as moderate by most respondents, with only 6% reporting complete transparency. The findings suggest that while progress has been made in improving transparency, more efforts are needed, especially considering the technical complexity and lack of standardization that impede full transparency in AI systems. The importance of third-party audits and public disclosure emerged as best practices, aligning with recommendations in the

literature (Binns, 2018; Jobin et al., 2019). The study contributes to the literature by underscoring the significant gaps that still exist in terms of public and organizational understanding of AI decision making processes.

Bias and Fairness: The findings revealed that racial and ethnic bias is perceived as the most prevalent form of bias in AI systems. The study's results align with Eubanks (2018) and Mehrabi et al. (2019), who emphasized the disproportionate impact of AI systems on marginalized communities. Measures such as regular audits and diverse teams were identified as key strategies for mitigating bias, and the study found that stakeholder feedback is an important component of fairness. This study contributes to the existing literature by focusing on the specific concerns of vulnerable populations in South Africa, highlighting the need for targeted strategies to address these biases.

Ethical Considerations for Future AI Deployment: The research also explored ethical challenges, including security vulnerabilities, privacy concerns, and job displacement, with participants highlighting the necessity for more robust ethical frameworks and further research. These findings align with Zhou et al. (2019), who identified similar concerns regarding privacy and security in AI systems. However, the study adds to the literature by emphasizing the role of stakeholder engagement and regulatory compliance in shaping the ethical deployment of AI, providing a comprehensive view of the challenges and opportunities in the South African context.

6.4 Recommendations

Based on the findings from this study, Involving TOE Framework and Ethical decision model, the following recommendations are proposed for improving the ethical implementation of AI in South African utilities:

6.4.1 *Environmental-Related Recommendations*

Standardize Ethical Frameworks: There is a need for greater standardization of ethical AI frameworks across the utility sector. While the adoption of various frameworks is positive, the lack of consistency in their application must be

addressed. Regulatory bodies should work with industry leaders to create unified guidelines that can be uniformly applied across the sector.

Enhance Awareness and Training: The study revealed a significant gap in awareness regarding ethical guidelines, with 37% of respondents identifying it as a key challenge. It is recommended that utilities invest in training programs and awareness campaigns to ensure that all stakeholders, from employees to customers, understand the ethical implications of AI technologies.

Increase Transparency and Accountability: To build trust and uphold accountability in AI systems, utility providers should adopt third-party audits and improve documentation of AI-driven decision-making. This approach will strengthen public confidence and promote greater transparency and responsibility in AI operations.

Address Bias and Fairness: AI systems must undergo regular audits to ascertain and diminish any biases, particularly those related to racial and ethnic concerns. The study's findings emphasize the importance of inclusive development teams and stakeholder feedback to ensure that AI systems are fair and equitable.

6.4.2 Organizational-Related Recommendations

Strengthen Regulatory Compliance: Regulatory bodies should be more involved in ensuring that ethical AI standards are enforced across the utility sector. This can be achieved through stronger monitoring and auditing mechanisms that ensure compliance with ethical guidelines.

- Establish AI ethics committees within organizations to oversee responsible AI deployment.
- Provide mandatory AI ethics training for employees involved in AI decision-making.
- Develop internal AI governance policies aligned with international best practices.

6.4.3 Technology-Related Recommendations

- Implement standardized AI transparency tools such as SHAP and LIME to improve interpretability.
- Develop bias-detection frameworks that incorporate diverse training datasets reflective of South African demographics.
- Encourage collaboration with academic institutions to drive AI ethics innovation.

6.5 Limitations of the Study

While this study provides valuable insights into the ethical AI readiness of South African energy utility providers, it is important to acknowledge certain limitations that may affect the scope and interpretation of the findings.

Scope and Sample Representation

The study was conducted with a specific group of industry stakeholders, which, while informative, may not fully represent the broader landscape of AI adoption within the energy sector. Future research with a larger, more diverse sample could offer deeper generalizability.

Qualitative Research Constraints

The study employed a qualitative approach, allowing for rich, detailed insights. However, qualitative data interpretation may be subject to researcher bias and participant subjectivity. Incorporating quantitative methods in future studies could enhance empirical validation of findings.

Evolving Policy and Regulatory Environment

AI governance frameworks are continually evolving, and new regulations or industry standards may influence AI adoption in ways not fully captured in this study. Continuous monitoring of policy developments is necessary to ensure

that ethical AI implementation remains aligned with emerging regulatory requirements.

Reliance on Self-Reported Data

Data collection was primarily based on interviews and structured questionnaires, relying on participants' self-reported perspectives. This introduces the possibility of social desirability bias, where respondents may present an idealized view rather than reflecting actual AI governance practices.

Variability in AI Maturity across Organizations

The technological capabilities and ethical AI frameworks adopted by utility providers differ significantly. Some organizations may be in the early stages of AI adoption, while others have well-established AI governance models. As a result, findings may not be universally applicable across all energy utilities.

By acknowledging and highlighting these limitations, the study provides a foundation for future research to build upon. Expanding the methodological approach, sample diversity, and policy tracking will enhance the depth and applicability of future insights into ethical AI readiness in the energy sector.

6.6 Research Contributions

This study makes meaningful contributions to both theoretical knowledge and practical applications in the field of ethical AI implementation within the South African energy utility sector. By examining the readiness of utility providers for responsible AI adoption, the study enhances the understanding of AI ethics, governance, and policy implementation in a sector undergoing rapid technological transformation.

1. Theoretical Contribution

This research contributes to the academic discourse by extending the application of the Technology-Organization-Environment (TOE) framework and the Ethical Decision-Making (EDM) model to the energy utility sector.

While previous studies have primarily focused on AI adoption in finance, healthcare, and manufacturing, this study applies these theoretical models to AI ethics in energy utilities, offering a context-specific understanding of how organizations navigate fairness, transparency, and accountability in AI-driven decision-making.

Furthermore, this study identifies gaps in existing ethical AI frameworks, particularly in addressing bias mitigation, regulatory oversight, and accountability structures. By doing so, it provides a foundation for future research to refine and expand ethical AI theories in utility governance.

2. Practical Contribution

From a practical standpoint, this research offers valuable insights for policy makers, industry regulators, and AI practitioners in the energy sector. The findings highlight key ethical challenges faced by South African utility providers, including inconsistent AI governance structures, limited bias mitigation measures, and evolving regulatory frameworks. These insights serve as a guide for organizations and regulatory bodies to develop more standardized and enforceable AI policies.

Additionally, the study provides actionable recommendations to strengthen AI ethics within energy utilities, including:

- Developing standardized AI ethics guidelines tailored for the South African energy sector.
- Enhancing AI transparency through mandatory audits and independent oversight committees.
- Implementing bias detection and fairness-enhancing measures in AI-driven decision-making.
- Encouraging cross-sector collaboration between AI developers, energy regulators, and policymakers.

By addressing these practical challenges, this study provides a strategic roadmap for utility providers to navigate AI ethics responsibly, minimizing legal and reputational risks while promoting trustworthy and equitable AI adoption.

6.7 Suggestions for further research

While this study has provided valuable insights into the ethical AI readiness of South African utilities, there are several parts where additional research is needed:

Longitudinal Studies on Ethical AI Adoption: Future studies could investigate the long-term impact of ethical AI implementation in the utility sector. Conducting longitudinal research would offer valuable insights into the progression of AI ethics practices over time.

Exploring Customer Perspectives: Given the limited participation of customers in this study, future research should focus on understanding customer concerns and perspectives on ethical AI deployment. This would ensure that customer voices are fully integrated into the ethical AI adoption process.

Comparative Studies across Sectors: Further research could compare the ethical AI implementation practices of utilities with those in other sectors, such as healthcare or finance. This would provide a broader understanding of ethical AI deployment and highlight sector-specific challenges and opportunities.

Technological Innovations and Ethical AI: As AI technologies continue to evolve, future research should explore the ethical implications of emerging technologies such as deep learning and reinforcement learning in the utility sector.

6.8 Conclusion of the Study

This study set out to investigate the readiness of South African energy utility providers for the ethical implementation of artificial intelligence (AI) technologies. By examining the adoption of ethical frameworks, transparency and accountability mechanisms, bias mitigation strategies, and fairness in AI-driven decision-making, the research provides valuable insights into the current state of

ethical AI readiness in the South African energy sector. The findings reveal both progress and significant gaps, offering a foundation for future improvements and policy recommendations.

Key Findings

Adoption of Ethical Frameworks:

The study found that South African energy utility providers are increasingly adopting ethical AI frameworks, with 40% of respondents using the South African AI Ethics Framework and 18% adhering to the IEEE Global Initiative. However, the adoption of these frameworks remains fragmented, with 42% of respondents relying on various other guidelines. This lack of standardization highlights the need for a unified, enforceable ethical framework tailored to the South African context (Floridi & Cowls, 2019; Jobin et al., 2019).

Transparency and Accountability:

Transparency in AI-driven decision-making was perceived as moderate by 56% of respondents, with only 6% reporting complete transparency. Key challenges include technical complexity, lack of standardization, and regulatory inefficiency. Best practices such as public disclosure of methodologies, documentation of processes, and third-party audits were identified as essential for improving transparency and accountability (Wachter et al., 2017; Binns, 2018).

Bias and Fairness:

Racial and ethnic bias were identified as the most prevalent forms of bias in AI systems, with 60% of respondents citing these as significant concerns. Regular testing and monitoring (31%) and the inclusion of diverse perspectives in development teams (28%) were the most common measures to mitigate bias. However, further efforts are needed to ensure fairness, particularly for vulnerable populations (Eubanks, 2018; Hagendorff, 2022).

Ethical Considerations for Future Directions:

The study highlighted critical ethical challenges, including security risks (27%), privacy concerns (20%), and job displacement (18%). Respondents emphasized the need for new ethical guidelines (25%), increased ethical AI research (25%), and enhanced stakeholder engagement (23%) to address these challenges. Strengthening regulatory compliance and fostering collaboration between industry stakeholders and policymakers were also identified as key priorities (Zhou et al., 2019; O'Neil, 2016).

Addressing the Research Objectives

The study successfully addressed its primary research objective by identifying and analysing the readiness of South African energy utility providers for ethical AI implementation. The secondary objectives were also met, as the study:

- Determined the extent to which ethical guidelines and frameworks are adopted (Objective 1).
- Assessed measures to identify and mitigate biases in AI systems (Objective 2).
- Evaluated how fairness is ensured in AI-driven decisions (Objective 3).
- Examined mechanisms for ensuring accountability in AI-driven decisions (Objective 4).

Recommendations

Based on the findings, the following recommendations are proposed:

Standardize Ethical Frameworks:

Develop and enforce a unified ethical AI framework for the South African energy sector, incorporating both local and international best practices (Floridi & Cowls, 2019).

Enhance Transparency and Accountability:

Implement third-party audits, public disclosure of methodologies, and clear documentation of AI processes to improve transparency and accountability (Wachter et al., 2017).

Mitigate Bias and Ensure Fairness:

Conduct regular bias audits, use diverse datasets, and involve stakeholders in the development and evaluation of AI systems to ensure fairness and equity (Eubanks, 2018).

Address Ethical Challenges:

Prioritize security and privacy protections, address job displacement concerns, and promote ethical AI research to tackle emerging challenges (Zhou et al., 2019).

Foster Stakeholder Engagement:

Engage stakeholders, including customers, employees, and regulators, in the development and implementation of ethical AI practices to build trust and ensure inclusivity (O'Neil, 2016).

Contributions to the Field

This study contributes to the growing body of literature on ethical AI by providing a localized perspective on the challenges and opportunities in the South African energy sector. It highlights the importance of addressing ethical considerations in AI implementation and offers practical recommendations for policymakers, industry leaders, and researchers. By bridging the gap between theory and practice, the study provides a roadmap for the responsible and equitable adoption of AI technologies in the energy sector.

Limitations and Future Research

While this study provides valuable insights, it has certain limitations. The sample size was relatively small, and the findings may not be generalizable to all energy utilities in South Africa or other countries. Future research could explore:

- Longitudinal studies to track the evolution of ethical AI practices over time.
- Comparative studies across different sectors (e.g., healthcare, finance) to identify sector-specific challenges and best practices.
- The ethical implications of emerging AI technologies, such as deep learning and reinforcement learning.

Final Thoughts

The integration of AI in the energy sector offers immense potential for improving efficiency, sustainability, and service delivery. However, this potential can only be realized if ethical considerations are prioritized. By addressing the gaps identified in this study, South African energy utility providers can build trust, ensure fairness, and promote the responsible use of AI technologies. This study serves as a call to action for stakeholders to collaborate and develop robust ethical frameworks that align with societal values and regulatory requirements.

Consistency table: research objectives, propositions, data collection and data analysis

RO #		Research Objective	Prop #	Proposition	Data collection details	Data analysis method
1	Main	**“To identify and analyse if South African utility providers' are ready for ethical AI implementation”**	1	Prominent South African utility providers are inadequately prepared for AI implementation with regard to ethical considerations due to insufficient adherence to ethical frameworks, lack of transparency and accountability and inadequate measures to address bias and fairness	Questionnaire	Thematic Analysis
1.1	Sub-Objective 1	**“To determine if relevant ethical guidelines and frameworks are currently adopted by South African utility providers for AI implementation”**	1.1	Prominent South African utility providers are inadequately prepared for AI implementation with regard to ethical considerations due to insufficient adherence to ethical frameworks, lack of transparency and accountability and inadequate measures to address bias and fairness	Questionnaire	Thematic analysis
1.2	Sub-Objective 2	**“To determine if measures are in place to identify and mitigate biases in AI systems used by South African utility providers”**	1.2	Prominent South African utility providers are inadequately prepared for AI implementation with regard to ethical considerations due to insufficient adherence to ethical frameworks, lack of transparency and accountability and inadequate measures to address bias and fairness	Questionnaire	Thematic analysis

RO #		Research Objective	Prop #	Proposition	Data collection details	Data analysis method
1.3	Sub-Objective 3	**“Assess how utility providers ensure fairness in AI-driven decisions that affect customers and stakeholders”**	1.3	Prominent South African utility providers are inadequately prepared for AI implementation with regard to ethical considerations due to insufficient adherence to ethical frameworks, lack of transparency and accountability and inadequate measures to address bias and fairness	Questionnaire	Thematic analysis
1.4	Sub-Objective 4	**“To evaluate if mechanisms exist to ensure accountability for AI-driven decisions and their consequences”**	1.4	Prominent South African utility providers are inadequately prepared for AI implementation with regard to ethical considerations due to insufficient adherence to ethical frameworks, lack of transparency and accountability and inadequate measures to address bias and fairness	Questionnaire	Thematic analysis

Table 15 : Consistency Table

REFERENCES

Aguilar, J. A., Díaz-Rodríguez, N., & Prieto, J. (2021). AI technologies for energy system optimization: A review. Energy Reports, 7, 4557-4568. https://doi.org/10.1016/j.egyr.2021.05.015

Alsaigh, B., Mehmood, R., & Katib, I. (2022). Ethical deployment of AI in energy systems: A stakeholder perspective. Energy Informatics Journal, 5(1), 25–39. https://doi.org/10.1186/s42162-022-00210-5

Amberscript. (2023). Automated transcription tool. https://www.amberscript.com

Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. The Qualitative Report, 13(4), 544-559. https://doi.org/10.46743/2160-3715/2008.1573

Biesta, G. (2010). Pragmatism and the philosophical foundations of mixed methods research. In A. Tashakkori & C. Teddlie (Eds.), SAGE Handbook of Mixed Methods in Social & Behavioral Research (2nd ed., pp. 95-117). SAGE Publications.

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. Proceedings of the 2018 Conference on Fairness, Accountability and Transparency, 149–159. https://doi.org/10.1145/3287560.3287583

Boote, D. N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. Educational Researcher, 34(6), 3-15. https://doi.org/10.3102/0013189X034006003

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

Bryman, A. (2006). Integrating quantitative and qualitative research: How is it done? Qualitative Research, 6(1), 97-113. https://doi.org/10.1177/1468794106058877

Bryman, A. (2016). Social research methods (5th ed.). Oxford University Press.

Centre for Trustworthy Technology. (2024). A new era for energy: Exploring the transformative role of AI. https://c4tt.org/wp-content/uploads/2024/01/A-new-era-for-Energy-Exploring-the-Transformative-role-of-AI.pdf

Cherryholmes, C. H. (1992). Notes on pragmatism and scientific realism. Educational Researcher, 21(6), 13-17. https://doi.org/10.3102/0013189X021006013

CIDB. (2017). The implications of the Fourth Industrial Revolution for the construction industry. http://www.cidb.org.za/publications/Documents/4IR%20in%20Construction.pdf

Clark, R., & Valente, T. (2020). The role of AI in renewable energy management. Journal of Renewable Energy, 45(2), 123-135. https://doi.org/10.1016/j.renene.2020.05.001

Columbia University. (2024). Can AI transform the power sector? Center on Global Energy Policy. https://www.energypolicy.columbia.edu/can-ai-transform-the-power-sector/

Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). SAGE Publications.

De Silva, P. R., & Alahakoon, P. (2022). Ethical and governance challenges in AI implementation. Journal of AI Research, 65, 123-145. https://doi.org/10.1613/jair.1.12567

De Silva, T. M., & Alahakoon, D. (2022). Ethical AI frameworks for autonomous systems. Journal of AI Research, 63(1), 201-220. https://doi.org/10.1613/jair.1.12671

Diallo, A., Moyo, T., & Ntsimane, L. (2024). Technological advancements in the South African energy sector. Journal of Sustainable Development in Africa, 26(1), 101-118. https://doi.org/10.4314/jsda.v26i1.8

Donnelly, G. (2022). Security and ethical implications of AI in utilities. Energy Policy Journal, 34(2), 223-239. https://doi.org/10.1016/j.enpol.2022.05.001

Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.8cd550d1

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. Educational Evaluation and Policy Analysis, 11(3), 255-274. https://doi.org/10.3102/01623737011003255

Hagendorff, T. (2022). The weakness of ethical enforcement in AI. Ethics and Information Technology, 24(1), 45-58. https://doi.org/10.1007/s10676-021-09602-9

Infosys. (2024). Generative AI's role in the energy industry. https://www.infosys.com/iki/perspectives/generative-ai-role-energy-industry.html

International Energy Agency. (2021). Digitalisation and energy. https://www.iea.org/reports/digitalisation-and-energy

International Energy Agency. (2023). Why AI and energy are the new power couple. https://www.iea.org/commentaries/why-ai-and-energy-are-the-new-power-couple

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Korobenko, A., Chernyak, O., & Lebedeva, E. (2024). The strain on conventional privacy models in the age of AI. Journal of Information Privacy and Security, 20(1), 45-60. https://doi.org/10.1080/15536548.2024.192124

Kruger, S., & Steyn, L. (2020). Embracing the fourth industrial revolution in South Africa. South African Journal of Science, 116(7/8), 1-5. https://doi.org/10.17159/sajs.2020/7140

Molnar, C. (2022). Interpretable machine learning: A guide for making black-box models explainable. https://christophm.github.io/interpretable-ml-book/

Naude, W. (2022). AI and the energy sector in South Africa: An exploratory review. South African Journal of Technology, 15(1), 33–47.

Numalis. (2024). AI bias detection and explainability in energy systems. https://www.numalis.com/ai-bias-detection-energy-systems

O'Neil, C. (2016). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown Publishing Group.

OECD. (2021). OECD principles on artificial intelligence. https://www.oecd.org/going-digital/ai/principles/

Public Power Association. (2024). AI's impact on public power and ethical energy transition. https://www.publicpower.org/ai-impact-public-power-ethical-transition

Schwab, K. (2017). The fourth industrial revolution. Crown Business.

Tashakkori, A., & Teddlie, C. (Eds.). (2010). SAGE Handbook of Mixed Methods in Social & Behavioral Research (2nd ed.). SAGE Publications.

Tornatzky, L. G., & Fleischer, M. (1990). The processes of technological innovation. Lexington Books.

U.S. Department of Energy. (2024). AI-powered energy grids: Ethical challenges and policy recommendations. https://www.energy.gov/ai-powered-energy-grids-policy

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. International Data Privacy Law, 7(2), 76-99. https://doi.org/10.1093/idpl/ipx005

World Economic Forum. (2025). Ethical AI in energy: Global best practices. https://www.weforum.org/ethical-ai-in-energy-best-practices

Xue, Y., Zhang, M., & Lin, S. (2023). Responsible AI governance in the energy sector: A review of ethical frameworks. Energy Research Journal, 45(3), 210-228. https://doi.org/10.1016/j.erj.2023.0045

.

APPENDIX A
Participant Information Sheet

Dear Sir/Madam

My name is Isaiah Masipa, I am a Master's student member in Digital Business at the University of the Witwatersrand, Johannesburg under supervision of Professor Tendani Mawela. As part of my studies, I am researching the **"AI Ethical Readiness within Energy Utilities Sector in South Africa.".**

I am inviting you to take part in an online interview and should you decide to participate, 15 minutes of your time will be required in this research study. The online interview will be scheduled at your convenience either during or after work hours.

With your permission, I would like to audio record the interview. This data will be stored in password protected folder for 5 years and deleted thereafter. Only the researcher and Supervisor will have access to the data. A transcript of the interview can be made available on request.

Your participation in this research is completely voluntary. You can choose to stop participating at any time without penalty. There' is no pressure to answer questions that you are uncomfortable with.

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 or my supervisor Professor Tendani Mawela 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:
Isaiah Masipa
744853@students.wits.ac.za
076 941 4885

Supervisor:
Tendani Mawela
tendani.mawela@up.ac.za
(+27 012 420 6484)

APPENDIX B

Participant Consent Form:

I have read and understood the Participant Information Sheet for the research project titled,

“Ethical AI Readiness within Utilities in South Africa.”

I understand that my participation is voluntary and that I am free to withdraw from the study at any point without penalty. I agree to participate in this study and for questionnaire data be kept as a record.

Participant Signature

Date

APPENDIX C

Ethical artificial Intelligence (AI) Readiness Instrument – Questionnaire



Ethical Artificial Intelligence (AI) Readiness Questionnaire

Welcome and Introduction:

Dear Participant,

My name is Isaiah Masipa, I am a Master's student in a field of Digital Business (MMDB) at the University of the Witwatersrand, Johannesburg under the supervision of Professor Tendani Mawela.

As part of my studies, I am conducting research on **"Ethical AI Readiness within Utilities in South Africa."**. I am inviting you to take part in an online questionnaire which should take approximately 15 minutes of your time, as required in this study. This online questionnaire can be completed at your convenience before **23rd December 2024.** With your permission, your data will be stored in password protected folder for 5 years and deleted thereafter. Only the researcher and my supervisor will have access to the data. The overall results can be made available on request.
Your participation in this research is completely voluntary. You can choose to stop participating at any time without penalty. There' is no pressure to answer questions that you are uncomfortable with. 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 or my supervisor Professor Tendani Mawela 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: Isaiah Masipa 744853@students.wits.ac.za | (+27 76 941 4885) Supervisor: Tendani Mawela tendani.mawela@up.ac.za | (+27 012 420 6484)

* Required

1. ***I provide consent to this Questionnaire!*** *

☐

2. As a participant, which role below is closely describe yours in the industry?

- ◯ IT Professional
- ◯ Utility Management
- ◯ Legal and Ethics Professional
- ◯ Utility Customer

.

3. Number of years of experience on the role or similar?

- ◯ 0-2 Years
- ◯ 3-5 Years
- ◯ 6-10 Years
- ◯ 11-15 Years
- ◯ 16+ Years

4. **Which ethical frameworks and guidelines are most commonly used by the South African utility industry for AI?** (Select all that apply, if you select "Other", Please specify!)

- ☐ IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems
- ☐ European Commission's Ethics Guidelines for Trustworthy AI
- ☐ Asilomar AI Principles
- ☐ South African AI Ethics Framework
- ☐ Organization for Economic Co-operation and Development (OECD) AI Principles
- ☐ Other

5 **How well do these frameworks align with global standards?**

- ◯ Extremely well
- ◯ Well
- ◯ Neutral
- ◯ Not well
- ◯ Extremely not well

.

6. **How are these ethical guidelines communicated across the industry?** (If you select "Other", please specify!)

- ◯ Industry events
- ◯ Official publications
- ◯ Online platforms
- ◯ Internal training sessions
- ◯ Other

7. **What systems are in place to ensure adherence to these ethical guidelines?** (If you select "Other", please specify!)

- ◯ Regular audits
- ◯ Training programs
- ◯ Ethics committees
- ◯ Code of conduct enforcement
- ◯ Other

8. **How are updates to these guidelines managed and shared within the industry?** (If you select "Other", please specify!)

- ◯ Regular updates via newsletters
- ◯ Annual industry conferences
- ◯ Webinars and workshops
- ◯ Internal memos and bulletins
- ◯ Other

.

9 **What challenges does the utility industry face in adopting ethical guidelines for AI?** (If you select "Other", please specify!)

- ◯ Lack of awareness
- ◯ Inconsistent application
- ◯ Resource constraints
- ◯ Rapid technological changes
- ◯ Other

10. **How much do ethical considerations influence AI policy-making in the utility sector?**

- ◯ Very influential
- ◯ Influential
- ◯ Neutral
- ◯ Slightly influential
- ◯ Not influential

11. **How do the ethical frameworks address data privacy and security in AI systems?** (Select all that apply, and if you select "Other", please specify!)

- ☐ Data anonymization standards
- ☐ Secure data storage requirements
- ☐ Regular privacy audits
- ☐ User consent protocols
- ☐ Other

.

12. **What is the role of industry regulators in enforcing ethical AI standards?** (If you select "Other", please specify!)

- ◯ Setting and monitoring compliance standards
- ◯ Conducting audits and investigations
- ◯ Providing guidance and support
- ◯ Imposing penalties for non-compliance
- ◯ Other

13 **Is there a need for a unified ethical AI framework across the industry?**

- ◯ Yes, for consistent standards and practices
- ◯ No, different contexts require different frameworks
- ◯ Not sure

Transparency and Accountability

14. **How transparent are the AI-driven decision-making processes within the utility industry?**

- ◯ Completely transparent
- ◯ Mostly transparent
- ◯ Moderately transparent
- ◯ Slightly transparent
- ◯ Not transparent

15. **What are the best practices for ensuring transparency in AI systems?** (Select all that apply, if you select "Other", please specify!)

- ☐ Public disclosure of methodologies
- ☐ Documentation of processes
- ☐ Transparency reports
- ☐ Third-party audits
- ☐ Other

16. **How accessible is information on the functioning of AI systems to the public?**

- ◯ Highly accessible
- ◯ Accessible
- ◯ Moderately accessible
- ◯ Slightly accessible
- ◯ Not accessible

17. **What challenges exist in maintaining transparency in AI operations?** (If you select "Other", please specify!)

- ◯ Technical complexity
- ◯ Proprietary concerns
- ◯ Lack of standardized reporting
- ◯ Insufficient regulatory requirements
- ◯ Other

18. **How does the industry ensure accountability for the outcomes of AI-driven decisions?** (If you select "Other", please specify!)

- ◯ Designated accountability officers
- ◯ Clear escalation procedures
- ◯ Oversight by regulatory bodies
- ◯ Transparency reports
- ◯ Other

19. **Have there been instances where a lack of transparency or accountability in AI decisions led to public concern? Please describe.**
(Yes/No, with a brief description if Yes)

20. **How are ethical concerns raised by stakeholders addressed in an AI development and deployment?** (If you select "Other", please specify!)

- ◯ Feedback mechanisms
- ◯ Regular stakeholder meetings
- ◯ Public consultations
- ◯ Ethics review boards
- ◯ Other

Bias and Fairness

21. **What role do third-party audits play in maintaining transparency and accountability?** (If you select "Other", please specify!)

- ◯ Essential for independent verification
- ◯ Useful but not critical
- ◯ Not currently used
- ◯ Other

22. **How does the industry handle ethical dilemmas arising from AI decision-making?** (If you select "Other", please specify!)

- ◯ Ethics committees
- ◯ Consultation with external experts
- ◯ Internal review processes
- ◯ Public disclosure and debate
- ◯ Other

23. **What improvements can be made to enhance transparency?** (Open-ended)

24. What improvement can be made to improve accountability in AI? (Open-ended)

25. **What measures are taken to identify and mitigate biases in AI systems in the utility sector?** (Select all that apply, and if you select "Other", please specify!)

- [ ] Bias audits
- [ ] Use of diverse datasets
- [ ] Inclusion of diverse perspectives in development teams
- [ ] Regular testing and monitoring
- [ ] Other

26. **How effective are these measures in reducing bias?**

- ◯ Very effective
- ◯ Effective
- ◯ Neutral
- ◯ Ineffective
- ◯ Very ineffective

27. **What types of biases are most commonly observed in AI systems?** (If you select "Other", please specify!)

- ◯ Gender bias
- ◯ Racial/ethnic bias
- ◯ Socio-economic bias
- ◯ Geographical bias
- ◯ Other

28. **How are decisions made by AI systems evaluated for fairness?** (If you select "Other", please specify!)

- ◯ Fairness metrics
- ◯ Stakeholder feedback
- ◯ Regular audits
- ◯ Comparison against human decisions
- ◯ Other

29. **What standards are used to ensure fairness in AI-driven decisions?**(If you select "Other", please specify!)

- ◯ International standards (e.g., ISO)
- ◯ National regulations
- ◯ Industry-specific guidelines
- ◯ Company policies
- ◯ Other

30. **How does the industry ensure that AI decisions do not disproportionately impact vulnerable populations?**
(Open-ended)

31. **Can you provide an example where a bias in an AI system was identified and addressed?** (Yes/No, with a brief description if Yes)

Ethical Considerations and Future Directions

32 **How frequently are AI systems reviewed for bias and fairness?**

- ◯ Regularly (e.g., quarterly)
- ◯ Annually
- ◯ As needed
- ◯ Not regularly reviewed

33. **How does the industry engage with external experts or stakeholders to assess bias and fairness?** (If you select "Other", please specify!)

- ◯ Regular consultations
- ◯ Collaborative research
- ◯ Public forums
- ◯ Ethics panels
- ◯ Other

34. **What additional measures could enhance fairness and reduce bias in AI systems?**

35.

1. **What ethical challenges does the utility industry foresee with the increasing use of AI?**
   (Select all that apply, and if you select "Other", please specify!)

. t i

- [ ] Privacy concerns
- [ ] Security risks
- [ ] Bias and fairness
- [ ] Transparency and accountability
- [ ] Job displacement
- [ ] Other

36. **How is the industry preparing to address these future ethical challenges?**
(Select all that apply, and if you select "Other", please specify!)

- [ ] Developing new ethical guidelines
- [ ] Investing in ethical AI research
- [ ] Enhancing stakeholder engagement
- [ ] Strengthening regulatory compliance
- [ ] Other

37. **How can ethical considerations be better integrated into AI Research and development?**
(Select all that apply, and if you select "Other", please specify!)

- [ ] Involving ethicists in project teams
- [ ] Regular ethics training
- [ ] Ethical impact assessments
- [ ] Public engagement initiatives
- [ ] Other

. t i

38 **How impor ant is it for the utility sector to collaborate with other ndustries in setting ethical AI standards?**

- ◯ Very important
- ◯ Important
- ◯ Neutral
- ◯ Slightly important
- ◯ Not important

39. **What role should government and regulators play in shaping the ethical landscape of AI?** (If you select "Other", please specify!)

- ◯ Strong regulatory oversight
- ◯ Providing ethical guidelines
- ◯ Supporting ethical AI research
- ◯ Facilitating industry collaboration
- ◯ Other

40. **How can the industry ensure continuous improvement in ethical AI practices?** (Select all that apply, and if you select "Other", please specify!)

- ☐ Ongoing training and development
- ☐ Regular audits and reviews
- ☐ Open dialogue with stakeholders
- ☐ Adopting new technologies responsibly
- ☐ Other

. t i

41. **What are the potential benefits and risks of AI adoption in the utility sector?** (If you select "Other", please specify!)

- ◯ Improved efficiency vs. data privacy risks
- ◯ Cost savings vs. job displacement
- ◯ Enhanced decision-making vs. potential bias
- ◯ Other

. t i

42 **How does he industry balance innovation with ethical considerat ons in AI deployment?** (If you select "Other", please specify!)

- ◯ Innovation driven by ethical frameworks
- ◯ Prioritizing ethical considerations over innovation
- ◯ Finding a middle ground
- ◯ Other

43. **What steps can be taken to improve stakeholder engagement in ethical AI deployment?** (Open-ended)

44. **How can ethical AI practices be standardized across the industry?** (Open-ended)

45. **How do you see the future of AI ethics in the South African utility industry?** (Open-ended)

46. **How can AI promote sustainable and equitable energy practices?** (If you select "Other", please specify!)

- ◯ Optimizing resource use
- ◯ Reducing emissions
- ◯ Improving access to energy
- ◯ Other

. t i

47 **What key e hical principles should guide AI deployment in the uti ity sector? (**If you select "Other", please specify!)

- ◯ Respect for privacy
- ◯ Fairness and non-discrimination
- ◯ Transparency and accountability
- ◯ Safety and security
- ◯ Other

48. **How does the industry measure the impact of ethical AI practices on business outcomes?**
(Open-ended)

49. **What are the main obstacles to achieving ethical AI implementation?**
(Select all that apply, and if you select "Other", please specify!)

- ☐ Lack of resources
- ☐ Rapid technological change
- ☐ Inadequate regulatory frameworks
- ☐ Resistance to change
- ☐ Other

50. **How can the industry build public trust in AI-driven decisions?**
(Select all that apply, and if you select "Other", please specify!)

- ☐ Transparency in decision-making processes
- ☐ Engaging with the public
- ☐ Demonstrating benefits of AI
- ☐ Addressing ethical concerns
- ☐ Other

. t i

APPENDIX D

Analysis tables and charts:

Table: Coding and Grouping for AI Ethics Analysis		
Category	**Themes**	**Frequency**
Ethical Guidelines Adoption	Adoption of AI Ethics Frameworks, Communication of Guidelines, Inconsistency in Application	20
Transparency and Accountability	Transparency in AI Decision-making, Third-party Audits, Technical Complexity, Regulatory Oversight	25
Bias and Fairness	Bias Mitigation, Fairness in AI, Stakeholder Feedback, Bias Audits	18
Future Ethical Concerns	Ethical AI Research	10

Table 16 : Coding and Grouping for AI Ethics Analysis

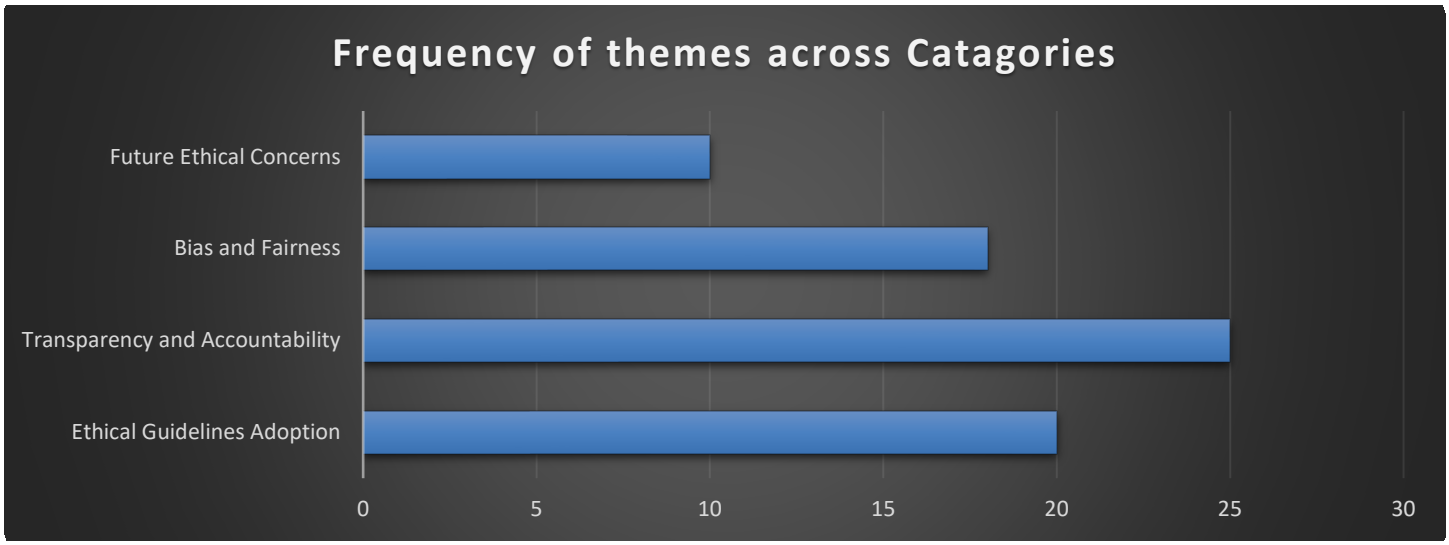


Figure 8 : Frequency of themes across categories

. t i



Figure 10 : Accountability Improvement in AI – Word-cloud

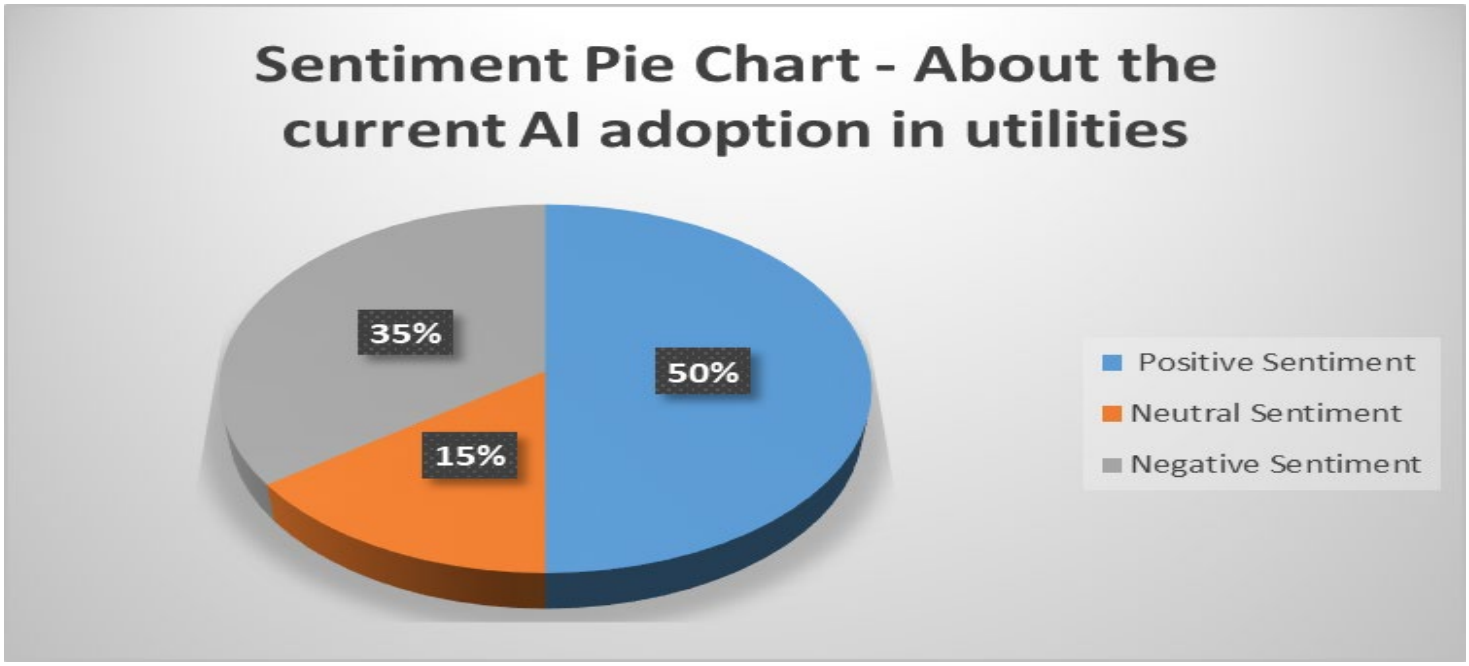


Figure 9 : Sentiment Pie Chart - About the current AI adoption in utilities

. t i

APPENDIX E

Ethics approval notification - Certificate

Graduate School of Business Administration
University of the Witwatersrand, Johannesburg

Wits Business School

Wits Business School Ethics Committee
Constituted under the University Human Research Ethics Committee (Non-Medical)

Ethics Clearance Certificate

Ethics protocol number: WBS/DB744853/960
This certificate is only valid with a legitimate ethics protocol number and signed by the Researcher (below).

Project title	AI ethical readiness within energy utilities sector in South Africa
Investigator / Researcher	Mr Isaiah Masipa
Nature of Project	MM (Digital Business)
Decision of the Committee	Approved, provided stakeholders and participants are guaranteed confidentialty.
Issue Date of Certificate	01/11/2024
Expiry date	Date of submission of the project / research report
Chairperson	Dr Ayanda Magida ☎ +27 11 717 3953 ✉ ayanda.magida@wits.ac.za

Declaration by Researcher

One copy must be signed by the Researcher and returned to the Chairperson of the Wits Business School Ethics Committee.

I fully understand the conditions under which I am 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 undertake to resubmit the protocol to the Committee.

Signature

02 / 11 / 2024
Date:

. t i

APPENDIX F

Title Approval - AI Readiness within Energy sector in South Africa.

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Private Bag 3 Wits, 2050
Fax:
Tel:

Reference: Ms Jennifer Mgolodela
E-mail: jennifer.mgolodela@wits.ac.za

16 January 2025
Person No: 744853
PAG

Mr IM Masipa
Unit 7 Village 1 Stone Arch Estate, Sunstone rd
Castleview
1401
South Africa

Dear Mr Isaiah Masipa

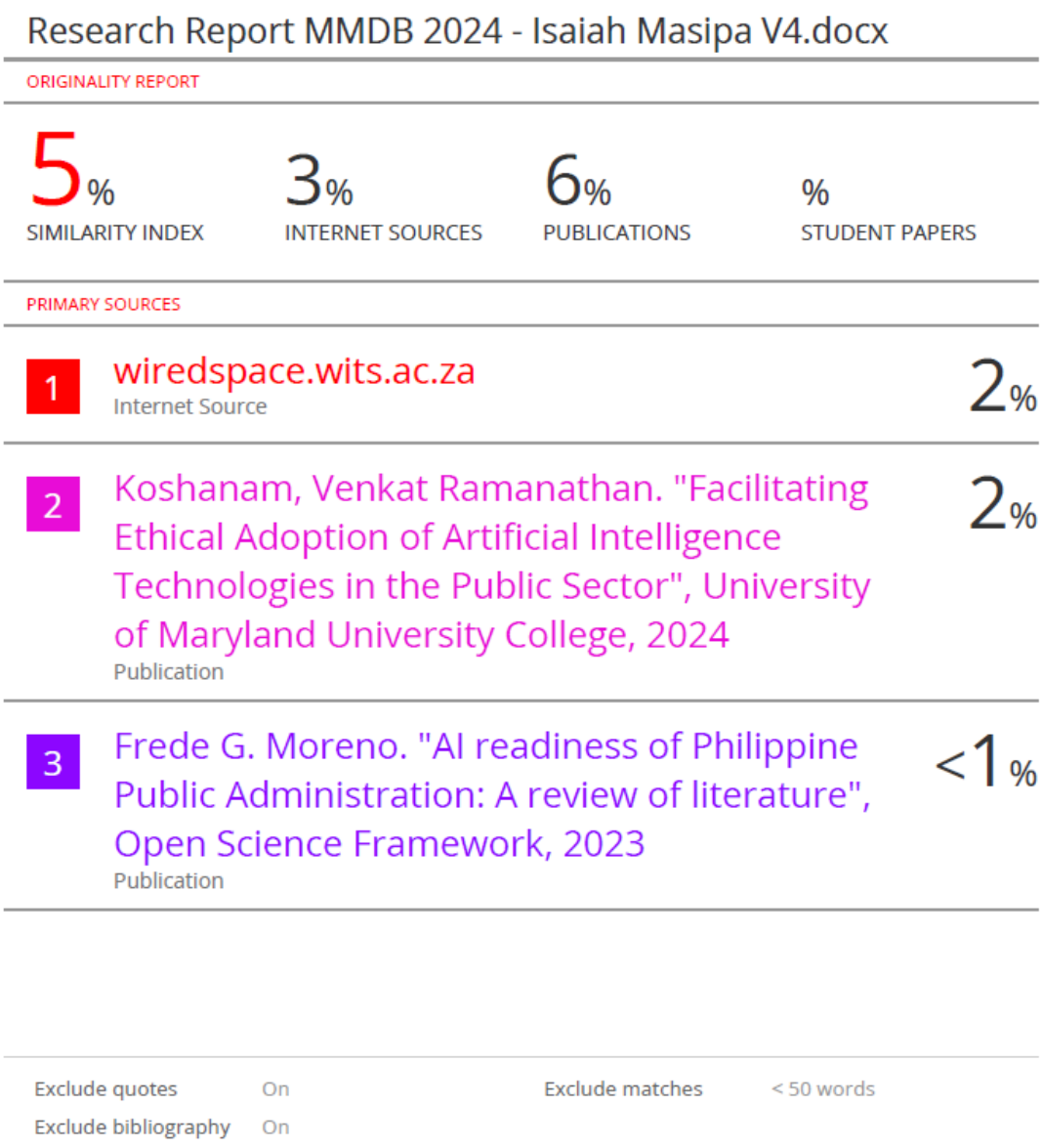
Master of Management: Approval of Title

We have pleasure in advising that your proposal entitled *AI ethical readiness within energy utilities sector in South Africa* has been approved. Please note that any amendments to this title have to be endorsed by the Faculty's higher degrees committee and formally approved.

Yours sincerely

Mrs Marike Bosman
Faculty Registrar
Faculty of Commerce, Law and Management

APPENDIX F



Research Report MMDB 2024 - Isaiah Masipa V4.docx

ORIGINALITY REPORT

5%	3%	6%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	wiredspace.wits.ac.za Internet Source	2%
2	Koshanam, Venkat Ramanathan. "Facilitating Ethical Adoption of Artificial Intelligence Technologies in the Public Sector", University of Maryland University College, 2024 Publication	2%
3	Frede G. Moreno. "AI readiness of Philippine Public Administration: A review of literature", Open Science Framework, 2023 Publication	<1%

Exclude quotes	On	Exclude matches	< 50 words
Exclude bibliography	On		

END