

https://dspace.org/

Perceptions of artificial intelligence (AI) professionals in South Africa on the ethical implications of AI

Siaga, Sinah

2025
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/47917
Siaga, Sinah. (2025). Perceptions of artificial intelligence (AI) professionals in South Africa on the ethical implications of AI [Master`s dissertation, University of the Witwatersrand, Johannesburg]. WIReDSpace. https://hdl.handle.net/10539/47917

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

Perceptions of artificial intelligence (AI) professionals in South Africa on the ethical implications of AI

Sinah Siaga

Supervisor: Prof Tendani Mawela

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, 2025

C2 General

ABSTRACT

This research explores the ethical implications of artificial intelligence (AI) adoption in South Africa, focusing on the private sector. Using a qualitative approach rooted in interpretivism and guided by thematic analysis, data was collected through semi-structured interviews with industry professionals. The study aimed to uncover how AI professionals perceive and respond to ethical considerations in the design and deployment of AI technologies across various industries.

Key ethical challenges identified include algorithmic bias, lack of transparency, data privacy concerns, and unequal access to AI technologies. Participants highlighted inconsistencies in ethical awareness, an absence of formalised guidelines, and a gap between personal values and institutional priorities. The study also found that regulatory ambiguity and profit-driven goals often overshadow ethical intentions, despite professionals recognising the societal impact.

The research proposes practical recommendations, including the establishment of inclusive AI regulatory frameworks, structured ethics training for developers, and the development of transparent governance models tailored to the South African context. These findings contribute to a deeper understanding of responsible AI integration and provide a foundation for future policy, education, and industry-wide ethical standards.

Keywords: AI Ethics, Privacy, Bias, Transparency, Accountability, Ethical AI, Consequentialist Ethics, Virtue Ethics, Regulatory Frameworks

C2 General

DECLARATION

I, Sinah Siaga , 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: Sinah Siaga Signature: 

Signed at Midrand…

On the 28th…day of February… 2025

C2 General

DEDICATION

I thank God for his blessings and guidance throughout the journey. This research paper is dedicated to my family, whose love, support, and unwavering belief in me inspire me every day. To my husband and sons, Phindulo, Phathutshedzo, and Andisani, you are my greatest motivation, reminding me that every challenge is an opportunity to learn and grow. I am deeply grateful for your love, encouragement, and the joy you bring into my life.

C2 General

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere gratitude to my supervisor, Prof Tendani Mawela, for her invaluable knowledge and guidance throughout this journey. Your unwavering support and expertise have challenged me to refine my work and elevate my thinking. I am also grateful to my employer for granting me the time and support to complete my studies.

To my dear friend and fellow master's researcher, I truly appreciate the support and words of encouragement that kept me motivated along the way. Lastly, thank my family and friends for their unwavering love, patience, and support. I am truly grateful.

C2 General

TABLE OF CONTENTS

LIST OF TABLES 10

LIST OF FIGURES 11

LIST OF ACRONYMS 12

CHAPTER 1. INTRODUCTION 13

1.1 Statement of Purpose 13
1.2 Background of the Study 13
1.3 Research Problem 14
1.4 Gaps Identified in the Study 17
1.5 Research Questions 18
1.6 Rationale 18
1.7 Delimitations of the Study 19
1.8 Definition of Terms 19
1.9 Assumptions 19

CHAPTER 2. LITERATURE REVIEW AND THEORETICAL FRAMEWORK 21

2.1 Introduction 21
2.2 AI Ethics Literature Overview 21
2.3 Digital Transformation and Emerging Disruptive Technologies 22
2.4 What is Artificial Intelligence 23
2.4.1 Types of AI 23
2.4.2 Opportunities for AI in Organisations and Society 24
2.4.3 Challenges of AI in Organisations and Society 25
2.5 Ethics and Values in Technology Design and Development 25
2.6 Ethics and Values in AI Design and Development 26
2.7 African Values and Principles in Technology 26
2.8 AI Professionals' Awareness 27
2.9 Frameworks for the Ethical Design, Development, and Use of AI. 28
2.10 Research Propositions 29
2.10.1 Proposition 1 29
2.10.2 Proposition 2 30

C2 General

2.10.3 Proposition 3 30
2.11 Analytical Framework 30
2.11.1 Theoretical Framework 30
2.11.2 Linking findings to Theoretical Framework 34
2.11.2.1 Proposition 1 34
2.11.2.2 Proposition 2 35
2.11.2.3 Proposition 3 35
2.11.2 Conceptual Framework 36
2.12 Conclusion of Literature Review 37

CHAPTER 3. RESEARCH METHODOLOGY 39

3.1 Research Paradigm 39
3.2 Research Approach 39
3.3 Research Design 40
3.4 Data Collection Methods 40
3.5 Population and Sample 41
3.5.1 Population 41
3.5.2 Sample and Sampling Method 41
3.6 The Research Instrument 43
3.7 Data Analysis Strategies and Interpretation 45
3.8 Possible Limitations and Challenges of the Study 46
3.9 Quality Assurance 46
3.9.1 Transferability 46
3.9.2 Credibility 47
3.9.3 Dependability 47
3.9.4 Confirmability 47
3.10 Ethical Considerations 48

CHAPTER 4. PRESENTATION OF RESULTS / FINDINGS 49

4.1 Introduction 49
4.2 Demographics of Participants 49
4.3 Concept Analysis 52
4.4 Themes 53
4.5 Results Pertaining to Proposition 1 53
4.5.1 Awareness and Perceptions of AI Ethics 54
4.5.2 Enablers for Integrating Ethics into AI Design 54
4.5.3 Barriers to Ethical AI Design 55
4.6 Results Pertaining to Proposition 2 56
4.6.1 AI Professionals' Awareness and Perceptions of AI Ethics 56
4.6.2 The Role of AI Professionals in Ethical AI Deployment 57
4.6.3 Enablers for Ethical AI Integration 58
4.7 Results Pertaining to Proposition 3 59
4.7.1 The Role of Ethical Guidelines in AI Development 59

C2 General

4.7.2 The Impact of Regulatory Frameworks on AI Ethics 60
4.7.3 Enablers for Ethical AI Development 61
4.8 Consequentialist and Virtue Ethics Theory 62
4.9 Summary of the Results/Findings 64

CHAPTER 5. DISCUSSION OF THE RESULTS OR FINDINGS .66

5.1 Introduction 66
5.2 Discussion Pertaining to Proposition 1 67
5.2.1 Personal Beliefs 67
5.3 Discussion Pertaining to Proposition 2 69
5.3.1 Awareness of AI Ethics 69
5.3.2 Perceptions of AI Ethics 71
5.4 Discussion pertaining to Proposition 3 73
5.4.1. Regulatory Framework 73
5.4.2. Challenges in Integrating Ethics 75
5.4.3. Enablers for Integrating Ethics 77
5.4.4. Collaboration of Stakeholders 79
5.5 Theory Driven Interpretation of findings 81
5.6 Conclusion 82

CHAPTER 6. CONCLUSIONS & RECOMMENDATIONS 84

6.1 Introduction 84
6.2 Addressing the Research Questions 84
6.2.1 Conclusions regarding the main research question 84
6.2.2 Conclusions regarding research sub-question 1 86
6.2.3 Conclusions regarding research sub-question 2 88
6.3 Research Contributions 89
6.4 Limitations of the Study 89
6.5 Recommendations 90
6.6 Suggestions for Further Research 91
6.6.1 Developing Ethical Frameworks for AI 91
6.6.2 Exploring the Role of AI in Specific Ethical Domains 92
6.6.3 Impact of Training and Education on AI Ethics 92
6.6.4 The Influence of Organisational Culture on AI Ethics 92
6.6.5 Future Technologies and Emerging Ethical Challenges 93
6.7 Implementation Guidance and Evaluation 93
6.8 Conclusion 94

C2 General

REFERENCES 95

APPENDIX (A) Semi-structured interview guide.....102

APPENDIX (B) Consent Form.....103

APPENDIX (C) Participation Information Sheet 104

APPENDIX (D) Ethics Approval Letter.....106

APPENDIX (E) Turnitin Report 107

C2 General

LIST OF TABLES

Table 1: Consequentialist and Virtue Ethics Table 63

Table 2: Summary of findings for the main research question 85

Table 3: Summary of findings for research sub-question 1 86

Table 4: Summary of findings for research sub-question 2 88

Table 5: Table of Recommendations 90

Table 6: Implementation Guidance and Evaluation Table 94

C2 General

LIST OF FIGURES

Figure 1: Proposed Framework that Combines Consequentialist and Virtue Approach to Ethical AI Design and Development. 36

Figure 2: Participants Experience 50

Figure 3: Participants Industry 50

Figure 4: Participants Gender 51

Figure 5: Employment Category 51

Figure 6: Concept Analysis 52

C2 General

LIST OF ACRONYMS

- ACM FAccT – Association for Computing Machinery Conference on Fairness, Accountability and Transparency
- AI – Artificial Intelligence
- FEAT – Fairness, Ethics, Accountability, and Transparency
- GDPR – General Data Protection Regulation
- IEEE – Institute of Electrical and Electronics Engineers
- NSF – National Science Foundation
- POPIA – Protection of Personal Information Act
- SA – South Africa
- UNESCO – United Nations Educational, Scientific and Cultural Organisation

C2 General

CHAPTER 1.INTRODUCTION

1.1 Statement of Purpose

The study aimed to investigate AI professionals' perceptions of the challenges related to ethical AI design and development across various sectors in South Africa.

1.2 Background of the Study

Governments and businesses strategically adopt emerging technologies such as artificial intelligence (AI) to improve productivity and increase efficiency. Artificial intelligence (AI) advancements can accelerate digital transformation across various sectors and socio-economic activities. The more economically developed countries are geared up to benefit from AI despite the threat of job losses. Countries in Africa need to address their market challenges relating to governance restrictions to harness the adoption of AI (Arakpogun et al., 2021).

The involvement of international communities in bridging the technological gaps in the global south can assist and support approaches to tackle the digital divide and technology disparities in Africa. The AI policies and strategies in Africa must be robust and tailor-made to cater to the African context and fully embrace the benefits of AI. A problem-oriented approach will allow the African countries to focus on their specific needs instead of adopting what has been implemented in the global north (Arakpogun et al., 2021).

According to Kiemde and Kora (2022), PWC suggests that the value added by AI technologies in Africa is $1.2 trillion. Some of the development challenges in Africa can be explored and solved through AI-based solutions. However, it is noted that most of the digital infrastructure in Africa is controlled and funded by international communities outside the continent, which can result in biases, unfairness, and unintended consequences of AI systems in society.

C2 General

Ayling and Chapman (2022) highlight that the misuse of predictive models and AI technologies is gaining traction across various domains, which include businesses, governments, and academia. The General Data Protection Regulation (GDPR) currently does not address AI technologies' societal and ethical issues (Ayling & Chapman, 2022). South Africa adopted the Protection of Personal Information Act (POPIA), which addresses ethical issues governing personal data collection, use, storage, and sharing.

1.3 Research Problem

The presence of bias and a lack of transparency and accountability resulting from the use of Artificial Intelligence (AI) systems raise ethical concerns and unintended consequences in data-driven innovation in society across various sectors, including decision-making resulting from the misuse of predictive models (Huriye, 2023).

This is a problem because it can reinforce societal inequalities resulting from biases in AI systems. According to Huriye (2023), the main ethical challenges in the development and use of AI are privacy, bias, transparency, and accountability. There is a need for collaboration among AI practitioners, policymakers, and local communities to ensure ethical policies and guidelines are documented and used in implementing AI systems. The African economic, political, and cultural context must be considered when creating policies. Adopting what has been implemented in Western countries could result in unfairness and biases in the African context.

Ethics is critical in designing and implementing AI technologies to minimise harmful effects and ensure positive outcomes when using AI systems. Ethical values must be in place to avoid discrimination and injustice in AI systems (Marwala, 2024). Ethical values can mitigate biases leading to discrimination by adhering to AI development principles that promote equality and fairness (Van Norren, 2022). Compliance and user trust can be achieved through the ethical

C2 General

use of AI systems. The ethical use of AI will enable human capabilities to be enhanced through technology without compromising human rights, justice, and privacy.

The problem of ethical AI design and implementation impacts various stakeholders, which include communities, minorities, public and private organisations, government, policymakers, regulators, technology companies, and the general society. All stakeholders who interact with the AI systems need to trust that the AI systems are transparent and fair and that their rights will not be compromised due to the use of the AI systems (Borenstein & Howard, 2021).

AI is integrated into our daily personal and professional lives, and the COVID-19 pandemic accelerated this. Trust and privacy are some of the ethical issues raised by using AI (Hamilton, Swart, & Stokes, 2021).

AI systems are sometimes used to monitor citizens' activities. For instance, organisations can use social media data to inform their decisions during the recruitment process. This can raise concerns about ethical implications and biased decision-making. AI systems are also designed and developed by humans, who can make errors. Bias needs to be addressed in the data and algorithms, as this will address the issue of biased outcomes (Borenstein & Howard, 2021).

To address the root cause of the problem, AI developers should consider the ethical aspects of the solutions (McLennan et al., 2022). Additionally, citizens must educate themselves about AI and understand how data is used in AI, even if they are not developers. This process can ensure that ethical principles are embedded in AI systems' design, development, and use. Various examples of AI misuse and unethical AI have been reported; for example, biases were identified in facial recognition technology where the users were not white, and this resulted in the wrongful arrest of a black man due to low accuracy rates (Borenstein & Howard, 2021).

C2 General

From a policy perspective, the ethical challenges in AI are being addressed using different initiatives, which include a comprehensive and applicable set of ethical principles to ensure that AI in education is developed with ethical considerations at its core (Nguyen et al., 2023). AI ethics have been documented. However, the impact of these standards in terms of new regulations is still unclear. Additionally, the ethics of AI have been documented by the IEEE, which is developing AI ethical standards. Ethical principles considered important by organisations such as Google and Deloitte include issues related to transparency and fairness of solutions (Borenstein & Howard, 2021).

From a scholarly perspective, ACM FAccT conferences have been created to focus on AI ethics-related issues. Fairness, Ethics, Accountability, and Transparency (FEAT) is supported by organisations like the National Science Foundation (NSF) (Borenstein & Howard, 2021). Also, Kiemde and Kora (2022) conducted a study on AI ethics in South Africa. They suggested that researchers in Africa should be involved in developing AI-based data sets to ensure awareness of the development process. Researchers have proposed solutions for ethical AI in Africa; however, these solutions are difficult to implement without training and proper integration of ethics. Education ensures that all stakeholders embrace AI systems' responsible and ethical use (Kamalov et al., 2023).

Additionally, UNESCO conducted a survey in Africa on AI, and 26 countries reported a skills gap in dealing with AI ethical implications (Kiemde & Kora, 2022). Education and training programs incorporating AI ethics can support the awareness and upskilling of members of AI communities and inform them of the dangers and impact of AI.

The study focused on AI solution designers, professionals, and implementers. It investigated the awareness and perceptions of this key stakeholder group regarding the factors that influence the integration of ethical principles in AI solutions.

C2 General

While AI presents transformative opportunities, its application in South Africa raised ethical concerns such as bias, privacy violations, lack of accountability, and opaque decision-making (Gwagwa et al., 2022). These risks were exacerbated by the absence of locally grounded ethical frameworks and underdeveloped regulatory oversight (Department of Communications and Digital Technologies, 2024). The problem is particularly pressing in the South African context, where the direct adoption of Western ethical standards may not address socio-cultural dynamics, thus risking unfair or discriminatory outcomes.

Existing literature documents ethical concerns broadly but offers limited insight into how AI professionals, who key stakeholders in AI creation perceive and navigate these ethical challenges. This study addressed this gap by exploring AI professionals' awareness, values, and behaviours related to ethical AI practices, with the aim of informing context-relevant, theory-driven solutions.

1.4 Gaps Identified in the Study

South Africa has begun to explore AI policy direction through documents like the Draft National AI Policy Framework (DCDT, 2024). However, this framework remains high-level and lacks detailed ethical guidance or mandatory implementation protocols. The absence of contextualised ethical frameworks makes it difficult for AI professionals to apply ethical principles practically in their work.

The gaps identified include the lack of empirical studies focusing on the ethical decision making of AI practitioners in South Africa and limited integration of ethical theory where virtue and consequentialist ethics are part of the analysis of AI development practices. The AI ethics guidelines are dominated by western values which are not global (Kiemde & Kora, 2022).

This research directly addresses these shortcomings by investigating the perceptions of AI professionals in South Africa through qualitative interviews. Applying virtue and consequentialist ethics to analyse how these professionals

C2 General

understand and respond to ethical challenges in practice, to generate practical insights and recommendations that could inform the implementation of ethical AI policies and contribute to a context specific ethical framework aligned with South African realities.

1.5 Research Questions

This study focused on the South African context and aimed to address the ethical considerations of AI by answering the following main research question:

- What are the ethical considerations for the design and development of AI technology in South Africa?

The following sub-questions support the above question:

I. What are the AI professionals' awareness and perceptions of AI ethics?
II. What are the enablers for integrating ethics into the design and development of AI solutions in South Africa?

1.6 Rationale

In South Africa, organisations are adopting AI to drive digital transformation. Although the deployment of AI in South Africa offers various opportunities, it also presents ethical challenges across various sectors. African countries face unique challenges when adopting emerging technologies compared to developed countries. Africa’s specific challenges need to be addressed to fully harness the potential of AI (Arakpogun et al., 2021). The study aimed to provide valuable insights into the ethical considerations of AI in South Africa to support AI professionals.

South Africa does not have policies that govern the use of AI. The adoption of emerging technologies is influenced by international standards, which can result in biases and ethical dilemmas in AI systems (Kiemde & Kora, 2022). Therefore, it was important to investigate how AI professionals could influence the adoption

C2 General

of ethical AI in the design and development of AI technologies, taking into consideration the South African context.

1.7 Delimitations of the Study

The scope focused on the South African context and included respondents who are professionals working in the design, development, and deployment of AI-based technology solutions.

1.8 Definition of Terms

AI: AI has evolved progressively into intelligent machines and algorithms that can reason and adapt based on sets of rules and environments that mimic human intelligence (McCarthy, 2007). Wang (2019) broadened the definition of AI, which can perform cognitive tasks, particularly learning and problem-solving, with exciting technological innovations such as machine learning, natural language processing, and neural networks (Zawacki-Richter et al., 2019).

Ethical AI: Ethical AI, also known as responsible AI, is the practice of using AI with good intentions to empower employees and businesses and moderately impact customers and society (Eitel-Porter, 2020).

Responsible AI: Responsible AI enables companies to engender trust and confidently scale AI. (Eitel-Porter, 2020).

1.9 Assumptions

I. Ethical challenges are evident in the design, development, and adoption of AI technologies in South Africa.
II. Limited AI policies and frameworks govern the design and development of AI for government and private institutions in South Africa.
III. AI professionals in South Africa are aware of ethical challenges in developing and adopting AI.

C2 General

IV. International standards and policies influence the adoption of AI in South Africa, which can result in biases, discrimination, and unfairness.

V. Ethical AI is required to build trust and credibility in the design, adoption, and use of AI in SA.

C2 General

CHAPTER 2.LITERATURE REVIEW AND THEORETICAL FRAMEWORK

2.1 Introduction

This literature review explored existing research on AI ethics in South Africa, highlighting the importance of the Protection of Personal Information Act (POPIA) in addressing ethical concerns relating to data privacy and protection. The extant literature related to the research questions was discussed in detail before exploring the theoretical framework for the study.

2.2 AI Ethics Literature Overview

According to Kiemde and Kora (2022), the AI literature emanating from Africa is limited. The AI ethics guidelines are dominated by Western values that are not global. African countries are exploring the use of AI technologies to solve development-related problems. For instance, in South Africa, AI is used in the health sector for rapid HIV testing. Countries such as Kenya and Uganda also receive international funding and grants from organisations like IBM to facilitate the development and use of AI technologies.

Africa scores very low in the key dimensions of AI: governance, data and infrastructure, and technology sector. South Africa ranked 59th in the Global Government AI 2020 ranking and was one of the top 5 countries in Africa. The other countries in the top five were Mauritius, Seychelles, Kenya, and Rwanda (Kiemde & Kora, 2022). Mauritius was the first country to develop an AI strategy in Africa and announce the establishment of a National AI strategy. The six principles for AI ethics in Africa were introduced by Arthur (2019) and include AI risks, opportunities, and African AI values to practice responsible AI. He recommends including ethical and socio-cultural values to manage the risk of AI adoption in Africa due to the limitations posed by global AI policies.

C2 General

Sallstrom et al. (2019) recommend that policymakers promote ethics by design through regulations. The authors emphasise that not all AI is the same, and AI policies must be detailed to reflect the values and cultures of each African nation. Researchers have proposed solutions for developing ethical AI in Africa; however, such policies are difficult to implement without training the key stakeholders involved. Education will be crucial in creating awareness of AI issues and promoting AI ethics on the continent.

Lottu et al. (2019) discuss adopting an ethical framework to guide and facilitate ethical AI development in IT systems. The framework must be implemented in every state of the AI development lifecycle. It focuses on the socio-technical impact on individuals and societies. The key components of the ethical framework are privacy, transparency, fairness, privacy, security, and accountability. The framework can be applied to diverse backgrounds to mitigate biases and discrimination.

Pant et al. (2024) argue that AI technologies are as ethical as the people developing the technologies. It is important to include the view of AI practitioners in ethical AI development as this can help to enhance the development of ethical AI. According to some AI practitioners, it is their responsibility to guarantee ethical AI development and take responsibility for any harm from AI systems. Others emphasise the necessity for precise definitions of accountability in AI development and propose that practitioners and users bear some of this obligation (Novelli et al., 2023).

2.3 Digital Transformation and Emerging Disruptive Technologies

Digital Transformation drives various business segments and is crucial in providing organisations with a competitive advantage in the market (Oludapo et al., 2024). Digital transformation can drive organisational agility through exploiting

C2 General

and exploring digital technologies to achieve organisational agility. Digital strategies can help organisations to respond to new opportunities in the digital economy (Kraus et al., 2021).

Emerging disruptive technologies, such as blockchain technology and the Internet of Things, can transform society, the economy, and daily life. These technologies challenge existing norms and regulations due to their rapid pace of innovation, offering both opportunities and risks. Governments and regulators must balance economic growth and improved quality of life with safety, privacy, and security risks. Addressing these challenges requires risk management, uncertainty, and effective regulatory responses (Taeihagh et al., 2021). One of the emerging and disruptive technologies that organisations are considering is artificial intelligence (AI).

2.4 What is Artificial Intelligence

Artificial Intelligence (AI) is the simulation of human intelligence processes by machines. The processes focus on learning, reasoning, and self-correction. AI uses machine learning, neural networks, natural language processing, and robotics to perform tasks typically requiring human intelligence. Artificial Intelligence (AI) comprises various technologies and methodologies, such as robotics, machine learning, neural networks, and natural language processing. These technologies enable robots to perform tasks that conventionally need human intelligence (Berente et al., 2021).

2.4.1 Types of AI

According to Berente et al. (2021), there are four types of Artificial Intelligence (AI) based on their functionalities and capabilities: Limited Memory AI, Reactive Machines, Self-Aware AI, and Theory of Mind AI. These types of AI represent a range of capabilities, ranging from basic reactive systems to self-aware entities.

C2 General

Every type of AI has unique applications and implications for various domains and sectors.

Limited Memory AI: This type of AI can investigate the past to make decisions. It learns from historical data to make informed decisions and predictions and to improve its performance. The systems can store and retrieve past information that can be applied to the decision-making process. Self-driving cars use limited memory AI to navigate through traffic safely by learning from experiences.

Reactive Machines AI: These are the most basic forms of AI that can perform tasks based on current information without memory or past experiences. These systems do not have memory or the ability to learn from past experiences or adapt to new situations. Examples include Deep Blue, IBM's chess-playing computer, which defeated world chess champion Garry Kasparov in 1997.

Self-aware AI is the most advanced form of AI with consciousness, awareness, and the ability to form representations about itself. It can form representations about itself and its internal states. Self-aware AI is speculative and has not been achieved yet.

Theory of Mind AI: This is a more advanced form of AI that can understand human intentions, emotions, beliefs, and thought processes. It can interact with humans more naturally and intuitively. While still in the research phase, it aims to be empathetic and intuitive.

2.4.2 Opportunities for AI in Organisations and Society

Stahl et al. (2021) discuss AI's opportunities, which include data processing and analysis, optimisation of choices, progress in science and knowledge generation, efficiency and productivity, personalised experiences, predictive analytics, healthcare innovations, and environmental sustainability.

C2 General

With the help of data-driven insights and solutions, AI can address complex global issues like social injustice and climate change. The capacity to evaluate enormous volumes of data can result in better-informed policy development and decision-making (Astobiza et al., 2021).

2.4.3 Challenges of AI in Organisations and Society

The challenges associated with AI are ethical concerns, risk management, data protection, impact assessments, organisational responses, trade-offs and decision-making, human rights implications, and legal and regulatory frameworks (Stahl et al., 2021).

Concerns regarding bias, discrimination, and the possibility of violating human rights are among the many ethical concerns surrounding artificial intelligence. As AI technologies grow more prevalent in decision-making procedures, it is imperative to guarantee equity and responsibility (Eke et al., 2023).

2.5 Ethics and Values in Technology Design and Development

According to Bednar and Spiekermann (2022), there is a high demand for technology design that meets ethical standards to address how technology can impact society, privacy, and behaviour manipulation. Technology designers and developers must consider ethics when creating new technologies. Technology must be considered a moral activity as it can affect human experience and people's moral decisions. Organisations have value principles as a guideline for ethical standards, which include transparency, fairness, and digital privacy. However, the effectiveness of dealing with the ethical implications of technology is questionable. Sustainable technology design ensures that environmental, ethical, and social values are considered when designing new technology. In Industry 5.0, the integration of ethics and values is important to design, implement, and use ethical and human-centred technologies (Longo et al., 2020).

C2 General

2.6 Ethics and Values in AI Design and Development

The ethical design involves inculcating ethical principles into the design and development of systems. Ethical AI design ensures that the design of AI systems incorporates ethical values that align with human rights and promotes fairness, accountability, and transparency. An ethical AI design considers the possible effects of technology on society. The objective of ethical design is to develop AI systems that are reliable, advantageous, and consistent with societal norms and values. Integrating AI ethics in the design process can promote responsible AI, mitigate potential risks, and address biases (Kieslich et al., 2022).

Ethical AI development incorporates ethical values and principles in designing, developing, and deploying artificial intelligence solutions. The social impact of AI systems must be considered to align with human rights, transparency, and fairness (Fukuda-Parr & Gibbons, 2021). The ethical development of AI ensures that AI technologies integrate ethical design and deployment. When ethical AI development is implemented, technologies are used fairly and responsibly (Brendel, 2021).

2.7 African Values and Principles in Technology

Kiemde and Kora (2022) discuss how African values and principles in technology consider ethics, cultural relevance, and community relevance in developing and applying technology applications. African technological values and principles are ubuntu, inclusivity, respect for cultural diversity, sustainability, and ethical considerations. Ubuntu is a traditional African philosophy and ethical concept based on humanity and interconnectedness.

Important African values and concepts exist, most notably the idea of Ubuntu, which emphasises community and interconnectivity and is summed up in the statement "a person is a person through other persons. " The common good, relational autonomy, solidarity, generosity, and cultural contextualisation are

C2 General

among the values highlighted. These ideals are crucial to creating ethical frameworks for AI that align with African societies' social and cultural reality (Gwagwa et al., 2022).

2.8 AI Professionals' Awareness

According to the study by Pant et al. (2024), AI professionals show moderate ethical awareness. AI professionals were most aware of "Privacy protection and security," which they knew about 64% of the time. "Human-centred values" and "reliability and safety" came next, each acknowledged by 53% of the participants. The awareness levels for other concepts, including "Accountability," "Fairness," and "Transparency and explainability," were 42%, 43%, and 45%, respectively. However, only 12% of participants knew every ethical concept, and 2% had never heard of any.

AI professionals frequently lack the knowledge and expertise to create morally and responsibly designed AI-based solutions. Numerous academics have determined that the problem of creating ethical AI-based systems is poorly understood, and professionals lack the necessary procedures, resources, and knowledge to deal with these challenges efficiently (Pink et al., 2024).

The challenges that AI professionals encounter in maintaining responsible AI values include the complexity of these values and the weight placed on individual professionals. Only a few experts are typically responsible for maintaining ethical AI values, which can result in an unfair workload and stress. Effective implementation may be hindered by the shallow engagement that many practitioners have with ethical AI values due to a lack of in-depth information or comprehension (Varanasi & Goyal, 2023).

C2 General

2.9 Frameworks for the Ethical Design, Development, and Use of AI.

The South African Department of Communications and Digital Technologies (DCDT) published the National Artificial Intelligence (AI) Policy Framework in August 2024. This document is the cornerstone upon which South Africa's AI laws and possibly an AI Act will be built. The DCDT asked stakeholders to provide feedback by 17:00 on Friday, November 29, 2024. The South Africa National Artificial Intelligence Policy Framework articulates a strategic approach to the responsible and ethical development of AI technologies (Department of Communications and Digital Technologies, 2024).

The framework emphasises a strong commitment to ethical AI by establishing detailed guidelines that promote transparency, accountability, and fairness in AI systems. The primary goal is to build trust among users and stakeholders while reducing biases and advancing social well-being. Essential ethical principles involve designing AI applications to enhance, rather than replace, human decision-making and prioritising human-centred AI solutions. This strategy upholds professional responsibility and ensures that AI advancements align with societal values in South Africa.

The Protection of Personal Information Act (POPIA) is the data protection and privacy legislation adopted in South Africa. The act regulates the government and private entities' collection, usage, sharing, and storage of personal information. The act aims to address the key ethical challenges of data associated with the adoption of emerging technologies such as AI in South Africa.

General Data Protection Regulation (GDPR) is a comprehensive data protection law enacted by the European Union (EU) on May 25, 2018. It aims to enhance people's control over personal information and harmonise data protection laws throughout EU member states. Organisations processing personal data must abide by several standards outlined in the GDPR, including

C2 General

integrity and confidentiality, lawfulness, fairness, and transparency (Dove & Chen, 2021).

POPIA and GDPR are key legislative frameworks that directly relate to the ethical considerations surrounding AI, particularly regarding data protection and privacy. These regulations ensure that AI systems handle personal data ethically.

The European Union has proposed the Artificial Intelligence Act, or AI Act, as a legislative framework to control the supply and usage of AI systems among its member states. In April 2021, the European Commission made the initial proposal, and in December 2023, a political agreement was obtained. This law is significant because it establishes a standard framework for the safe and moral application of AI technology, making it the first legally enforceable global horizontal legislation on AI (Madiega, 2021).

Ashok et al. (2022) present a four-dimensional ontological framework for AI, focusing on physical, cognitive, information, and governance domains. The framework emphasises the need to consider digital ethics to address ethical challenges in AI use in digital technologies. Governance is identified as a critical principle in the ethical use of AI. The framework explores the interrelations between regulations and ethical considerations in AI technology development and deployment.

2.10 Research Propositions

The following propositions were derived based on the literature review:

2.10.1 Proposition 1

The integration of the consequentialist and virtue ethics theories will lead to the ethical design and deployment of AI solutions that are contextually relevant and ethically sound.

C2 General

2.10.2 Proposition 2

AI professionals play a crucial role in the ethical integration of AI into African societies. Their awareness and adherence to ethical principles significantly impact the responsible deployment of AI technologies.

2.10.3 Proposition 3

Enablers such as ethical guidelines and regulatory frameworks will facilitate the successful incorporation of ethics into the design and development of AI solutions, ensuring that AI solutions are contextually relevant and ethically sound.

2.11 Analytical Framework

2.11.1 Theoretical Framework

Different philosophers developed and influenced various significant theories of ethics. These theories include deontological, consequentialist, utilitarian, virtue, and contractualistic ethical approaches. Deontological ethics is based on rigid laws, obligations, or requirements. It emphasises abiding by universal standards that establish what is ethically acceptable or unacceptable (Hagendorff, 2020).

Consequentialist ethics is based on the notion that the morality of actions depends entirely on the consequences of those actions. Utilitarianism is one of the most used consequentialist theories. It was created by philosophers John Stuart Mill and Jeremy Bentham (Anderson & Anderson, 2020).

According to Anderson and Anderson (2020), Utilitarianism holds that a course of action is morally justified if it has the most significant potential to benefit everyone involved or cause the least harm. The theory involves performing "moral arithmetic," which entails weighing the relative benefits of different courses of action to decide which one yields the best results.

C2 General

Alternatively, the Contractualistic theory is predicated on individual agreements to construct moral norms and social contracts. It strongly emphasises equity and the notion that all parties involved should concur on moral values. (Hagendorff, 2020).

The Virtue theory strongly emphasises developing moral qualities and the moral agent's nature above laws or penalties. The theory highlights the value of moral qualities and the formation of positive character traits, arguing that moral behaviour results from being a decent person (Hogenhout, 2021).

The Consequentialist and Virtue ethical approaches underpin this research study.

Consequentialism emphasises that an action's morality is only judged by the outcomes it generates. When an action produces positive results, it is deemed morally good; when it produces negative results, it is deemed morally wrong. Utilitarianism is one of the most used consequentialist theories. It was created by philosophers John Stuart Mill and Jeremy Bentham (Anderson & Anderson, 2020). Actions that increase general happiness or usefulness are supported by utilitarianism. According to this theory, the optimal course of action is the one that benefits the most significant number of individuals (Pettit, 2000).

The consequentialist approach presents an organised method of assessing moral conduct based on results. Consequentialism is very effective when applied to clearly defined situations where possible outcomes and rules are known. However, a consequentialist approach may become less effective in real-life scenarios where information is frequently inadequate, and other actors' conduct is uncertain. The method can be further separated into two primary categories: rule utilitarianism, which analyses the implications of adhering to broad norms that regulate behaviour, and act utilitarianism, which assesses the repercussions of specific actions (Bench-Capon, 2020).

Consequentialism encourages decision-makers to broaden their horizons by placing a higher priority on the outcomes of their decisions. This is especially important in AI since the results of algorithms have the potential to have a

C2 General

significant impact on people's lives and society (Hogenhout, 2021). AI systems can be designed to function with the goal of maximising positive outcomes for the most significant possible number of people. A self-driving car, for instance, may be configured to prioritise the safety of most pedestrians and passengers, even if doing so means jeopardising a single person's life in an emergency (Anderson & Anderson, 2020).

Virtue ethics theory places more emphasis on a person's qualities and character than on the rules that must be followed or the results of an action. This method has its roots in the philosophy of Aristotle, who maintained that moral behaviour results from developing virtues, or positive character traits, which allow people to live morally good lives (Alibašić, 2023). Virtues are positive characteristics or attributes that are morally righteous. Courage, honesty, kindness, and wisdom are a few examples. People are encouraged by virtue ethics to develop these qualities since they are necessary for attaining good morals (Hogenhout, 2021).

Virtue ethics can be applied to AI in several ways to foster moral behaviour in AI systems and their interactions with people. AI systems can be designed with virtues like justice, empathy, and honesty. For example, an AI assistant might prioritise operational transparency, ensuring users know the decision-making process. This design, which embodies the principle of honesty, can encourage users to trust the system (Anderson & Anderson, 2020).

The integration of virtue ethics into AI design fosters cooperation between various stakeholders, including technologists and ethicists. Using an interdisciplinary approach, it is possible to ensure that AI systems are designed with a thorough awareness of the ethical considerations involved (Johansson et al., 2022).

The consequentialist theory can be applied to identify enablers and considerations for integrating ethics into the design and development of AI solutions in South Africa, which is one of the study's objectives. Consequentialist ethics, which evaluate the morality of actions based on their outcomes, can provide a framework for addressing the ethical considerations and enablers for

C2 General

integrating ethics into AI design and development in South Africa. AI technologies in South Africa should be designed to benefit society, reduce inequalities, and improve quality of life. AI professionals can ensure that AI systems are transparent and that their decision-making processes can be understood and scrutinised. This helps build trust and promote ethical AI in South Africa.

Consequentialist ethics can be used to pinpoint important enablers that ensure ethical issues are considered when developing AI. A thorough approach to addressing ethical considerations and enablers in South Africa's AI ecosystem can be supported by consequentialist ethics, which focuses on the implications of AI development and strives to maximise beneficial results while reducing harm.

Virtue theory encourages moral decision-making among individuals and institutions creating and implementing AI applications (Hagendorff, 2020). The theory will assist the researcher in evaluating the AI professional’s awareness and perceptions of AI ethics in South Africa. Virtue ethics encourages AI professionals to develop accountability, honesty, and integrity. If they prioritise these virtues, AI professionals are more likely to be conscious of the ethical consequences of their work and try to comprehend and address ethical issues in AI research. AI professionals are more likely to be conscious of the ethical consequences of their work and try to comprehend and address ethical issues in AI research if they prioritise these virtues.

AI professionals who uphold virtue ethics are likely to regularly reflect on the moral implications of their work, which can result in a heightened awareness and more complex understanding of AI ethics. Virtue ethics emphasise how communities influence moral behaviour. Since AI professionals are impacted by the common beliefs and behaviours of their society, a professional culture that upholds and encourages virtues would help them become more ethically conscious.

C2 General

2.11.2 Linking findings to Theoretical Framework

This section reflects on how the thematic findings of the study relate to the three theoretical propositions underpinned by consequentialist and virtue ethics. The goal is to assess whether the empirical evidence supports or challenges the assumptions underlying the ethical integration of AI in South Africa.

2.11.2.1 Proposition 1

The integration of the consequentialist and virtue ethics theories will lead to the ethical design and deployment of AI solutions that are contextually relevant and ethically sound.

Empirical Linkage and Interpretation: The data partially supported this proposition. Participants demonstrated awareness of ethical consequences, particularly related to public trust, fairness, and potential harm which are core on consequentialist ethics. For example, many interviewees spoke about the importance of avoiding bias and ensuring fairness in algorithmic design, indicating a concern for long term societal outcomes.

However, the practical application of virtue ethics was inconsistent. While some participants described personal values such as integrity, responsibility, and transparency, these virtues were often overridden by institutional priorities like speed to market and profitability. This shows that while the theoretical integration is conceptually sound, in practice, there is limited support unless reinforced by structures that enable ethical behaviour.

Conclusion: This proposition is conditionally supported. Theories alone are insufficient without operational support, culture, and accountability mechanisms to guide ethical integration.

C2 General

2.11.2.2 Proposition 2

AI professionals play a crucial role in the ethical integration of AI into African societies. Their awareness and adherence to ethical principles significantly impact the responsible deployment of AI technologies.

Empirical Linkage and Interpretation: This proposition is strongly supported. Across interviews, AI professionals expressed a keen understanding of the ethical risks associated with AI systems, including privacy violations, algorithmic bias, and lack of explainability. Participants identified themselves as “ethical gatekeepers” and emphasised the moral responsibility they feel in shaping the outcomes of AI systems.

Several professionals admitted that ethical awareness is often intuitive or based on personal beliefs rather than grounded in formal training. This gap underscores the need for structured ethics education and professional codes of conduct.

Conclusion: The findings validate this proposition. AI professionals have the potential to serve as key enablers of ethical AI integration, but their effectiveness is limited without institutionalised ethical support.

2.11.2.3 Proposition 3

Enablers such as ethical guidelines and regulatory frameworks will facilitate the successful incorporation of ethics into the design and development of AI solutions, ensuring that AI solutions are contextually relevant and ethically sound.

Empirical Linkage and Interpretation: This proposition is broadly supported by the data. Participants consistently pointed to the lack of clear regulatory guidance and ethical standards as a barrier to effective ethical practice. Many expressed a desire for context-specific frameworks that go beyond high-level principles to offer actionable guidance suited to South African realities.

C2 General

Interviewees also identified enablers such as organisational culture, leadership support, and interdepartmental collaboration as critical to implementing ethical principles. These findings align with both virtue ethics (which emphasises organisational character and leadership) and consequentialist ethics (which prioritise outcomes like fairness and harm reduction).

Conclusion: The findings affirm that enabling structures such as guidelines, frameworks, and leadership are essential to embedding ethics into AI systems in Africa. However, such mechanisms are currently underdeveloped and urgently needed.

2.11.2 Conceptual Framework

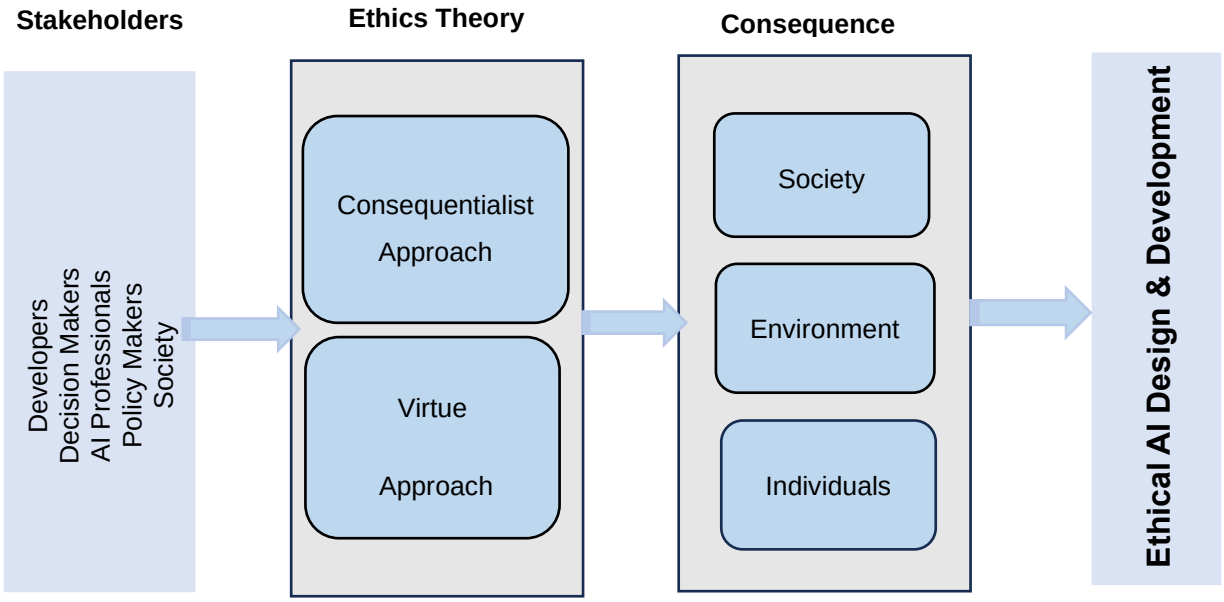


Figure 1: Proposed Framework that Combines Consequentialist and Virtue Approach to Ethical AI Design and Development.

The stakeholders are all individuals affected by AI systems, including AI professionals, developers, decision-makers, policymakers, and society. They play a pivotal role in defining and guiding the ethical considerations of AI, ensuring that diverse perspectives and needs are addressed (Hagendorff, 2020).

C2 General

The consequentialist approach evaluates the morality of AI actions based on their outcomes, emphasising the importance of maximising positive impacts and minimising harm. This approach considers the impact on society, aiming to create AI systems that enhance public safety, reduce inequality, and promote social well-being. It also considers the environmental impact, ensuring that AI technologies are developed with sustainability to prevent ecological damage (Anderson & Anderson, 2020).

The virtue approach focuses on AI professionals' moral character and virtues, fostering qualities such as honesty, empathy, and justice. This approach is crucial in promoting ethical decision-making within AI development, as it encourages AI professionals to uphold high ethical standards and reflect these virtues in their work. The virtue approach also highlights the impact on individuals prioritising user well-being, privacy, and trust. When combined with the consequentialist approach, it provides a comprehensive ethical framework where both the outcomes of AI systems and the moral integrity of their creators are considered (Hogenhout, 2021).

Together, these approaches ensure that ethical AI design and development is mindful of its broader implications, integrating stakeholders' perspectives and balancing the various impacts on society, the environment, and individuals. By considering both the results of AI actions and the virtues embodied in their creation, ethical AI systems can be designed to benefit society, promoting trust and sustainability (Johansson et al., 2022). This holistic approach ensures that AI technologies are practical, efficient, and aligned with ethical values that address the needs and concerns of all stakeholders involved.

2.12 Conclusion of Literature Review

The literature review highlighted the complexities and challenges associated with the ethical design, development, and use of AI in South Africa. It emphasises the critical role of the Protection of Personal Information Act (POPIA) in addressing

C2 General

data privacy concerns. Theoretical insights, including consequentialist and virtue ethics, provide a robust framework for addressing ethical concerns in AI development.

This review also underscores the importance of ethical frameworks tailored to African contexts, as proposed by various scholars. Such frameworks must integrate ethics throughout the AI lifecycle, addressing key components such as privacy, transparency, fairness, security, and accountability. Additionally, the role of AI professionals is pivotal, as their awareness and values significantly influence the ethical integration of AI into society.

The publication of the National Artificial Intelligence (AI) Policy Framework by the South African Department of Communications and Digital Technologies (DCDT) in August 2024 marks a significant step toward establishing a structured regulatory environment for AI in the country. This framework provides strategic direction for responsible and ethical AI development and signals the potential for a dedicated AI Act in the future. The DCDT demonstrates an inclusive approach by seeking stakeholder feedback, ensuring diverse perspectives shape the evolving AI landscape.

In conclusion, AI's ethical development and deployment in South Africa require a multifaceted approach that blends global ethical standards with African cultural values. Integrating ethics into AI is not merely a technical challenge but a socio-cultural imperative that demands the active participation of all stakeholders, including policymakers, AI professionals, and the broader society.

C2 General

CHAPTER 3.RESEARCH METHODOLOGY

This chapter describes the research approach followed to obtain insights into the perceptions of AI professionals in South Africa on the ethical implications of AI. This section examines the research methodology employed to conduct the investigation, outlines the design, sampling method, and research instruments, and includes data collection, analysis, and ethical considerations.

3.1 Research Paradigm

The research paradigm that guided this study is interpretivism. Interpretivism is a qualitative research paradigm that focuses on understanding and interpreting the meaning people assign to their social environment, considering their perceptions and experiences (Alharahsheh & Pius, 2020). It is suited for examining ethical dilemmas in AI research, where the goal is to comprehend the varying perspectives, values, and beliefs of the stakeholders involved in designing and developing AI technologies.

3.2 Research Approach

Qualitative research entails understanding human behaviour complexities and experiences, allowing the researcher to understand the research subject (Ritter et al., 2023). A qualitative method was suitable for this study and provided valuable insights from South African AI professionals on their perceptions of ethical AI design and development.

The qualitative research approach uses techniques such as interviews, focus groups, and observations to collect more descriptive data that provides detailed insights into the research topic (Queirós, Faria, & Almeida, 2017). The use of the participants' words in a qualitative approach reduces the risk of misinterpretations and incorrect conclusions. It allows the researcher to use direct citations from the data-gathering process (Trochim, 2002).

C2 General

Having established the qualitative approach grounded in interpretivism, this study adopts an exploratory case study design to allow for a rich, contextualised understanding of AI ethics. The next section outlines how this design was operationalised through data collection methods, sampling, and analysis strategies tailored to the South African AI landscape

3.3 Research Design

The research design for this study is exploratory and case study-based, aligning with the qualitative and deductive approaches under the interpretive paradigm. This design enabled an in-depth, contextualised exploration of ethical considerations in AI design and development in South Africa, providing rich insights into how AI professionals navigate ethical challenges. It is a case study of South Africa focusing on AI professionals.

Using case studies allowed the researcher to examine the interplay of ethical theories, such as consequentialist and virtue ethics, within real-world AI development projects, thereby contributing to developing grounded, context-specific knowledge in AI ethics. Case studies are beneficial for evaluating complex dilemmas within real-life contexts to gain a comprehensive knowledge of the interplay between elements driving the development of ethical AI (Yin, 2018).

With the design established, the following section explains how participants were selected and how data was collected.

3.4 Data Collection Methods

The study's data collection methods included semi-structured interviews structured around the key themes of consequentialist and virtue ethics. This ensured that the data collected was aligned with the study’s conceptual framework, allowing for a focused exploration of ethical considerations (Johansson et al., 2022). Diverse participants were selected to ensure a range of

C2 General

perspectives that reflect different aspects of AI development, from technical to ethical considerations. This diversity assisted in understanding how ethical theories are applied across different contexts and roles.

Through interviews, the researcher gained in-depth knowledge of ethical issues and the qualities of AI professionals. This approach enabled a thorough investigation of the use of ethical concepts in AI development and design. The consequences of AI technologies and the moral principles that guide ethical AI were the main discussion topics during the interviews. Interviews were conducted with various AI professionals in South Africa.

3.5 Population and Sample

3.5.1 Population

The research focused on AI professionals within South Africa. The population of interest was people who design, develop, and deploy AI-based solutions. It is noted that their job titles varied depending on their context and the organisations they were affiliated with. The study population included Data Scientists, Machine Learning Engineers, AI Researchers, AI Product Managers, AI developers, Robotic Process Automation Specialists, and AI and ethics Regulatory Experts. The population included skilled AI professionals in various industries across genders, age groups, and demographics.

3.5.2 Sample and Sampling Method

Qualitative research aims to explore complex human experiences and social contexts in depth rather than to generalise findings statistically. As such, smaller, purposefully selected samples are typical and appropriate (Creswell & Poth, 2018). The goal is to reach thematic saturation, the point at which no new themes or insights emerge from additional data (Guest, Bunce & Johnson, 2006).

C2 General

This study involved 12 AI professionals from various sectors in South Africa. By the 10th interview, themes such as regulatory uncertainty, ethical awareness, and organisational values began to recur. The last two interviews confirmed data saturation, as no substantially new perspectives were introduced. This indicates that the sample size was sufficient to fulfil the study's exploratory aims and to support meaningful thematic analysis.

While the findings were insightful and aligned with qualitative research norms, the study's scope was limited. The sample may not fully capture sectoral or regional diversity across South Africa. Future research should consider engaging a broader and more demographically varied group of AI professionals. This would enhance generalisability and allow for a more comprehensive understanding of sector-specific ethical practices.

Participants were primarily sourced from the researcher's professional and academic network, a common practice in qualitative research that prioritises access to information-rich cases. However, this introduced the risk of selection bias. To reduce its effect, efforts were made to include individuals from multiple industries (e.g., health, finance, education, public sector) and varied professional roles (e.g., developers, managers, data scientists).

Nevertheless, this limitation is acknowledged. Future studies should aim for more randomised or stratified sampling methods to enhance the representativeness and credibility of the findings.

The study adopted a purposive sampling method involving AI professionals across various industries in South Africa. Purposive sampling is a beneficial strategy that allows researchers to purposefully select participants based on specific characteristics related to the research objectives, enabling researchers to acquire in-depth and meaningful data for analysis and interpretation (Rai et al., 2004). The study aimed to interview 12 participants, and ultimately, 12 candidates were successfully interviewed as part of the research process. This alignment

C2 General

between the research objective and execution strengthened the validity of the findings.

The purposive sampling approach ensured that the sample consisted of people actively involved in AI development and qualified to comment on the ethical frameworks applied in the study. The selection of diverse participants captured a wide range of viewpoints. The sample included individuals from various industries with different experience levels and responsibilities within the AI development process. This made it easier to comprehend how ethical theories are used in various contexts within South Africa's AI industry.

3.6 The Research Instrument

The integration of virtue ethics and consequentialist approaches shaped the structure of this study's data collection instrument. The instrument's development involved ensuring that the questions were aligned with virtue ethics and consequentialist ethical theories. This alignment helped gather relevant data directly connected to the conceptual framework and research questions, making it easier to analyse the data through virtue ethics and consequentialist approaches.

Consequentialist ethics interview questions were designed to assess how AI professionals evaluate the outcomes of AI systems. The questions probed into AI's societal, environmental, and individual impacts, which are central concerns in consequentialist ethics (Anderson & Anderson, 2020). This approach ensured that the data collected revealed how participants consider the broader consequences of their AI work.

The interview questions included scenario-based questions that required participants to evaluate different outcomes and explain their decision-making processes. This allowed the researcher to understand how AI professionals apply consequentialist thinking. The method aligns with the consequentialist focus on the outcomes of actions (Hagendorff, 2020).

C2 General

Additionally, questions informed by virtue ethics to explore the personal virtues that guide AI professionals. This includes examining how values such as honesty, fairness, and responsibility influence their work (Hogenhout, 2021). Virtue ethics emphasises the individual's moral character, so questions addressed how these virtues are nurtured or challenged in their professional environment. These questions assisted in uncovering how virtues influence ethical decision-making in the context of AI design (Alibašić, 2023).

The interviews were organised around the key themes of consequentialist and virtue ethics. This ensured that the collected data aligned with the study's conceptual framework, enabling a focused exploration of ethical considerations (Johansson et al., 2022). A diverse group of participants was selected to provide a range of perspectives that reflect various aspects of AI development, from technical to ethical considerations. This diversity allowed for understanding how ethical theories are applied across different contexts and roles. The following steps were used as part of the data collection process.

- The ethics certificate was obtained from the university's Research Ethics Committee.
- An email introducing the researcher and explaining the purpose of the research was sent to the participants identified through the researcher's professional network.
- Participants provided consent to proceed with the interview before the interview.
- Invites were sent to participants, providing an option to select face-to-face or virtual interview sessions.
- Interview questions were shared with participants before the interview session.
- Obtained permission for the participants to be audio recorded for data collection.

C2 General

3.7 Data Analysis Strategies and Interpretation

In this study, thematic analysis was employed for data analysis and interpretation. The interview transcripts provided rich qualitative data that captured the respondents' views and perceptions. According to Braun and Clarke (2006), thematic analysis is valuable for assessing qualitative data since it is conceptually flexible and easily comprehensible. Beginners can quickly adopt the method.

The responses were coded and categorised based on the themes derived from consequentialist and virtue ethics. For example, data related to the impacts of AI systems were analysed through a consequentialist lens. At the same time, responses about personal values and organisational support for ethical behaviour were examined using virtue ethics (Hagendorff, 2020).

The analysis compared how AI professionals prioritise outcomes versus virtues in their ethical decision-making. This provided insights into the interplay between these ethical frameworks in real-world AI development (Johansson et al., 2022). Applying the conceptual framework was central to interpreting the study's findings. By grounding the research in established ethical theories, the study provided a robust analysis of how AI professionals navigate ethical challenges in their work.

Participant responses were analysed using deductive thematic analysis using the Virtue and Consequentialist ethics framework, following the guidelines outlined by Braun and Clarke (2006):

Step 1: The interviews were audio recorded and converted into written transcripts. The transcripts were thoroughly read to understand the participants' perspectives regarding AI ethics.

Step 2: The written transcripts were summarised and examined to systematically and meaningfully organise the data and code the significant portions of the content.

C2 General

Step 3: Similar codes were grouped or discarded if deemed irrelevant to the study.

Step 4: Themes and sub-themes were identified by grouping similar codes. These were then reviewed for relevance, resulting in some themes and sub-themes being further divided, merged, created, or eliminated.

Step 5: Using the thematic map to verify their relationships with the study, understanding the connections between themes.

Step 6: Compile the report detailing the findings by interpreting findings about the theoretical framework.

3.8 Possible Limitations and Challenges of the Study

- The study relied on participants' perceptions, ideas, and experiences on ethical AI.
- The purposive sampling method, which involves choosing participants based on specific criteria, might lead to a selection bias. This could result in an unrepresentative sample that may not fully capture the diversity of perspectives within the AI field.
- Analysing qualitative data through the virtue and consequentialist ethics lenses can be complex and subjective.
- Discussing ethical considerations in AI may lead to disclosing sensitive or confidential information.

3.9 Quality Assurance

3.9.1 Transferability

In qualitative research, transferability refers to the extent to which the research findings can be applied to another context involving different participants. Through intentional sampling, researchers can facilitate judgements and help

C2 General

users make decisions about transferability (Anney, 2014). This research was specific to the perceptions of South African AI professionals in the ethical design and development of AI solutions. The purposeful sampling method was used to target AI professionals across different industries in SA. The findings may not be directly transferable to other countries; however, the results may yield several engaging lessons and contributions to the literature.

3.9.2 Credibility

Credibility is the confidence level in the accuracy of the study findings in qualitative research. Credibility determines whether the research findings accurately reflect the original opinions of the participants and that the views of the participants have been interpreted correctly from the original data. Accurate and reliable data can be achieved by conducting interviews skillfully and sensitively (Anney, 2014). The researcher ensured that professional interview guidelines were applied to ensure the collection of accurate data. Also, the researcher recorded the interviews to ensure the analysis reflects what the participants said.

3.9.3 Dependability

Dependability in qualitative research refers to the consistency and stability of the research findings across time and under various conditions. It entails proving that the study's findings are trustworthy and reproducible by other researchers. The researcher can achieve dependability by keeping records of the research process and proving that they have enough data for others to validate their findings (Anney, 2014). In this study, creating an audit trail enables the researcher to keep records of all the interview transcripts and the research process.

3.9.4 Confirmability

Confirmability is the extent to which the conclusions are not based on the researcher's biases, preferences, or preconceptions but more on participant

C2 General

replies and facts. It is about ensuring the research findings are supported by the data and are not overly impacted by the researcher. The data gathered from participants was directly reflected in the research conclusions. This indicates that the research's evidence backs the conclusions and interpretations and is not shaped by the researcher's biases. The interview transcripts and detailed notes support this.

3.10 Ethical Considerations

- The ethical research guidelines established by the Wits Business School were followed in this study.
- Consent from the participants was acquired either verbally or in writing to record and document the interview.
- Participants were not required to provide sensitive information and could opt out anytime.
- The interview participants knew the study's objectives, goals, and scope.
- Participant names and personal information are confidential.

C2 General

CHAPTER 4.PRESENTATION OF RESULTS / FINDINGS

4.1 Introduction

This chapter presents the study's findings, which focused on the perceptions of AI professionals within South Africa regarding the ethical use of AI. The chapter provides the outcomes from the semi-structured interviews. It highlights the key themes and patterns identified, offering a comprehensive understanding of the experiences and contributions of South African AI professionals.

4.2 Demographics of Participants

The findings draw on insights from a diverse group of skilled AI professionals, encompassing various genders, age groups, and demographics. The population of interest includes individuals involved in designing, developing, and deploying AI-based solutions across various industries. These professionals, whose job titles vary depending on their organisational context, include Data Scientists, Machine Learning Engineers, AI Researchers, AI Product Managers, AI Developers, Robotic Process Automation Specialists, and AI and ethics Regulatory Experts.

The study examined the demographics of AI professionals, focusing on their work experience, industry affiliation, gender distribution, and professional roles. Most participants had between 15 and 20 years of work experience, highlighting a seasoned cohort of AI practitioners. Regarding industry representation, most participants were employed within the Information Technology (IT) and Telecommunications sectors, reflecting the strong presence of AI applications in these fields. Gender distribution within the sample was significantly skewed, with eleven male participants and only one female participant, emphasising the gender disparity in AI-related professions.

C2 General

The sample included a mix of corporate and entrepreneurial AI professionals. Ten participants were engaged in corporate AI roles, while two participants were identified as tech entrepreneurs.

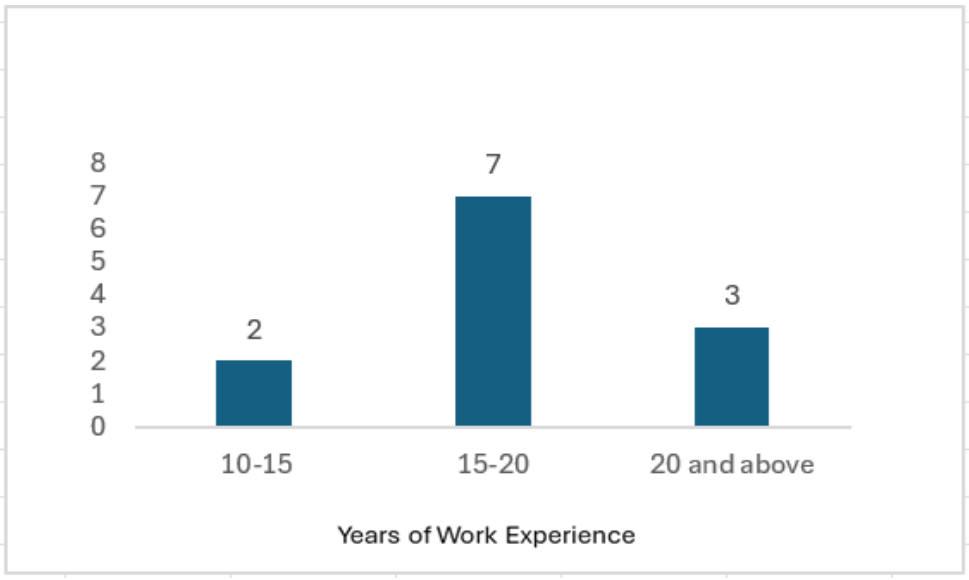


Figure 2: Participants Experience

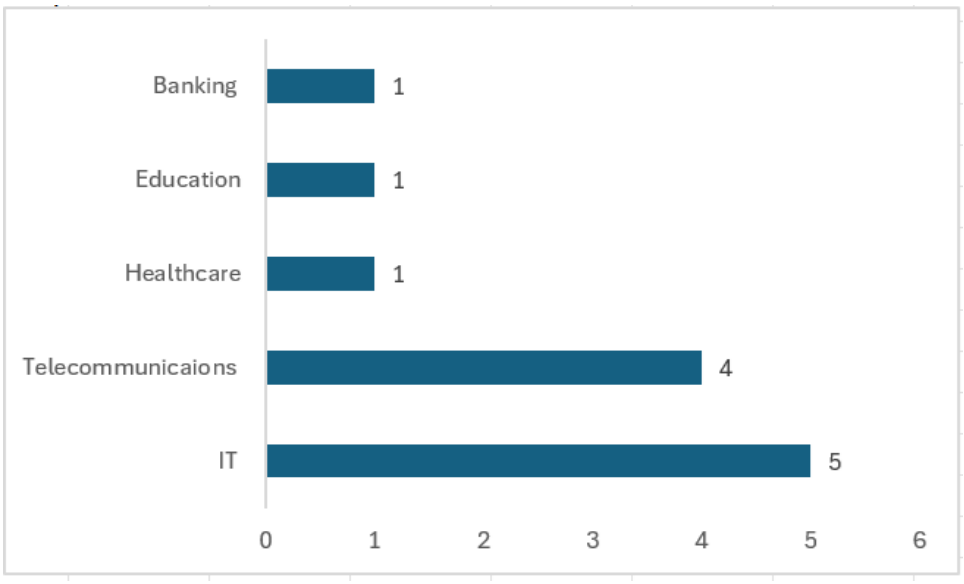


Figure 3: Participants Industry

C2 General

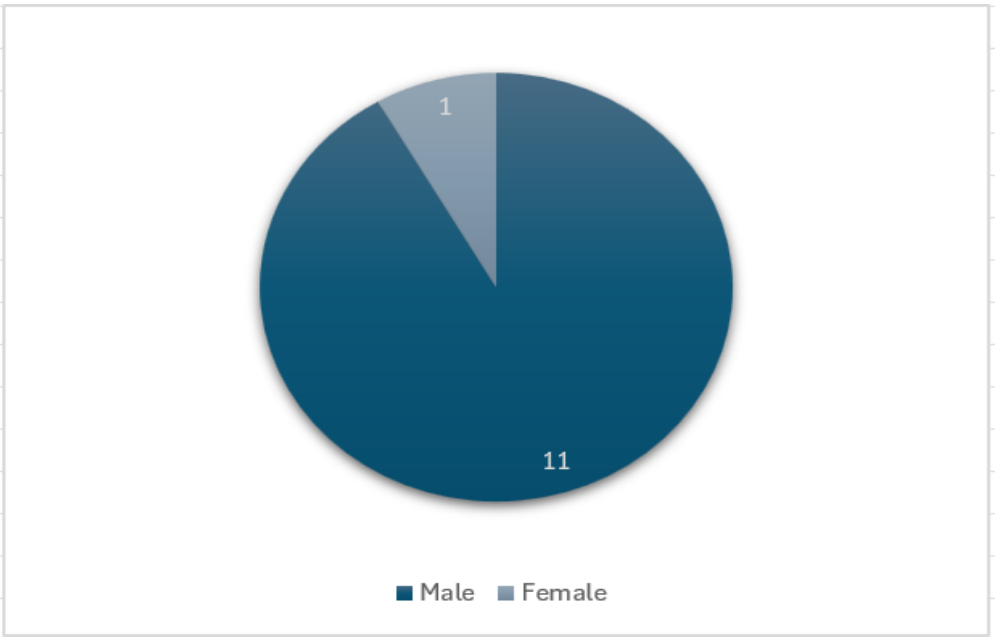


Figure 4: Participants Gender

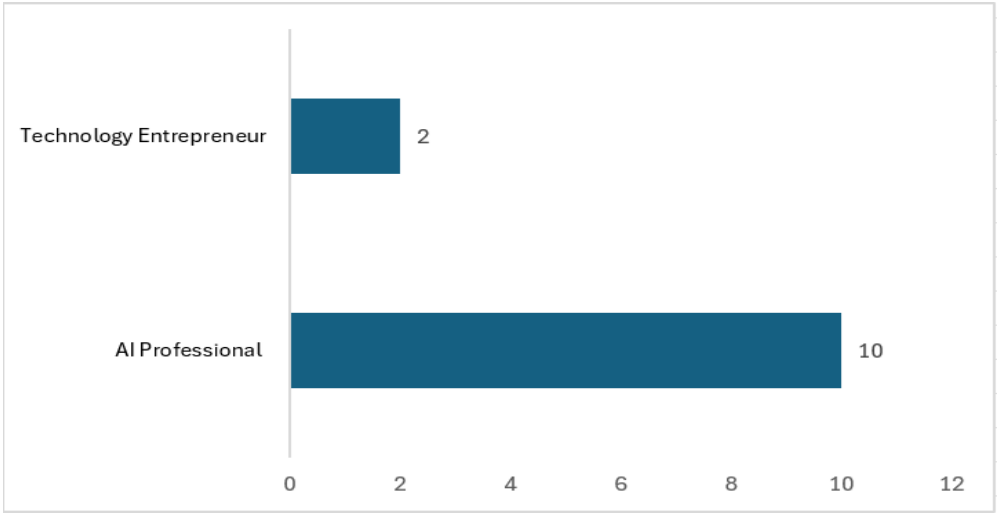


Figure 5: Employment Category

C2 General

4.3 Concept Analysis

The word cloud in Figure 6 visually represents key themes that emerged from the interview participants. Larger words indicate more frequent mentions, highlighting critical concepts such as AI, ethics, ethical, people, systems, solutions, and impact. These terms reflect the core discussions around ethical AI development, its implications, and the need for responsible decision-making.

The word cloud effectively captures the multifaceted discussions surrounding AI ethics. It highlights the need for responsible AI development, balancing technological innovation with human-centred values, regulatory oversight, and environmental sustainability. The participants' insights reinforce that ethical AI is an ongoing conversation that requires collaboration among developers, businesses, regulators, and society.



Figure 6: Concept Analysis

C2 General

4.4 Themes

This section presents a thematic analysis of participant responses, identifying key themes that emerge in discussions on ethical considerations in AI development. The findings highlight the complexity of ethical engagement, shaped by various factors such as professional roles, organisational policies, and individual perspectives. The research paper analyses the themes from participant interview responses exploring ethical considerations in AI development. Insights highlight varying levels of understanding and engagement with ethics in AI, influenced by professional, organisational, and personal contexts. The themes include:

- Awareness of AI ethics
- Perceptions of AI ethics
- Regulatory frameworks
- Personal beliefs and values
- Collaboration of stakeholders
- Challenges in integrating ethics
- Enablers for integrating ethics

4.5 Results Pertaining to Proposition 1

Proposition 1: The integration of the consequentialist and virtue ethics theories will lead to the ethical design and deployment of AI solutions that are contextually relevant and ethically sound.

The responses from AI professionals in South Africa indicate a mixed understanding and application of ethical considerations in AI development. While some participants acknowledged ethical concerns, most viewed AI as an extension of traditional software development rather than a distinct field requiring specialised ethical frameworks.

A key finding is that many AI professionals prioritise profitability, efficiency, and customer requirements over explicit ethical considerations. However, some

C2 General

participants reflect on their personal values as a guiding principle in ethical AI design.

4.5.1 Awareness and Perceptions of AI Ethics

Several participants do not differentiate AI ethics from general technology ethics. Ethics is considered a secondary concern, addressed only when issues arise rather than as an integral part of the design process. Some professionals dismiss ethicists and regulators as out of touch with technological advancements, arguing that regulations are behind in innovation.

A few professionals recognise that AI can introduce bias and unintended consequences, but solutions are often developed without formal ethical guidelines.

Code	Example Quotes
Participant 1	“Ethics don’t change because AI is just a technological tool.”
Participant 2	“There is no such thing as AI ethics. It’s about software ethics in general.”
Participant 5	“We try to hire ethical people and develop solutions that benefit customers, but there is no complex ethical evaluation.”
Participant 8	“Ethics is more about the overarching company values than something specifically tied to AI.”
Participant 10	“There’s no dedicated AI ethics manual; we rely on existing ethical guidelines.”
Participant 12	“Our ethical approach isn’t specific to AI; it’s about aligning with our general principles.”

4.5.2 Enablers for Integrating Ethics into AI Design

The strongest ethical check appears to be customer rejection of AI solutions that conflict with their values. Some participants apply their personal values (e.g., thinking about their children’s future) when making ethical decisions. A few

C2 General

participants believe that collaborating with regulators and ethicists could improve AI ethics, but they emphasise the need for more practical guidelines.

Code	Example Quotes
Participant 1	"If I build something unethical, my customers will reject it. That's the only measure that matters."
Participant 3	"We need strong ethical guidelines, collaboration with stakeholders, and organisational culture to support ethical AI."
Participant 6	"Ethical accountability is important, but we don't have dedicated people to review AI ethics."
Participant 8	"Having clear ethical guidelines and a strong organisational culture helps maintain ethical integrity in AI projects
Participant 11	"Supportive leadership is essential to drive ethical AI development forward."
Participant 10	"Working with external stakeholders, including ethicists and regulatory bodies, adds value to our ethical practices."

4.5.3 Barriers to Ethical AI Design

Many participants do not use a formal AI ethics framework. Some rely on general business ethics or common sense. South Africa lacks AI-specific regulations, making ethical integration inconsistent across industries. Some participants noted the lack of structured datasets on ethical considerations for diverse communities. AI systems prioritise business needs and efficiency over ethical concerns.

Code	Example Quotes
Participant 5	"Regulations are behind. By the time they catch up, AI has already moved forward."

C2 General

Participant 7	“I use company policies and common sense, but there is no AI ethics ‘bible’ guiding me.”
Participant 8	“One of the biggest challenges is the absence of clear ethical guidelines specifically for AI.”
Participant 10	“We struggle with access to resources that specifically address ethical considerations in AI.”
Participant 11	“Implementing ethical AI is difficult when the guidance is not clear or readily available.”
Participant 12	“Balancing innovation with ethical responsibilities is tough, especially with new technologies.”

While the consequentialist approach (focusing on outcomes such as customer satisfaction and profitability) is widely applied, virtue ethics (grounded in personal values) appears to be an informal guide rather than a structured approach in AI development.

To ensure ethical AI in South Africa, there is a need for formal AI ethics frameworks that are industry specific. Training for AI professionals on ethical considerations beyond just business priorities is required and stronger collaboration between AI developers, ethicists, and regulators to bridge the gap between technological advancement and ethical governance.

4.6 Results Pertaining to Proposition 2

Proposition 2: AI professionals play a crucial role in the ethical integration of AI into African societies, as their awareness and adherence to ethical principles significantly impact the responsible deployment of AI technologies.

4.6.1 AI Professionals’ Awareness and Perceptions of AI Ethics

The interviews revealed varied levels of awareness regarding AI ethics among professionals. Some participants acknowledged ethical considerations, while others saw no distinction between AI ethics and general software ethics. Most

C2 General

participants did not receive specialised training in AI ethics and relied on common sense or existing company policies.

Ethical considerations arise reactively rather than proactively in AI development. A few participants actively reflect on ethical considerations, citing personal values as an influence.

Code	Example Quotes
Participant 1	"We don't establish ethical standards as a separate process. It's just part of building technology".
Participant 3	"We develop ethical guidelines internally, but AI-specific ethics frameworks are missing."
Participant 5	"I follow a simple rule: do no harm. But AI ethics is complex, and we don't always have the tools to manage it properly".
Participant 9	"We try to ensure our solutions benefit users while avoiding harm."
Participant 10	"Our approach to AI ethics is guided by our understanding of broader business ethics."
Participant 12	"It's always a balance between innovation and maintaining ethical responsibility."

4.6.2 The Role of AI Professionals in Ethical AI Deployment

While AI professionals are the primary decision-makers in AI development, their impact on ethical deployment is influenced by multiple factors, including organisational priorities, customer expectations, and regulatory gaps. Most professionals do not consult ethicists or regulators, believing they lack technical understanding or are too slow to keep up with AI advancements.

Ethics is often driven by customer demands rather than internal ethical frameworks. Some AI professionals incorporate ethical principles into their work, especially in areas like bias mitigation and transparency.

C2 General

Code	Example Quotes
Participant 1	“We don’t sit around discussing ethics. We focus on building what the client wants”.
Participant 6	“We need structured AI ethics guidelines because right now, we rely too much on company policies that aren’t specific to AI”.
Participant 7	“I think about ethics in my work, but there is no universal framework guiding AI developers in South Africa”.
Participant 8	“Accountability lies with us, not with the AI.”
Participant 9	“Transparency with stakeholders is crucial in ethical AI deployment.”
Participant 12	“Ensuring fairness and unbiased results in AI systems is part of our role.”

4.6.3 Enablers for Ethical AI Integration

Several factors could enhance the role of AI professionals in ensuring ethical AI deployment in South Africa. The need for AI-specific ethics training was highlighted as a major gap. Regulatory clarity is required, but many professionals believe current efforts are too slow to be effective. Stakeholder collaboration (between developers, ethicists, and regulators) could improve ethical standards but is not widely practiced. Industry-wide AI ethics frameworks would provide clearer guidelines for AI professionals.

Code	Example Quotes
Participant 3	“Strong ethical guidelines and collaboration with regulators will help improve ethical AI development”.
Participant 5	“We need AI frameworks that are both practical and aligned with business needs.”
Participant 6	“We need ethical experts embedded in AI teams to guide us in real time”.

C2 General

Participant 8	“Clear policies and regular training make it easier to integrate ethics into AI.”
Participant 11	“Access to industry best practices supports ethical AI development.”
Participant 12	“A structured organisational approach makes ethical decision-making smoother.”

The findings suggest that AI professionals do influence ethical AI deployment, but their role is limited by a lack of formal ethical training, weak regulatory frameworks, and business-driven priorities. While some professionals prioritise ethical considerations, many rely on customer feedback or company policies rather than structured ethical frameworks. AI professionals do play a role in ethical AI deployment, but their awareness and adherence to ethical principles vary due to the lack of formal ethics frameworks, training, and regulatory clarity in South Africa.

4.7 Results Pertaining to Proposition 3

Proposition 3: Enablers such as ethical guidelines and regulatory frameworks will facilitate the successful incorporation of ethics into the design and development of AI solutions ensuring that AI solutions are both contextually relevant and ethically sound.

4.7.1 The Role of Ethical Guidelines in AI Development

The interviews revealed that while some AI professionals acknowledge the importance of ethical guidelines, their actual use in AI development is limited due to a lack of standardised frameworks. Participants noted that their organisations do not have specific AI ethics guidelines, instead they rely on general business ethics and software development best practices.

Ethical Considerations are often informal as many professionals incorporate ethics based on personal judgment or customer requirements rather than

C2 General

structured guidelines. Participants suggested that clearer ethical frameworks would improve AI development by providing consistent standards across the industry.

Code	Example Quotes
Participant 3	AI ethics should be built into guidelines that developers can follow. Right now, we rely on existing company policies, which aren't specific to AI".
Participant 6	"We don't have AI ethics specialists in our company. We rely on policies, but they are generic."
Participant 7	"There is no AI ethics 'manual' to guide us. We use company policies and common sense".
Participant 8	"Clear policies and regular training make it easier to integrate ethics into AI."
Participant 9	"Collaborating with experts keeps us aligned with ethical standards."
Participant 11	"Access to industry best practices supports ethical AI development."

4.7.2 The Impact of Regulatory Frameworks on AI Ethics

Regulatory frameworks could play a crucial role in shaping ethical AI development, but South Africa currently lacks clear AI regulations. Participants highlighted that AI-specific regulations are still in early development, leaving a governance gap in the industry. AI professionals expressed concerns that regulatory processes are too slow to keep up with rapid AI advancements.

AI professionals believe that regulators act reactively rather than proactively, making it difficult to align AI development with ethical and legal standards.

C2 General

Code	Example Quotes
Participant 1	“Regulations are not ahead of innovation; they follow it. AI ethics will only be regulated after problems emerge”.
Participant 2	“Where are the AI regulations? They don’t exist yet, so we follow what we think is right”.
Participant 5	“Regulators move too slowly. By the time they create regulations, AI has already evolved beyond them”.
Participant 8	“Regulatory frameworks are often outdated by the time they are implemented.”
Participant 10	“Clear regulations would provide much-needed guidance for ethical AI deployment.”
Participant 12	“The absence of formal regulations makes it challenging to standardise ethical practices.”

4.7.3 Enablers for Ethical AI Development

Several factors could support the integration of ethics into AI development, including collaboration, training, and access to ethical resources. Some participants suggested that stronger collaboration between AI developers, regulators, and ethicists could improve ethical AI development. Many professionals lack formal training in AI ethics and would benefit from workshops, guidelines, and organisational policies tailored to AI development.

Participants mentioned industry standards like IEEE as potential reference points, but these are not widely adopted in their workplaces.

Code	Example Quotes
Participant 3	“We need structured ethical training for AI professionals. Right now, ethics is not a key focus in AI education”.
Participant 5	“We need AI ethics frameworks that are both practical and aligned with business needs”.

C2 General

Participant 6	“Collaboration between AI developers and ethicists would improve ethical integration, but we don’t have this in place”.
Participant 8	“Ethical development is supported by clear guidelines and a strong ethical culture.”
Participant 10	“We need a proactive approach to integrating ethics into AI development.”
Participant 11	“Regular training and exposure to ethical discussions are crucial.”

The findings suggest that ethical guidelines and regulatory frameworks could facilitate ethical AI development, but their absence currently hinders consistent ethical integration. While some AI professionals incorporate ethics into their work, their approaches are informal and inconsistent due to a lack of clear frameworks. Ethical guidelines and regulations could be enablers, but their current absence or inefficiency in South Africa limits their effectiveness in ensuring contextually relevant and ethically sound AI solutions.

4.8 Consequentialist and Virtue Ethics Theory

The study used consequentialist ethics and virtue ethics as theoretical frameworks. The responses were grouped under predefined themes to explore how these ethical theories are applied in practice. The comparative table below demonstrates how consequentialist and virtue ethics perspectives shape the ethical considerations of AI professionals, supported by direct quotes from the interviews.

C2 General

Table 1: Consequentialist and Virtue Ethics Table

Aspect	Consequentialist Ethics	Virtue Ethics
Focus	**Outcomes and social impact**	**Individual values and moral character**
Awareness of AI Ethics	Awareness focuses on societal outcomes and optimisation. Quote: "*AI is just another technology that optimises solutions. The ethical considerations don't differ much from other technologies." Participant 2*	Awareness is guided by personal responsibility and customer needs. Quote: *"The resources are more my own exposure and knowledge of my customer." Participant 1*
Perceptions of AI Ethics	Ethics integrated into functional and societal benefits of AI. Quote: "*The benefits of AI far outweigh the disadvantages, and most use cases are straightforward." Participant 2*	Ethics guided by personal beliefs and the developer's moral responsibility. *Quote: "I think about my kids when building solutions. I want to leave behind a world where the products they use are ethical." Participant 1*
Regulatory Framework	Compliance ensures fairness and accountability. *Quote: "We follow frameworks aligned with the POPI Act and international guidelines to ensure fairness and accountability." Participant 4*	Relies on personal accountability in the absence of robust regulation. *Quote: "Regulators are always five years behind. I focus on building solutions that align with my customers' expectations." Participant 1*
Collaboration	Collective efforts with stakeholders ensure better societal outcomes. *Quote: "Stakeholders like legal, regulatory, and ethics teams provide a complete view of how the solution can impact users." Participant* 4	Prioritises personal judgment and customer input. *Quote: "I don't consult regulators or ethicists; they're out of tune. My customers' needs guide my ethical decisions." Participant 1*
Challenges in Integrating Ethics	Balancing innovation with ethical outcomes is difficult. Quote: *"Balancing innovation and business outcomes with ethical imperatives is difficult because they often conflict." Participant 4*	Limited by individual knowledge and values. Quote: *"There isn't a database for the ethics spectrum of society. We have to infer how users might receive our solutions." Participant 1*
Enablers in Integrating Ethics	Policies, monitoring, and interdisciplinary teams drive ethical outcomes.	Personal integrity and partnerships with experts guide ethical integration.

C2 General

	Quote: *"Strong AI policies, interdisciplinary teams, and regular monitoring are critical for integrating ethics." Participant 4*	Quote: *"Partnerships with universities and my moral compass help me build ethical solutions." Participant 4*
Personal Beliefs and Values	Personal values influence design decisions to ensure better societal outcomes. Quote: *"Your personal values influence how you design systems, but professional guidance is necessary to address blind spots." Participant 4*	Personal beliefs and moral responsibility guide ethical practices in AI development. Quote *"I think about my kids when building solutions. I want to leave behind a world where the products they use are ethical." Participant 1*

4.9 Summary of the Results/Findings

This chapter presents the findings on the perceptions of AI professionals in South Africa regarding the ethical use of AI. The research was based on semi-structured interviews and highlighted key themes, including awareness of AI ethics, regulatory challenges, and the role of personal values in ethical decision-making.

AI professionals apply ethical considerations inconsistently, often prioritising business goals over ethics. The consequentialist approach (profitability and efficiency) is dominant, while virtue ethics (personal values) serves as an informal guide rather than a structured framework.

AI professionals influence ethical AI deployment, but their awareness of and adherence to ethical principles vary. Ethics is often considered reactively rather than proactively, and there is limited formal training or engagement with ethicists and regulators.

Ethical guidelines and regulatory frameworks could enable ethical AI integration, but their absence or inefficiency in South Africa limits their effectiveness. AI professionals rely on company policies, common sense, or customer feedback rather than formal ethical guidelines.

C2 General

The chapter concludes by applying consequentialist and virtue ethics theories to AI professionals' perspectives. While consequentialist ethics emphasise societal outcomes and regulatory compliance, virtue ethics highlight personal values and moral responsibility. Both approaches face challenges due to the lack of structured ethical frameworks and slow regulatory advancements.

The study identified several key findings regarding AI professionals' ethical perceptions and practices in South Africa. Most AI professionals do not distinguish AI ethics from general technology ethics. Ethical considerations are often secondary, addressed only when issues arise. Some professionals acknowledge AI's potential for bias and unintended consequences but lack formal ethical guidelines.

The absence of structured AI ethics frameworks results in inconsistent ethical application across industries. South Africa lacks AI-specific regulations, fragmenting governance. AI professionals often rely on company policies and common sense instead of formal ethical training.

The study highlights the need for structured AI ethics frameworks, regulatory clarity, and interdisciplinary collaboration to ensure ethical AI development in South Africa. While AI professionals play a key role in ethical AI deployment, their approaches vary due to the lack of formal guidance and training. Addressing these gaps is crucial for balancing technological innovation with ethical responsibility.

C2 General

CHAPTER 5.DISCUSSION OF THE RESULTS OR FINDINGS

5.1 Introduction

This chapter discusses the research findings about existing literature on AI ethics. The analysis explored key themes identified in the study, including awareness of AI ethics, perceptions of AI ethics, the regulatory framework, personal beliefs, collaboration among stakeholders, challenges in integrating ethics, and enablers for ethical AI integration. By linking these findings to prior research, this chapter highlighted areas of alignment and emerging gaps in ethical AI implementation

Existing literature emphasises the critical role of AI ethics in mitigating risks such as bias and privacy violations. Scholars argue that ethical AI frameworks should be embedded in the development process, with clear regulatory guidance and stakeholder collaboration (Floridi et al., 2018). However, the findings of this study suggest that AI professionals often perceive ethics as secondary to business objectives, with some participants stating that ethical considerations do not differ from general technology ethics. This challenges the prevailing academic perspective that AI ethics requires distinct frameworks due to the technology's capacity for autonomous decision-making and large-scale impact.

Participants highlighted a lack of AI-specific regulations in South Africa, leading to reliance on internal policies and ad hoc ethical reasoning. This regulatory gap contrasts with global efforts in regions like the European Union, where initiatives like the EU AI Act seek to establish legal guidelines for responsible AI development. The absence of such frameworks locally implies a risk of inconsistent ethical practices across organisations, leaving ethical decision-making mainly in the hands of developers and business leaders.

C2 General

5.2 Discussion Pertaining to Proposition 1

Proposition 1: The integration of the consequentialist and virtue ethics theories will lead to the ethical design and deployment of AI solutions that are contextually relevant and ethically sound.

5.2.1 Personal Beliefs

The transcripts indicate that AI professionals in South Africa frequently rely on their personal values and experiences when making ethical decisions regarding AI development. Several participants noted that they consider the long-term societal consequences of their work, particularly in relation to their families and future generations. This aligns with virtue ethics, which emphasises the importance of individual moral character in ethical decision-making (Hogenhout, 2021). AI professionals who prioritise personal responsibility are more likely to integrate ethical considerations into their AI solutions, even in the absence of formal regulatory frameworks.

However, the study also found inconsistencies in ethical decision-making among AI professionals. While some rely on virtue-based moral reasoning, others prioritise business success and customer expectations over predefined ethical principles. This perspective aligns more with consequentialist ethics, which evaluate the morality of AI development based on its outcomes. AI professionals who adopt this approach view ethics as a function of market acceptance and business viability rather than moral obligations (Anderson & Anderson, 2020).

These findings from this study are like Pant et al. (2024), who found that AI technologies are as ethical as the people developing them, highlighting how personal beliefs shape ethical considerations. Additionally, Novelli et al. (2023) identified that inconsistent ethical practices among AI professionals often stem from varied interpretations of ethical responsibilities, a challenge also noted by participants in this study.

C2 General

Moreover, Floridi et al. (2018) argue from a consequentialist perspective that AI ethics should be evaluated based on societal benefits and risks, aligning with the study's findings that some AI professionals prioritise market-driven ethics. Pink et al. (2024) further supports this, emphasising that ethical compliance is often seen as a business constraint rather than a long-term advantage, particularly in environments with regulatory uncertainty, as seen in South Africa.

The reliance on personal beliefs can lead to inconsistent ethical practices, as individual interpretations of ethical responsibilities vary. This inconsistency might undermine ethical AI development, especially when formal guidelines are absent. The lack of formal AI ethics frameworks in South Africa places a greater burden on individual professionals to make ethical decisions. This fosters a reactive approach to ethics, where ethical issues are addressed only when they arise, contrasting with the ethics-by-design approach recommended by Sallstrom et al. (2019).

The study suggests that business pressures often lead AI professionals to prioritise profitability over ethical considerations, aligning with consequentialist ethics but challenging virtue ethics principles. This dynamic implies that ethical AI development might not be achieved consistently, potentially leading to ethical compromises in pursuit of market success.

The findings emphasise the need for clear regulatory frameworks and ethics training programs that provide consistent ethical guidelines for AI professionals. By addressing these challenges, South Africa can enhance the ethical integrity of AI systems, balance business objectives with moral responsibilities, and align practices with both virtue and consequentialist ethical principles.

C2 General

5.3 Discussion Pertaining to Proposition 2

Proposition 2: AI professionals play a crucial role in the ethical integration of AI into African societies, as their awareness and adherence to ethical principles significantly impact the responsible deployment of AI technologies.

5.3.1 Awareness of AI Ethics

AI professionals play a crucial role in the ethical integration of AI into African societies, as their awareness and adherence to ethical principles significantly impact the responsible deployment of AI technologies. The study revealed that AI professionals in South Africa exhibit varying levels of awareness regarding AI ethics. Some professionals demonstrated a strong understanding of ethical principles, particularly concerning data privacy, security, and human-centred values. However, many AI professionals lack deep engagement with ethical principles beyond basic compliance with regulations such as POPIA.

This suggests that while regulatory requirements drive awareness of data protection ethics, broader ethical considerations, such as accountability, fairness, and transparency, are less understood or inconsistently applied. Furthermore, several participants expressed uncertainty about who holds responsibility for AI ethics, reflecting a lack of structured ethical guidelines in AI development. This finding aligns with Pant et al. (2024), who found that AI professionals often lack the necessary knowledge and expertise to create ethically responsible AI solutions.

Additionally, Varanasi and Goyal (2023) argue that AI professionals struggle to implement ethical AI due to a lack of engagement with ethical principles beyond regulatory compliance. This aligns with the study's finding that AI ethics awareness is primarily driven by regulations like POPIA rather than structured ethical training. However, Lottu et al. (2019) advocate for continuous training throughout the AI lifecycle to ensure meaningful integration of ethics, which contrasts with the current practices in South Africa, where formal ethics education appears scarce.

C2 General

The transcripts showcase mixed levels of ethical awareness among AI professionals. These insights align with Stahl et al. (2021), who argue that regulatory uncertainty can hinder the meaningful integration of ethics into AI development, leading professionals to default to basic regulatory compliance rather than a broader ethical approach.

The absence of structured ethics education leads to varying interpretations of ethical responsibilities, potentially resulting in ethical inconsistencies in AI development. While regulations like POPIA promote data privacy ethics, the narrow focus on compliance means that broader ethical principles (e.g., fairness, accountability) are often overlooked, which might limit the social acceptability of AI technologies. Many AI professionals adopt a consequentialist approach, where ethical decisions are based on business outcomes rather than moral principles, which could compromise ethical standards in high-pressure market environments.

The study suggests that AI ethics awareness in South Africa is mostly incidental rather than systematically developed, which could undermine the ethical maturity of the AI industry. The lack of formal AI ethics frameworks contributes to confusion among AI professionals, highlighting the need for national guidelines that clearly define ethical responsibilities in AI development. To enhance ethical awareness, there is a need for structured education programs that not only focus on regulatory compliance but also broaden ethical considerations to include accountability, transparency, and fairness.

The study emphasises the critical role of AI professionals in ensuring ethical AI development but also reveals gaps in ethical awareness due to regulatory uncertainties, business pressures, and a lack of formal training. South Africa can strengthen the ethical foundation of its AI industry, promoting responsible innovation that aligns with both local values and global ethical standards by addressing these challenges through structured ethical education and the development of clear regulatory frameworks.

C2 General

5.3.2 Perceptions of AI Ethics

The study found that AI professionals hold diverse perceptions about AI ethics, influenced by business priorities, regulatory constraints, and personal values. Some participants view AI ethics as a secondary concern, believing that ethical considerations naturally align with market demand and regulatory requirements. This perspective aligns with consequentialist ethics, where AI ethics is judged based on outcomes rather than predefined moral principles (Anderson & Anderson, 2020). However, other professionals express concern about ethical gaps in AI development, particularly regarding bias, fairness, and transparency.

The study's findings align with Floridi et al. (2018), who argue that AI professionals often define ethics based on market-driven or regulatory compliance approaches rather than moral obligations. Additionally, Novelli et al. (2023) found that while some AI professionals take personal responsibility for AI ethics, others believe that ethical accountability should be shared between developers, organisations, and policymakers. This aligns with the study's finding that many AI professionals in South Africa feel uncertain about who holds ethical responsibility.

The study also aligns with Pink et al. (2024), who noted that insufficient training leads to inconsistent ethical practices, as many AI professionals base their ethical decisions on business needs and compliance requirements rather than proactive ethical engagement. Varanasi & Goyal (2023) highlighted similar challenges, arguing that AI professionals often struggle to understand and implement ethical principles, particularly when structured ethical training is lacking.

Furthermore, Sallstrom et al. (2019) advocate for ethics-by-design approaches, suggesting that ethics should be embedded in AI development from the outset rather than being an afterthought. However, the study suggests South African AI professionals do not consistently integrate ethics into AI development, primarily due to regulatory gaps and limited AI ethics training.

C2 General

The lack of a national AI ethics framework in South Africa contributes to inconsistent ethical practices. AI professionals often default to compliance-based ethics rather than proactive ethical engagement, similar to findings by Mashishi & Gungubele (2023). The study found that market-driven ethics often lead AI professionals to prioritise profitability over ethical considerations, a challenge also highlighted by Oludapo et al. (2024), who noted that business needs often override ethical integrity in technology development.

The study supports Lottu et al. (2019), who argue that continuous ethical training is crucial throughout the AI lifecycle, as many AI professionals in South Africa reported limited formalised AI ethics training, making ethical decision-making heavily influenced by individual values and business considerations/

The absence of structured AI ethics training and clear regulatory frameworks leads to varying ethical practices among AI professionals, potentially undermining the ethical integration of AI in African societies. The prevalence of consequentialist ethics means that ethical decisions are often reactive, based on market acceptance and business viability, which could compromise ethical standards in high-pressure environments. The lack of clarity about who holds ethical responsibility in AI development contributes to confusion among AI professionals, indicating a need for national AI ethics frameworks to define ethical roles and responsibilities clearly.

The study highlights a gap between ideal ethical practices (e.g., ethics-by-design) and the current practices in South Africa, where ethics is often seen as a business or regulatory function rather than a moral obligation. AI professionals require structured AI ethics training that goes beyond compliance, emphasising broader ethical principles like accountability, fairness, and transparency. Implementing an ethics-by-design approach would help AI professionals embed ethical considerations from the beginning of AI development, promoting responsible AI deployment in Africa.

C2 General

5.4 Discussion pertaining to Proposition 3

Proposition 3: Enablers such as ethical guidelines and regulatory frameworks will facilitate the successful incorporation of ethics into the design and development of AI solutions, ensuring that AI solutions are both contextually relevant and ethically sound.

5.4.1. Regulatory Framework

The study reveals that South Africa previously lacked a comprehensive AI regulatory framework, leaving AI professionals without clear guidelines on ethical AI development. While data protection laws such as POPIA guided privacy and security, they did not address broader AI ethics concerns such as fairness, accountability, and transparency.

However, in August 2024, the South African Department of Communications and Digital Technologies (DCDT) published the National Artificial Intelligence (AI) Policy Framework, marking a significant step forward in AI governance (Department of Communications and Digital Technologies, 2024). This policy framework lays the foundation for future AI laws, potentially leading to the development of a South African AI Act. It emphasises ethical AI principles, including transparency, accountability, and fairness, aiming to build trust among users, reduce biases, and promote social well-being.

The framework introduces guidelines that promote human-centred AI solutions, ensuring that AI technologies are designed to enhance human decision-making rather than replace it. This approach aligns with virtue ethics, emphasising professional responsibility and ensuring that AI advancements are consistent with societal values in South Africa. These findings align with Mashishi & Gungubele (2023), who previously noted that South Africa was developing a National AI Framework, which has now materialised into a formal policy. The new framework addresses gaps highlighted by Stahl et al. (2021), who argued that regulatory uncertainty can hinder ethical integration in AI development.

C2 General

The framework's focus on transparency and accountability is also consistent with Sallstrom et al. (2019), who emphasise that AI professionals should adopt ethics-by-design approaches, embedding ethical considerations into AI development. The guidelines outlined in the policy framework align with Sallstrom et al.'s (2019) recommendation for a proactive approach to AI ethics, moving South Africa closer to global best practices.

The study's findings also support Novelli et al. (2023), who found that clear regulatory frameworks help AI professionals balance ethical considerations with business pressures, fostering an environment where ethical compliance is seen as a strategic advantage rather than a constraint.

The absence of a national AI ethics framework previously contributed to inconsistent ethical practices. Still, the new policy framework aims to address this gap, providing AI professionals with clear ethical guidelines. The framework offers structured guidance, potentially reducing the influence of market demands on ethical decision-making, a challenge previously highlighted by Oludapo et al. (2024). The policy framework extends beyond data privacy, addressing fairness, accountability, and transparency, thus filling critical gaps left by POPIA.

The framework's detailed guidelines will likely lead to more consistent ethical practices, helping AI professionals align with standardised ethical principles. By embedding ethics into AI systems from the development stage, the framework supports proactive ethical integration, as Sallstrom et al. (2019) recommended. The emphasis on human-centred AI aligns with virtue ethics, promoting professional responsibility and ensuring that AI developments are beneficial to society

The framework's strategic approach to ethical AI development aims to build trust among users and advance social well-being, potentially enhancing public confidence in AI technologies. The new framework transforms South Africa's AI governance landscape, establishing clear ethical benchmarks and reducing reliance on international guidelines that may not fit the local context.

C2 General

The shift towards an ethics-by-design approach could lead to more robust ethical practices, as AI professionals are encouraged to consider ethical impacts throughout the AI lifecycle, not just as a compliance exercise. The framework's focus on enhancing human decision-making supports the integration of virtue ethics in AI practices, ensuring that AI technologies are aligned with societal values.

The introduction of the National AI Policy Framework marks a pivotal development in South Africa's AI governance offering a strategic pathway to future AI legislation, potentially leading to a South African AI Act. By addressing past regulatory gaps, the framework ensures that AI technologies are transparent, fair, and accountable, contributing to ethical innovation that aligns with local values and international standards.

5.4.2. Challenges in Integrating Ethics

The study identifies several challenges in integrating ethics into AI development, including:

- **Regulatory Uncertainty** – The absence of AI-specific guidelines leaves AI professionals unsure of ethical best practices.
- **Business Pressures** – Many AI professionals prioritise profitability and market demand over ethics, as ethical compliance is often perceived as slowing down innovation.
- **Lack of AI-specific training** – AI professionals reported limited formal training on AI ethics, leading to varied interpretations of ethical responsibilities.

These findings indicate that AI ethics is often deprioritised due to regulatory gaps, business constraints, and a lack of structured ethical training. Stahl et al. (2021) argue that without clear policies, AI professionals struggle to integrate ethics meaningfully into AI development. Similarly, Kiemde and Kora (2022) emphasise that the dominance of Western ethical frameworks in AI guidelines can create

C2 General

confusion for African AI professionals, who may find these frameworks culturally irrelevant.

The study's findings align with other research, such as Lottu et al. (2019), who also identified regulatory uncertainty, business pressures, and inadequate training as critical barriers to ethical AI development. Moreover, Pant et al. (2024) highlighted that AI professionals often lack comprehensive training in ethics, resulting in inconsistent ethical practices The challenges related to business pressures were also noted by Oludapo et al. (2024), who argued that the competitive digital economy often drives organisations to prioritise profitability over ethical considerations, mirroring sentiments expressed by participants in this study.

Unlike Sallstrom et al. (2019), who advocate for an "ethics-by-design" approach that embeds ethics throughout the AI lifecycle, the study found that South African AI professionals often adopt a reactive approach to ethics, addressing concerns only when they arise. This contrasts with the proactive integration of ethics recommended in the literature, highlighting a gap between ideal ethical practices and the current practices observed in the South African AI industry.

Without clear AI-specific guidelines, ethical practices in AI development remain inconsistent, potentially leading to ethical breaches, biases in AI systems, and a lack of accountability. This uncertainty can undermine public trust in AI technologies and may result in reputational and legal risks for businesses. When profitability and market demands take precedence over ethical considerations, it can exacerbate issues such as algorithmic bias, data privacy violations, and the creation of AI tools that may harm vulnerable communities.

The absence of structured training on AI ethics means that professionals may not fully understand or apply ethical principles in their work. This gap could lead to inconsistent ethical standards, making it difficult to implement a unified approach to responsible AI development across the industry. The reactive approach observed in South African AI professionals, where ethics is addressed only when

C2 General

issues arise, contrasts with the “ethics-by-design” philosophy. This reactive stance increases the risk of ethical oversights during the AI development process and may lead to costly and complex remediation efforts after deployment.

The findings highlight critical gaps in the integration of ethics into AI development, suggesting a need for regulatory reform, business incentives to promote ethical practices, and tailored ethical training programs. Addressing these challenges could lead to more responsible AI innovations, enhanced public trust, and the development of AI technologies that align with both global ethical standards and local cultural contexts.

5.4.3. Enablers for Integrating Ethics

The study identifies several enablers that could facilitate the integration of ethics into AI development:

- **Government and Industry Collaboration** – AI professionals suggest that stronger government-industry partnerships could help develop practical AI ethics guidelines.
- **AI-Specific Ethical Guidelines** – AI professionals believe establishing formal AI ethics policies would provide clarity and consistency in ethical decision-making.
- **Ethics Training and Awareness Programs** – Many professionals support the introduction of structured AI ethics training to ensure that ethics is embedded in AI development from the start.

These findings align with Lottu et al. (2019), who argue that ethical AI development requires an ethics framework applied throughout the AI lifecycle. Similarly, Ashok et al. (2022) emphasise the need for AI governance structures that integrate regulatory oversight with digital ethics.

The findings are similar to those of Brendel (2021), who found that effective integration of ethics into AI systems is facilitated by strong partnerships between government bodies and industry leaders, which help translate ethical principles

C2 General

into practical guidelines. Additionally, Varanasi & Goyal (2023) highlighted similar enablers, including the need for regulatory support, structured ethical training, and collaborative frameworks to strengthen ethical practices among AI professionals.

Unlike Sallstrom et al. (2019), who advocate for ethics-by-design, the study suggests that AI professionals in South Africa need external regulatory support to integrate ethics effectively. This indicates that ethics-by-design alone may not be sufficient in the local context. This finding reflects the unique challenges of South Africa's AI landscape, where regulatory gaps and business pressures influence ethical practices.

Government and industry collaboration can be achieved by establishing partnerships, regulatory bodies, and industry stakeholders to co-create practical ethical guidelines. Providing AI professionals with formal education on ethical principles can help embed ethics into AI systems from the development stage, promoting a proactive approach. The National AI Policy Framework offers a structured approach to AI ethics, potentially leading to a South African AI Act that institutionalises ethical practices.

The framework's detailed guidelines will likely lead to more consistent ethical practices, helping AI professionals align with standardised ethical principles. Through ethics training and awareness programs, AI professionals can adopt ethics-by-design approaches, ensuring that ethical considerations are integrated throughout the AI lifecycle.

The study highlights the importance of structured enablers, such as government support, industry partnerships, and formal ethics training, in integrating ethics into AI development. The collaborative approach proposed by the National AI Policy Framework could shift the focus from market-driven to principle-driven ethics, promoting ethical integrity in South Africa's AI. Establishing formal ethical guidelines can enhance professional responsibility, ensuring that AI technologies are transparent, accountable, and aligned with societal values.

C2 General

5.4.4. Collaboration of Stakeholders

The study highlights the importance of collaboration between government, industry, academia, and civil society to ensure ethical AI development. AI professionals emphasise that collaborative efforts are crucial for creating ethical guidelines, providing ethics training, and developing AI solutions that are contextually relevant and ethically sound.

The study also finds that AI ethics should be localised to align with South African cultural and societal values, as highlighted in the literature on Ubuntu and African AI ethics (Gwagwa et al., 2022). The principles of Ubuntu, which emphasise community, interconnectedness, and ethical responsibility, provide a valuable framework for shaping AI governance in Africa, promoting ethical practices aligned with local values.

These findings from this study are similar to those of Arthur (2019), who recommends the integration of ethical and sociocultural values in AI development, suggesting that AI systems should reflect the social realities and cultural contexts of African societies. Also, Kiemde & Kora (2022) and Gwagwa et al. (2022) highlighted similar challenges, noting that AI ethics frameworks often rely on Western values, which may not be suitable for African contexts, reinforcing the need for local stakeholder collaboration to develop culturally relevant ethical standards.

Unlike Sallstrom et al. (2019), who advocate for an ethics-by-design approach driven primarily by industry standards, the study suggests that South African AI professionals require multi-stakeholder collaboration to effectively integrate ethics, particularly in a regulatory environment that is still evolving.

C2 General

Regulatory bodies and industry leaders can develop practical ethical guidelines, ensuring that AI professionals have clear standards to follow. Research institutions can provide empirical insights into ethical AI practices, supporting the development of training programs and ethical frameworks. Engaging with society ensures that AI systems are aligned with community needs, promoting transparency, fairness, and accountability in AI technologies.

Ubuntu's principles offer a unique ethical lens, promoting community-focused AI development that enhances social well-being. Multi-stakeholder collaboration helps create a balanced approach to AI ethics, integrating regulatory oversight, industry innovation, and societal values, which is crucial for building trust in AI systems.

Through collaboration with academia and society, AI ethics frameworks can be localised to reflect African cultural and socio-ethical contexts, reducing ethical inconsistencies and biases. Stakeholder collaboration enhances the practical application of the ethics-by-design approach, helping AI professionals embed ethical principles throughout the AI development lifecycle. The engagement of government entities in collaborative initiatives ensures that AI regulations are clear, practical, and aligned with ethical expectations, mitigating business pressures that might undermine ethical practices.

The study underscores the importance of collaboration in bridging gaps between theoretical ethics and practical implementation, showing that AI professionals' benefit from input from diverse stakeholders, including regulators, researchers, and community representatives. The integration of Ubuntu principles into AI ethics frameworks offers a distinctive advantage for South Africa, promoting ethical AI development that is rooted in local values and enhances community trust.

Collaborative efforts help mitigate regulatory gaps, providing AI professionals with clear ethical guidelines and supporting training programs that raise awareness about ethical responsibilities. By engaging all stakeholders, South

C2 General

Africa can develop AI technologies that are not only innovative but also socially responsible, transparent, and fair, contributing to positive societal impacts.

The study emphasises that collaboration between stakeholders is a key enabler for integrating ethics into AI development, including:

- Government and industry partnerships that translate ethical principles into practical guidelines.
- Academic research contributions that support the development of ethics training programs.
- Society involvement to ensure AI solutions are inclusive, fair, and aligned with societal needs.
- Incorporating Ubuntu principles, promoting a community-centred approach to ethical AI development.

By fostering stakeholder collaboration, South Africa can address regulatory gaps, reduce business pressures, and enhance ethical awareness, ensuring that AI solutions are contextually relevant and ethically sound. The National AI Policy Framework provides a strong foundation for building these partnerships, promoting ethics-by-design approaches, and embedding ethical principles throughout the AI lifecycle.

5.5 Theory Driven Interpretation of findings

While the findings of this study reveal a general awareness of ethical considerations among AI professionals, a deeper analysis reveals that this awareness does not consistently translate into ethically guided actions. Thematic data show that while participants refer to values such as fairness, accountability, and integrity, aligning with virtue ethics, these are often subordinated to organisational pressures, resource constraints, or lack of regulatory enforcement. This indicates that virtue ethics, though conceptually resonant, lacks practical traction in environments where systemic support is weak.

C2 General

Similarly, consequentialist reasoning appeared in participants' concern about long-term harm, reputational damage, and public trust. However, ethical decision-making was often reactive, emerging after incidents rather than being proactively embedded in design phases. This reactive tendency undermines the forward-looking nature of consequentialism, highlighting the limits of applying such theories without structural accountability mechanisms.

The divergence between ethical intent and practical application suggests that ethical theory alone is insufficient for guiding AI development in the South African context. Instead, ethical practice must be constructed through institutional culture, enforceable guidelines, and education.

5.6 Conclusion

This chapter analysed the research findings on AI ethics in South Africa, linking them to existing literature. The study finds that AI professionals in South Africa often rely on personal beliefs and experiences when making ethical decisions, demonstrating a mix of virtue and consequentialist ethics. While some professionals prioritise societal impacts and moral character, others focus on business success and market acceptance.

The reliance on personal values creates inconsistent ethical practices influenced by a lack of formal AI ethics frameworks in South Africa. The study suggests that business pressures often push professionals toward profitability, challenging virtue ethics principles and risking ethical compromises. It concludes that clear regulatory frameworks and ethics training programs must provide consistent ethical guidelines to balance business objectives and moral responsibilities.

AI professionals significantly impact ethical AI deployment through their awareness of and adherence to ethical principles. The absence of structured ethics education contributes to inconsistent ethical practices, with many professionals adopting a consequentialist approach focused on business outcomes. The study emphasises the need for structured ethical education that

C2 General

broadens considerations beyond compliance, promoting accountability, transparency, and fairness.

The introduction of South Africa's National AI Policy Framework in August 2024 addresses previous regulatory gaps by providing guidelines that promote transparency, accountability, and fairness in AI development. This framework aligns with global best practices and supports a shift toward an ethics-by-design approach, embedding ethics into AI systems from development. However, challenges such as regulatory uncertainty, business pressures, and limited AI-specific training remain.

The study identifies key enablers for integrating ethics into AI development, including government-industry collaboration, formal ethical guidelines, and structured ethics training programs. The principles of Ubuntu and multi-stakeholder collaboration offer a culturally relevant framework for ethical AI, promoting community-centred development aligned with South African values.

C2 General

CHAPTER 6.CONCLUSIONS & RECOMMENDATIONS

6.1 Introduction

This chapter concludes the research on the perceptions of AI professionals in South Africa regarding the ethical implications of AI. It provides a comprehensive summary of how the research questions were addressed, highlights the key findings, discusses the study's contributions, outlines the study's limitations and recommendations, and suggests areas for future research.

6.2 Addressing the Research Questions

This study focused on the South African context and aimed to address the ethical considerations of AI by answering the following main research question:

- What are the ethical considerations for the design and development of AI technology in South Africa?

The following sub-questions support the above question:

III. What are the AI professionals' awareness and perceptions of AI ethics?
IV. What are the enablers for integrating ethics into the design and development of AI solutions in South Africa?

6.2.1 Conclusions regarding the main research question

The study found that the ethical considerations for the design and development of AI technology in South Africa encompass key principles such as privacy, bias, transparency, and accountability. These considerations are crucial for ensuring that AI technologies align with societal values, promote fairness, and respect human rights.

C2 General

Table 2: Summary of findings for the main research question

Principles	Key Findings
Privacy	The Protection of Personal Information Act (POPIA) is a critical regulatory framework in South Africa that governs data privacy and security. It ensures that organisations manage personal data ethically by setting standards for data collection, processing, and storage. AI systems, which often rely on large datasets, must comply with these regulations to avoid the misuse of personal information. The study highlighted that many AI professionals are aware of privacy concerns, particularly in scenarios where AI systems manage sensitive data. However, practical implementation of privacy-preserving technique was inconsistent across the industry.
Bias	AI technologies are prone to bias, especially when trained on datasets that reflect historical or societal inequalities. The study identified cases where biased AI systems led to unintended and discriminatory outcomes. One example discussed was the use of facial recognition technology that performed poorly with non-white individuals, resulting in the wrongful arrest due to low accuracy rates. This case underscores the need for bias mitigation strategies in AI development, including the use of diverse datasets, bias detection tools, and regular audits of AI systems.
Transparency	Transparency in AI involves making the decision-making processes of AI systems understandable and interpretable to users. The study found that many AI systems in South Africa operate as “black boxes,” meaning their internal logic and processing are not easily decipherable by users or even developers. This lack of transparency poses risks to accountability and trust. AI professionals expressed a need for tools and methods that enhance the explainability of AI models, such as explainable AI

C2 General

	techniques, which can help stakeholders understand how AI systems generate outputs.
Accountability	Accountability is about determining who is responsible when AI systems cause harm or produce unethical outcomes. The research indicated that there is currently a gap in clear accountability structures within organisations. Many AI professionals rely on personal ethics or informal guidelines rather than structured accountability frameworks. There is a need for organisational policies that define roles and responsibilities related to AI ethics, ensuring that accountability is not only a matter of individual conscience but also supported by institutional policies and governance frameworks

The study concluded that while South African AI professionals generally understand the importance of these ethical considerations, there is a significant gap in formal ethical frameworks and regulatory support to guide AI development. Addressing these gaps through stronger regulations, formalised ethical guidelines, and enhanced training could lead to more consistent and robust ethical practices in the AI industry.

6.2.2 Conclusions regarding research sub-question 1

Table 3: Summary of findings for research sub-question 1

Item	Key Findings
Differentiation Between AI and General Technology Ethics	A significant finding was that many AI professionals did not see a distinction between AI ethics and general technology ethics. For these professionals, ethical considerations were often treated as an extension of traditional software development ethics, without recognising the unique challenges posed by AI technologies, such as autonomous decision-making and algorithmic bias. This indicates a critical need for targeted training that addresses the specific ethical implications of AI systems.

C2 General

Reactive vs. Proactive Approaches	The study found that ethical considerations were often approached reactively. Many professionals only addressed ethical issues when they manifested, rather than embedding ethical principles throughout the AI development lifecycle. This reactive approach is problematic because it can lead to ethical oversights during critical phases such as data collection, model training, and deployment. Proactive integration of ethics into AI design, through frameworks like ethics-by-design, could mitigate such risks.
Influence of Personal Values	Some AI professionals relied heavily on their personal values, such as honesty, integrity, and fairness, to guide ethical decisions. While this individualised approach to ethics can contribute positively to ethical AI development, it also creates inconsistencies, as not all professionals may share the same ethical standards. The study suggested that formal ethical training and organisational support could help standardise ethical practices.
Perception of Regulations	AI professionals expressed criticism towards regulatory bodies and ethicists, often perceiving them as out of touch with rapid technological advancements. This criticism may lead to resistance against adopting formal ethical guidelines and regulations. Bridging this gap requires regulatory frameworks that are not only robust but also agile and adaptable to technological changes

While there is a foundational awareness of general ethics among AI professionals, there is a clear need for enhanced education and practical tools specific to AI ethics. Developing AI-specific ethical frameworks, regulatory policies, and training programs could improve the consistency and depth of ethical awareness in the AI industry.

C2 General

6.2.3 Conclusions regarding research sub-question 2

Table 4: Summary of findings for research sub-question 2

Item	Key Findings
Ethical Guidelines and Frameworks	The National Artificial Intelligence Policy Framework provides a strategic foundation for promoting transparency, accountability, and fairness in AI systems. These guidelines encourage organisations to adopt ethical-by-design approaches, where ethical considerations are integrated throughout the AI development process, from conceptualisation to deployment.
Regulatory Support	Regulatory frameworks like POPIA play a vital role in setting the legal and ethical boundaries for AI development. POPIA enforces strict data privacy and security standards, which directly contribute to ethical AI practices by ensuring that AI systems handle personal data responsibly.
Stakeholder Collaboration	Effective collaboration among AI professionals, policymakers, and ethicists was identified as a critical enabler. Such collaboration fosters a multidisciplinary approach to ethical AI, combining technical expertise with ethical and legal insights to create well-rounded solutions that meet both technological and ethical standards.
Education and Training	Providing AI professionals with formal training on ethical considerations beyond business priorities could bridge the gap between technological innovation and ethical governance. The study highlighted the need for workshops, seminars, and certification programs that focus on AI ethics, bias mitigation, and transparent AI development practices.
Personal Virtues and Organisational Culture	Many professionals integrate virtues such as honesty, empathy, and accountability into their work. Supporting this through formalised ethical frameworks and a strong organisational culture of ethics could enhance ethical practices.

While enablers for ethical AI integration exist, their potential impact is limited by inconsistent application and the need for comprehensive regulatory and

C2 General

educational support. The study recommends enhancing ethical frameworks, providing robust training, and promoting cross-sector collaboration to strengthen ethical AI development in South Africa.

6.3 Research Contributions

This research contributes to the extant literature on the ethical design and implementation of AI. The research highlighted the need for formalised ethical frameworks specific to AI in South Africa. It provided insights into the ethical challenges that AI professionals face, such as balancing business objectives with ethical obligations. The study also emphasised the importance of regulatory and educational support to enhance ethical awareness among AI professionals.

The study applied consequentialist and virtue ethics theories to evaluate ethical considerations in AI. The consequentialist approach emphasised evaluating the morality of actions based on their outcomes, promoting transparency and fairness in AI systems. The virtue ethics approach focuses on the personal virtues of AI professionals, such as honesty, empathy, and justice, demonstrating how these values influence ethical AI development.

The research employed a qualitative, interpretivist approach using semi-structured interviews. This methodological approach provided rich, contextual insights into the ethical perceptions of AI professionals and demonstrated the value of qualitative research methods in exploring complex ethical issues in technology development

6.4 Limitations of the Study

This interpretivist and qualitative study focused on South Africa and was limited to interviews with 12 AI professionals. Future studies may seek to expand the research by collecting data from a wider sample across South Africa. Additionally, it could explore the views of citizens and users of AI solutions and compile interviews or surveys with users.

C2 General

6.5 Recommendations

The research identified several key stakeholders critical to the ethical design and development of AI technology in South Africa. These stakeholders include AI professionals, executives and corporations, policymakers and regulators, educational institutions, and non-profit organisations. Each stakeholder has a unique role in promoting ethical AI practices. The recommendations are tailored to address their specific needs and responsibilities. Stakeholders can contribute to a sustainable AI ecosystem that balances technological innovation with ethical integrity and social responsibility by following the recommendations below.

Table 5: Table of Recommendations

Stakeholder	Recommendations
AI Professionals	AI Professionals including Data Scientists, Machine Learning Engineers, AI Researchers, AI Product Managers, AI Developers, Robotic Process Automation Specialists, and AI & Ethics Regulatory Experts, are at the forefront of AI development. Their decisions directly influence the ethical outcomes of AI technologies.
Ethicists and Legal Experts	AI professionals should collaborate with ethicists and legal experts during the AI development process. This multidisciplinary approach ensures that AI technologies are not only technically robust but also ethically sound.
Executives and Corporations	Executives and corporations play a critical role in setting organisational priorities and fostering a culture of ethical practices. Their influence extends over corporate strategies, investment decisions, and operational policies.
Policymakers and Regulators	Policymakers should maintain an open dialogue with AI professionals, corporations, and society to understand industry challenges and adjust regulations accordingly. Hosting conferences, workshops, and public consultations can support this engagement.

C2 General

Educational Institutions	Educational institutions, including universities, technical colleges, and research centres, are pivotal in building a knowledgeable workforce equipped with both technical skills and ethical awareness.
Industry and Government	Educational institutions should collaborate with industry leaders and government bodies to align academic programs with market needs and regulatory requirements. This includes internship programs, joint research projects, and knowledge-sharing platforms.
Non-profit Organisations	Non-profit organisations, advocacy groups, and community-based organisations, play a crucial role in protecting public interests and promoting ethical AI practices.
Collaborate with Other Stakeholders	Engage with policymakers, corporations, and educational institutions to provide insights from the community perspective, helping to shape policies and practices that are ethical and inclusive.

6.6 Suggestions for Further Research

The findings of this study on the ethical implications of AI in South Africa have highlighted several areas for future research. While the study addressed important questions about AI ethics, awareness, and integration enablers, it also opened new questions and opportunities for further exploration. These suggestions aim to build on the current research and guide future studies toward expanding knowledge in the field of ethical AI development.

6.6.1 Developing Ethical Frameworks for AI

The research revealed a gap in formalised ethical frameworks specific to AI development in South Africa. Further research could study how well the National Artificial Intelligence Policy Framework and POPIA are being implemented and understood by AI practitioners and organisations, identifying gaps and areas for improvement. Investigate the practical impact of ethical guidelines on AI project outcomes, particularly focusing on bias mitigation, transparency, and accountability.

C2 General

6.6.2 Exploring the Role of AI in Specific Ethical Domains

The study touched on key ethical principles such as privacy, bias, transparency, and accountability. Future research could delve deeper into these domains:

- **Data Privacy and AI:** Explore how AI technologies manage personal data under POPIA, including case studies on compliance and violations, and their impact on public trust.
- **Bias and Fairness in AI Systems:** Conduct experimental studies on AI algorithms to assess how training data and model design choices affect bias and fairness in AI outputs.

6.6.3 Impact of Training and Education on AI Ethics

The research identified education as a key enabler for ethical AI development. Further studies could:

- **Evaluate Training Programs**: Assess the effectiveness of existing training programs for AI professionals in AI ethics, identifying knowledge gaps and training needs.
- **Develop Educational Curriculum:** Collaborate with academic institutions to create a curriculum that integrates AI ethics with technical training, promoting a balanced approach to AI development.

6.6.4 The Influence of Organisational Culture on AI Ethics

The research indicated that corporate culture plays a significant role in ethical AI development:

- **Case Studies on Ethical Leadership:** Analyse organisations with strong ethical cultures to identify practices that support ethical AI development and share best practices.

C2 General

- **Impact of Corporate Governance:** Study how governance structures, such as ethics committees and board-level oversight, influence AI development practices.
- **Ethics and Profitability:** Investigate whether ethically developed AI systems contribute to business success, examining the relationship between ethical practices and corporate performance.

6.6.5 Future Technologies and Emerging Ethical Challenges

AI technologies are rapidly evolving, presenting new ethical challenges:

- **AI and Emerging Technologies:** Explore the ethical implications of AI integration with other technologies, such as blockchain, the Internet of Things (IoT), and biometric systems.
- **Ethical Considerations of Autonomous AI:** Research the unique ethical issues associated with autonomous AI systems, including self-driving vehicles, automated healthcare diagnostics, and predictive policing technologies.
- **Predictive Models and Decision-Making:** Study how predictive AI models are used in decision-making processes across industries, focusing on ethical implications related to accountability and human oversight.

6.7 Implementation Guidance and Evaluation

The following table provides a structured summary of each key recommendation, along with specific guidance for implementation and mechanisms to evaluate its effectiveness. This ensures the practical utility of the research findings and supports targeted action by various stakeholders.

C2 General

Table 6: Implementation Guidance and Evaluation Table

Recommendation	Implementation Guidance	Evaluation Mechanisms
Develop and Institutionalise AI Ethics Training	• Integrate ethics into tech education and certifications. • Offer industry-led training workshops. • Require ethics training for public-sector AI contractors.	• Pre- and post-training surveys. • Annual ethics audits. • Monitoring ethical incident rates.
Establish Contextualised Ethical Guidelines	• Create a multi-stakeholder working group. • Localise global frameworks to reflect South African values.	• Stakeholder reviews every two years. • Adoption rates across sectors. • Decision impact assessments.
Strengthen Regulatory and Governance Structures	• Finalise and implement the Draft AI Policy Framework. • Mandate algorithmic transparency in public projects. • Create a national AI Ethics Advisory Committee.	• Number of projects under ethical review. • Public satisfaction surveys. • Independent compliance reports.
Embed Ethics in Organisational Culture	• Appoint AI ethics champions. • Incorporate ethical KPIs in performance reviews. • Encourage internal ethical dialogue.	• Internal ethics climate surveys. • Innovations driven by ethics. • Ethical concern resolution rates

6.8 Conclusion

The study addressed existing gaps in ethical AI research within South Africa and opened new avenues for further exploration. These suggestions for future research aim to broaden understanding, influence policy development, and enhance ethical practices in the rapidly evolving field of artificial intelligence. By exploring these new research opportunities, academics, policymakers, and industry leaders can contribute to building a responsible AI ecosystem that prioritises ethical integrity, human rights, and societal well-being.

C2 General

REFERENCES

Alharahsheh, H. H., & Pius, A. (2020). A review of key paradigms: Positivism vs interpretivism. *Global Academic Journal of Humanities and Social Sciences*, 2(3), 39–43. https://doi.org/10.36348/gajhss.2020.v02i03.001

Alibašić, H. (2023). Developing an ethical framework for responsible artificial intelligence (AI) and machine learning (ML) applications in cryptocurrency trading: A consequentialism ethics analysis. *FinTech*, 2(3), 430–443. https://doi.org/10.3390/fintech2030024

Queirós, A., Faria, D., & Almeida, F. (2017). Strengths and limitations of qualitative and quantitative research methods. *European Journal of Education Studies*, 3(9), 369–387. https://doi.org/10.5281/zenodo.887089

Anderson, S. L., & Anderson, M. (2021). AI and ethics. AI and Ethics, 1(1), 27–31. https://doi.org/10.1007/s43681-020-00003-6

Anney, V. N. (2014). Ensuring the quality of the findings of qualitative research: Looking at trustworthiness criteria. *Journal of Emerging Trends in Educational Research and Policy Studies*, 5(2), 272–281. https://www.scholarlinkinstitute.org/jeteraps/articles/Ensuring%20The%20Quality%20Of%20The%20Findings%20new.pdf

Arakpogun, E. O., Elsahn, Z., Olan, F., & Elsahn, F. (2021). Artificial intelligence in Africa: Challenges and opportunities. In A. Hamdan, A. E. Hassanien, A. Razzaque, & B. Alareeni (Eds.), *The Fourth Industrial Revolution: Implementation of artificial intelligence for growing business success* (pp. 375–388). Springer. https://doi.org/10.1007/978-3-030-62796-6_22

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for artificial intelligence and digital technologies. *International Journal of Information Management*, 62. Elsevier Ltd. https://doi.org/10.1016/j.ijinfomgt.2021.102433

Ayling, J., & Chapman, A. (2022). Putting AI ethics to work: Are the tools fit for purpose? *AI and Ethics*, 2(3), 405–429. https://doi.org/10.1007/s43681-021-00084-x

Bednar, K., & Spiekermann, S. (2022). Eliciting values for technology design with moral philosophy: An empirical exploration of effects and shortcomings. *Science, Technology & Human Values*, 016224392211225. https://doi.org/10.1177/01622439221122595

Bench-Capon, T. J. M. (2020). Ethical approaches and autonomous systems. *Artificial Intelligence*, 281. https://doi.org/10.1016/j.artint.2020.103239

Bleher, H., & Braun, M. (2023). Reflections on putting AI ethics into practice: How three AI ethics approaches conceptualize theory and practice. *Science and Engineering Ethics*, 29(3). https://doi.org/10.1007/s11948-023-00443-3

Borenstein, J., & Howard, A. (2021). Emerging challenges in AI and the need for AI ethics education. *AI and Ethics,* 1(1), 61–65. https://doi.org/10.1007/s43681-020-00002-7

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

Brendel, A. B., Mirbabaie, M., Lembcke, T. B., & Hofeditz, L. (2021). Ethical management of artificial intelligence. *Sustainability*, 13(4), 1974. https://doi.org/10.3390/su13041974

Creswell, J. W., & Poth, C. N. (2018). Qualitative inquiry and research design: Choosing among five approaches (4th ed.). Sage Publications.

Department of Communications and Digital Technologies. (2024). *South Africa National AI Policy Framework.* Government of South Africa. https://www.dcdt.gov.za/sa-national-ai-policy-framework/file/338-sa-national-ai-policy-framework.html

Dove, E.S., & Chen, J. (2021). What does it mean for a data subject to make their personal data *“manifestly public”?* An analysis of GDPR Article 9(2)(e).

International Data Privacy Law, 11(2), 107–124. https://doi.org/10.1093/idpl/ipab005

Eitel-Porter, R. (2021). Beyond the promise: Implementing ethical AI. *AI and Ethics*, 1(1), 73–80. https://doi.org/10.1007/s43681-020-00011-6

Eke, D. O., Wakunuma, K., & Akintoye, S. (Eds.). (2023). *Responsible AI in Africa: Challenges and opportunities.* Palgrave Macmillan. https://link.springer.com/book/10.1007/978-3-031-08215-3

Figueras, C., Verhagen, H., & Cerratto-Pargman, T. (2022). Exploring tensions in responsible AI in practice: An interview study on AI practices in and for Swedish public organizations. *Scandinavian Journal of Information Systems*, 34(2), Article 6. https://aisel.aisnet.org/sjis/vol34/iss2/6/

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Fukuda-Parr, S., & Gibbons, E. (2021). Emerging consensus on 'ethical AI': Human rights critique of stakeholder guidelines. *Global Policy*, 12(S6), 32–44. https://doi.org/10.1111/1758-5899.12965

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. Field Methods, 18(1), 59–82. https://doi.org/10.1177/1525822X05279903

Gwagwa, A., Kazim, E., & Hilliard, A. (2022). The role of the African value of Ubuntu in global AI inclusion discourse: A normative ethics perspective. *Patterns*, 3(4). Cell Press. https://doi.org/10.1016/j.patter.2022.100462

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hamilton, C., Swart, W., & Stokes, G. M. (2021). Developing a measure of social, ethical, and legal content for intelligent cognitive assistants. *Journal of Strategic Innovation and Sustainability*, 16(3), 1–16. https://doi.org/10.33423/jsis.v16i3

Hogenhout, L. (2021). *A framework for ethical AI at the United Nations*. https://doi.org/10.48550/arXiv.2104.12547

Huriye, A. Z. (2023). The ethics of artificial intelligence: Examining the ethical considerations surrounding the development and use of AI. *American Journal of Technology,* 2(1), 37–45. https://doi.org/10.58425/ajt.v2i1.142

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. Sustainability *(Switzerland),* 15(16). https://doi.org/10.3390/su151612451

Kazim, E., & Koshiyama, A. S. (2021). A high-level overview of AI ethics. *Patterns*, 2(9). Cell Press. https://doi.org/10.1016/j.patter.2021.100314

Kiemde, S. M. A., & Kora, A. D. (2022). Towards an ethics of AI in Africa: Rule of education. *AI and Ethics*, 2(1), 35–40. https://doi.org/10.1007/s43681-021-00106-8

Kieslich, K., Keller, B., & Starke, C. (2022). Artificial intelligence ethics by design: Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data and Society*, 9(1). https://doi.org/10.1177/20539517221092956

Kraus, S., Jones, P., Kailer, N., Weinmann, A., Chaparro-Banegas, N., & Roig-Tierno, N. (2021). Digital transformation: An overview of the current state of the art of research. *SAGE Open*, 11(3). https://doi.org/10.1177/21582440211047576

Longo, F., Padovano, A., & Umbrello, S. (2020). Value-oriented and ethical technology engineering in industry 5.0: A human-centric perspective for the design of the factory of the future. Applied Sciences *(Switzerland)*, 10(12), 1–25. https://doi.org/10.3390/APP10124182

Madiega, T. (2021). Artificial intelligence act. *European Parliamentary Research Service*.https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf

Marwala, T., & Edu, W. U. (2024). *United Nations University Framework for the Governance of Artificial Intelligence.* United Nations University. https://www.unu.edu

Mashishi, A., & Gungubele, M. (2023). *South Africa's Artificial Intelligence (AI) Planning: Adoption of AI by Government.* Department of Communications and Digital Technologies. https://www.dcdt.gov.za/images/phocadownload/AI_Government_Summit/National_AI_Government_Summit_Discussion_Document.pdf

McLennan, S., Fiske, A., Tigard, D., Müller, R., Haddadin, S., & Buyx, A. (2022). Embedded ethics: A proposal for integrating ethics into the development of medical AI. *BMC Medical Ethics*, 23(1), Article 43. https://doi.org/10.1186/s12910-022-00746-3

Morley, J., Kinsey, L., Elhalal, A., Garcia, F., Ziosi, M., & Floridi, L. (2023). Operationalising AI ethics: Barriers, enablers, and next steps. *AI and Society*, 38(1), 411–423. https://doi.org/10.1007/s00146-021-01308-8

Munn, L. (2023). The uselessness of AI ethics. *AI and Ethics*, 3(3), 869–877. https://doi.org/10.1007/s43681-022-00209-w

Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies,* 28(4), 4221–4241. https://doi.org/10.1007/s10639-022-11316-w

Nassar, A., & Kamal, M. (2021). Ethical dilemmas in AI-powered decision-making: A deep dive into big data-driven ethical considerations. *International Journal of Responsible Artificial Intelligence*, 11(8), 1–11.https://neuralslate.com/index.php/Journal-of-Responsible-AI/article/view/43

Novelli, C., Taddeo, M., & Floridi, L. (2023). Accountability in artificial intelligence: What it is and how it works. *AI and Society*. https://doi.org/10.1007/s00146-023-01635-y

Oludapo, S., Carroll, N., & Helfert, M. (2024). Why do so many digital transformations fail? A bibliometric analysis and future research agenda.

Journal of Business Research, 174, Article 114528. https://doi.org/10.1016/j.jbusres.2024.114528

Pant, A., Hoda, R., Spiegler, S. V., Tantithamthavorn, C., & Turhan, B. (2024). Ethics in the age of AI: An analysis of AI practitioners' awareness and challenges. *ACM Transactions on Software Engineering and Methodology,* 33(3). https://doi.org/10.1145/3635715

Pink, S., Quilty, E., Grundy, J., & Hoda, R. (2024). Trust, artificial intelligence and software practitioners: An interdisciplinary agenda. *AI and Society.* https://doi.org/10.1007/s00146-024-01882-7

Rai, N., Thapa, B., Joshi, A. B., Banjara, M. R., Myneni, S., & Kothari, C. (2004*). A study on purposive sampling method in research. Kathmandu School of Law Style Guide*. https://www.academia.edu/28087388/A_STUDY_ON_PURPOSIVE_SAMPLING_METHOD_IN_RESEARCH

Rai, N., Thapa, B., Joshi, A. B., Banjara, M. R., Myneni, S., & Kothari, C. (2023). A study on purposive sampling method in research. *Kathmandu School of Law Journal*, 21(3), 45–58. https://www.academia.edu/28087388/A_STUDY_ON_PURPOSIVE_SAMPLING_METHOD_IN_RESEARCH

Ritter, C., Koralesky, K. E., Saraceni, J., Roche, S., Vaarst, M., & Kelton, D. (2023). Invited review: Qualitative research in dairy science—A narrative review. *Journal of Dairy Science*, 106(9), 5880–5895. https://doi.org/10.3168/jds.2022-23125

Shaver, R., Baron, M. W., Pettit, P., & Slote, M. (2000). Three Methods of Ethics: A Debate. *The Philosophical Review*, 109(1), 125. https://doi.org/10.2307/2693569

Stahl, B. C., Antoniou, J., Ryan, M., Macnish, K., & Jiya, T. (2022). Organisational responses to the ethical issues of artificial intelligence. *AI and Society*, 37(1), 23–37. https://doi.org/10.1007/s00146-021-01148-6

Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society*, 40(2), 137–157. https://doi.org/10.1080/14494035.2021.1928377

Trochim, W. M. K. (2002). The research methods knowledge base. William M. K. Trochim. https://www.researchgate.net/publication/243783609

van Norren, D. E. (2022). The ethics of artificial intelligence, UNESCO and the African Ubuntu perspective. *Journal of Information, Communication and Ethics in Society,* 21(1), 112–128. https://doi.org/10.1108/JICES-04-2022-0037

Varanasi, R. A., & Goyal, N. (2023, April 19). It is currently hodgepodge: Examining AI/ML practitioners' challenges during co-production of responsible AI values. *Proceedings of the Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3544548.3580903

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1). https://doi.org/10.1186/s41239-019-0171-0

APPENDIX (A) Semi-structured interview guide

1. How do you examine the potential societal, environmental, and individual impact of the AI systems you develop?
2. What measures do you take to ensure that the outcomes of your AI projects align with ethical standards?
3. How do you evaluate the potential impact of AI systems you develop?
4. What resources are available to you to ensure that AI solutions are ethically designed?
5. Can collaboration with stakeholders such as ethicists and regulators influence the ethical design of AI technologies? How so?
6. How do you reflect on your own ethical practices and decision making as an AI professional?
7. In what ways do your personal beliefs and values influence the design and development of AI systems?
8. What challenges do you face in integrating ethics into the design and development of AI solutions?
9. What factors would enable or support you in integrating ethics into the design and development of AI solutions?

APPENDIX (B) Consent Form

Participant Consent Form for Research Study

Perceptions of artificial intelligence (AI) professionals in South Africa on the ethical implications of AI

Name of researcher: Sinah Siaga

I, ……., agree to participate in this research project.

I agree to the following:

(Please circle the relevant options below)

The research study was explained to me. I understand what this study is about.	YES/NO
I understand that I can volunteer to take part in the study	YES/NO
I agree that the interview will be audio-recorded	YES /NO
I agree that direct quotations from my interview may be used by the researcher in their research report	YES/NO
I agree that my participation will remain anonymous (my name or other identifying data will not be used by the researcher in their research report)	YES/NO

Name of participant……
……
Participant Signature ……
Date……

Name of researcher/person seeking consent…….
Signature……
Date…….

APPENDIX (C) Participation Information Sheet

Participant Information Sheet

Dear Sir/Madam

My name is Sinah Siaga. 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 conducting a research study on the perceptions of artificial intelligence (AI) professionals in South Africa on the ethical implications of AI. The purpose of this study is to investigate the perceptions of AI professionals on the challenges related to ethical AI design and development across various sectors in South Africa.

I would like to invite you to take part in an interview. If you decide to participate, your involvement will last about one hour. The interview will take place online based on your availability. With your permission, I would like to audio record the interview. This data will be securely stored for a period specified by the university. Only the researcher will have access to the data.

Your participation will be confidential and anonymous. When sharing the results of the study, your name and any identifying information will not be included.

Participation in this study is entirely voluntary. You may choose not to take part, and you can withdraw from the study at any time without any consequences. You are also free to skip any questions you do not wish to answer. There will be no direct benefits to you for participating in this study. The risks involved in this study are minimal and are no more than what you would encounter in everyday life. However, some questions may cause discomfort. If you feel upset at any point, we can pause or reschedule the interview.

The findings of this research will be compiled into a report and made available on the university library website. If you would like to receive a summary of the report, I will be happy to send it to you.

If you have any questions about this research study, please feel free to contact me or my supervisor using the details below. For any concerns or complaints about the ethical aspects of this research, you are welcome to contact the University Human Research Ethics Committee (Non-Medical) at +27(0) 11 717 1408 or email hrecnon-medical@wits.ac.za.

Yours sincerely,

Sinah Siaga
Researcher
2395519@students.wits.ac.za
082 922 1194

Prof. Tendani Mawela
Supervisor
tendani.mawela@wits.ac.za
012 420 6484

APPENDIX (D) Ethics Approval Letter

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/DB2395519/432

This certificate is only valid with a legitimate ethics protocol number and signed by the Researcher (below).

This certificate is only valid if accompanied by formal permission from the relevant stakeholder(s).

Project title	Perceptions of artificial intelligence (AI) professionals in South Africa on the ethical implications of AI
Investigator / Researcher	Mrs Sinah Siaga
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

01/11/2024
Date:

APPENDIX (E) Turnitin Report



Turnitin Report 28 Feb 2025.pdf