University of Pretoria

FACULTY OF HUMANITIES

DEPARTMENT OF PHILOSOPHY

PHD RESEARCH

The Role of the Board in Governance of Artificial Intelligence Ethics – A Case for JSE Listed Companies

by

XITSHEMBHISO RUSSEL MULAMULA

Student Number: 17306699

Supervisor:

PROF. EMMA RUTTKAMP-BLOEM

ACKNOWLEDGMENTS

To GOD Almighty, thank you for giving me the strength and wisdom to complete this doctoral thesis.

To my dear wife Xingwavila Khosa-Mulamula, I could not have completed this thesis without your prayers, support, sacrifices, and understanding that I had to take some of the family time from you and our children. I pray that God continues to bless your loving, caring, and supportive heart.

My mentor in scholarship, Prof. Emma Ruttkamp-Bloem, you have been my beacon through this academic journey. Your leadership, patience, and insightful guidance has not only shaped this thesis but has also sculpted my philosophical enquiry. I am eternally grateful for your unwavering support.

To each participant who contributed to this research, your voices have been the cornerstone of this study. Your willingness to share your expertise and experiences has granted this work value and substance.

The Limpopo Treasury Department, your financial backing, and recognition of the pursuit of knowledge beyond career aspirations have enabled me to undertake this scholarly path.

To my siblings, Godfrey, Iren, Arrel and John, thank you for believing in me. In heaven, our mother, Gertrude, watches over us with a smile, knowing her legacy of unity continues to thrive. This PhD is as much yours as it is mine.

Last by not least, to my friends who have been my grounding forces, Bra Dan Neluheni, Martin Ramaboea, Magasela Mzobe, Nhlamulo Mabasa, your unwavering support has often served as a reminder of the greater purpose beyond the pages of my thesis.

To our beautiful children Xivono Kayla Mulamula and Risima Xitshembhiso Mulamula, I dedicated this doctoral thesis to you. May you grow to follow your parents' academic footsteps. I pray that God may continue to protect and bless you abundantly.
INKOMU!

DECLARATION

I, XITSHEMBHISO RUSSEL MULAMULA declare that this thesis/dissertation/mini-dissertation is my own original work. Where secondary material is used, this has been carefully acknowledged and referenced in accordance with university requirements.

I understand what plagiarism is and am aware of university policy and implication in this regards.

15 February 2024

ETHICS STATEMENT

I, XITSHEMBHISO RUSSEL MULAMULA, have obtained, for the research described in this work, the applicable research ethics approval.

Date of approval: 26 AUGUST 2021

Research ethics number: 17306699 (HUM021/0621)

I, XITSHEMBHISO RUSSEL MULAMULA, declare that I have observed the ethical standards required in terms of the University of Pretoria's Code of Ethics for Researchers and Policy Guidelines for Responsible Research.

15 February 2024

ABSTRACT

This interdisciplinary study enriches academic discourse by merging corporate governance and AI ethics insights, offering a comprehensive perspective on the ethical issues surrounding AI development, deployment, and implementation. This study aims to investigate how corporate boards can lead and set a tone for ethically managing AI technologies, ensuring that these technologies align with ethically sound strategic objectives. It contributes uniquely to scholarly dialogue by examining the board's role in AI ethics and governance through the lens of both AI ethics literature, corporate governance principles and AI technology's impact on society.

To date, the existing literature on AI ethics has largely centred on software developers' tasks or the broader regulatory aspects of AI technologies, paying minimal attention to the board's influence or their oversight responsibilities within private sector companies. Boards play a pivotal role in fostering an ethical AI culture, managing AI opportunities and the potential societal risks AI poses, including ethical dilemmas, biases, and reputational threats.

A qualitative phenomenological method was employed to understand and interpret board members' experiences, with the aim of uncovering the meanings within their lived experiences. The findings of the empirical study highlight the board's crucial role in establishing ethical guidelines for AI governance and were corroborated by the literature review.

For responsible AI ethics governance, it is imperative for boards to establish robust structures, processes, and rational mechanisms for responsible AI ethics governance to mitigate the significant risks associated with AI technologies, such as bias, discrimination, and privacy violations. The study adopted a stakeholder-inclusive approach, which guided the recommendations, which includes setting up an AI Centre of Excellence in private sector corporations, enhancing the Social and Ethics Committee for AI oversight, establishing a board technology committee and AI governance framework principles and having AI expert(s) at the board level.

Keywords: Corporate Governance, Boards of Directors, Artificial Intelligence (AI), AI Governance, AI Ethics, Structures, Processes, Rational mechanism, AI Governance Framework.

TABLE OF CONTENT

FIGURES

TABLES

LIST OF ABBREVIATIONS

ACM	Association of Computing Machinery
AI	Artificial Intelligence
AIG	Artificial Intelligence Governance
APP(s)	Applications
CEO	Chief Executive Officer
CG	Corporate Governance
CIO	Chief Information Officer
CNN	convolutional neural network
COBIT	Framework created by ISACA for IT management and IT governance
DL	Deep Learning
GDP	Gross Domestic Product
GPT	Generative Pre-training Transformer
IBM	International Business Machines Corporation
ICARFT	Integrity, Competence, Responsibility, Accountability, Fairness and Transparency
ICT	Information Communication Technology
IEEE	The Institute of Electrical and Electronics Engineers
IODSA	Institute of Directors South Africa
IoT	Internet of Things
ISO 38500	International standard for Corporate governance of information technology
IT	Information Technology
ITG	Information Technology Governance
ITIL	Information Technology Infrastructure Library
JSE	Johannesburg Stock Exchange
ML	Machine Learning
MOI	Memorandum of Incorporation
NASA	The National Aeronautics and Space Administration
NGO	Non-governmental Organisations
OECD	Organisation for Economic Co-operation and Development
POPIA	The Protection of Personal Information Act
RNN	recurrent neural network
UNESCO	The United Nations Educational, Scientific and Cultural Organization
XAI	Explainable AI

CHAPTER 1: INTRODUCTION

1.1. INTRODUCTION

Most businesses rely on Artificial Intelligence (AI) technologies for efficient operation and support functions in the current global information era (Weill & Woerner, 2017; Perifanis & Kitsios, 2023). As a result of globalisation and the creation of a global village, emergent technology in general and AI technologies, in particular, have become an essential element of business operations and a platform for businesses to interact with one another globally. AI technologies are becoming more popular as they become more powerful. Their application is gaining preference in everyday use processes and operations as these technologies use a methodology based on algorithms that require less human monitoring and supervision and can make predictions and generate decisions based on huge amounts of data in a much shorter time than humans can.

"AI can now recognise objects in images and videos; transcribe speech; translate between languages; beat humans at Jeopardy, at Go, and at poker; paint in the style of van Gogh; write Beatles-like music; help prepare legal documents; trade stocks; drive cars; fly drones; write its own encryption language; identify cancer in tissues; and solve the quantum state of many particles at once. In the coming years, it is expected that AI will reach and exceed human performance on many more, and increasingly complex, tasks" (Liao, 2020, p.1).

AI technologies have significantly achieved in areas such as education, science, health care and climate change. Some of the achievements include, amongst others, the Ochsner Health[1] AI system, which predicts hospital patients who are at imminent risk of cardiac arrest, allowing doctors and nurses to intervene in time to prevent an emergency. AI technologies promise enormous benefits for economic growth, social development, as well as human well-being and safety improvement. However, these technologies perform routine tasks independently and make difficult life-and-death decisions and other decisions that directly impact individuals in society, such as financial risk-related decisions, on their own (Borenstein & Howard, 2021). Such autonomous decision-making raises concerns about the collection, use, and quality of data used to train these technologies, as well as about the

---

[1] Ochsner Health based in New Orleans, US is a large health care system with unique clinical environments that span the spectrum from small community hospitals to large quaternary academic centers (Fixler, et al., 2023).

very notion of allowing AI technologies to make decisions that pose significant risks for humanity and societies.

Over the years, there have been growing fears that governments might use AI technologies to create a "big brother state" through mass surveillance programmes. Fears that these technologies might replace jobs for human workers, that autonomous weapons may be misused at the hands of malevolent characters, that their development eludes accountability, and that AI technologies inadvertently disseminate algorithmic bias and thereby undermine allocation and representational fairness abound.[2] These fears have been exacerbated by cases such as algorithmic discrimination in facial recognition systems, which have been found to have misidentified Africans and Asian people more often than white people[3]. Also, for instance, Amazon's machine learning recruiting algorithm was found to be biased against women[4]. The algorithm was trained on the number of resumés submitted over the past ten years and the candidates hired, where most of the candidates were men; thus, the algorithm also favoured men over women. Other examples of current ethical concerns about AI technologies include that of the IBM Watson-Texas MD partnership, where AI systems failed to give the right medicine to patients[5], and the Microsoft developed AI chatbot that was released but soon withdrawn after it learned Nazi propaganda and racist insults[6].

These and other cases have prompted AI ethics scholars (Wachter, Mittelstadt & Floridi, 2017; Buolamwini & Gebru, 2018; Noble, 2018; Richardson, Schultz, & Crawford, 2019; Suresh & Guttag, 2019) to raise a debate about the lack of transparency, poor accountability, rising inequity and amplification of structural bias by algorithms in the context of the research, design, development, deployment and use of AI technologies to the extent of undermining fundamental human rights. For example, Noble (2018), has raised a debate on how search engine algorithms can reflect and reinforce systemic inequalities present in

---

[2] Allocation harm is an immediate and economic harm and speaks to withholding access to certain services or goods based on some kind of structural bias in the training data set. Representational harm speaks to the underlying harm of categorizing society in terms of stereotyping and identity prejudice. See Crawford (2017).
[3] Many Facial-Recognition Systems Are Biased, Says U.S. Study
https://www.nytimes.com/2019/12/19/technology/facial-recognition-bias.html
[4] Amazon scraps a secret A.I. recruiting tool that showed bias against women
https://www.cnbc.com/2018/10/10/amazon-scraps-a-secret-ai-recruiting-tool-that-showed-bias-against-women.html
[5] How IBM's Watson Went From the Future of Health Care to Sold Off for Parts
https://slate.com/technology/2022/01/ibm-watson-health-failure-artificial-intelligence.html#:~:text=One%20of%20IBM%27s%20high%2Dprofile,What%20went%20wrong%3F
[6] Microsoft's new AI chatbot has been saying some 'crazy and unhinged things'
https://www.npr.org/2023/03/02/1159895892/ai-microsoft-bing-chatbot

society, highlighting the role of human prejudices embedded in AI technologies. The study by Suresh and Guttag (2019), pointed to various forms of bias in machine learning, such as historical bias, representation bias, and measurement bias, which can lead to discriminatory outcomes, particularly in areas like criminal justice and hiring practices.

These ethical issues and social concerns, among others, have made some groups in society shift from thinking purely about the functional capabilities of AI technologies to the ethics behind creating such powerful and potentially life-consequential technologies. International organisations, governments and big corporations have started to respond to these societal fears by developing AI governance and ethical frameworks. These include, amongst others, the European Commission high-level expert group on AI, the OECD expert group on AI in society, the Singapore Advisory Council on the ethical use of AI and Data, the UNESCO bio-ethics and ethics of Science and Technology team, the United Kingdom select committee on AI, and most recently, the UN Secretary-General's High Level Advisory Body on AI. Companies such as Google and SAP have publicly released AI guidelines and principles. Non-profit organisations such as the Association of Computing Machinery (ACM), Access Now, and Amnesty International have issued declarations and recommendations on the responsible use of IA (Schiff et al., 2021). Also, professional bodies, such as the IEEE, are playing core roles in this context.

Consequently, the comprehensive integration of AI technologies in private sector organisations has also put a spotlight on the board of directors' role in AI governance and ethics. Over the years, boards have increased strategic involvement and oversight in AI matters. Accordingly, the board has a fiduciary duty to ensure that adequate AI governance structures and ethics frameworks are in place to ensure transparency, accountability, fairness, privacy, robustness and safety. Inadequate oversight at the leadership level heightens the risk of various bad AI ethics practices (Ruttkamp-Bloem, 2020), such as copyright infringement, privacy violations, bias, etc.

As a result, this study proposes investigating how public listed companies boards perform their functions over the ethical application and governance of Artificial Intelligence technologies through a case study of Johannesburg Stock Exchange (JSE) Listed Companies and suggesting changes where ethical governance seems lacking. The JSE offered an excellent case study for analysing the function of the board in governing AI ethics due to its rigorous regulatory environment. Unlike private companies, JSE-listed companies

are obliged to publish annual integrated reports, which ensure the availability of substantial and reliable public information. This was crucial given the limited research on AI within South African companies and allowed for a comprehensive examination of corporate governance practices and how boards address AI ethics, providing a robust foundation for this study. Moreover, JSE-listed companies are mandated to adopt the King IV principles of corporate governance, which emphasise ethical leadership and corporate citizenship. These principles align closely with the ethical considerations necessary for AI governance and ethics. Additionally, Companies Act 71 of 2008 requires these companies to have both an audit committee and a social and ethics committee. These committees play a crucial role in overseeing ethical practices which impact society.

This investigation, being situated at the intersection of philosophy and business management and driven by research in the ethics of AI[7], aims to inform recommendations on the responsibility of boards in private sector companies in terms of AI ethics regulation implementation and governance.

1.2. RESEARCH PROBLEM STATEMENT

Over recent years, the media, both non-governmental organisations (NGO's) and intergovernmental organisations (such as UNESCO and the OECD), scholars, and members of civil society have raised concerns over the growing erosion of ethics in the research, development, deployment and use of AI systems and applications. The rapid increase in the significance of AI ethics and governance due to recent cases of algorithmic bias in facial recognition systems, Cambridge Analytica seeking to manipulate elections, and other examples of AI-driven decisions negatively impacting human rights, have highlighted the need for algorithms which are ethical, accountable, explainable, auditable, and transparent (Whittaker et al., 2018).

AI applications are evolving with increased sophistication, sparking considerable complex questions related to its technology's social impact, governance, and ethics. These questions are particularly salient as accountability mechanisms for algorithms are yet in a nascent stage even given recent activity in the US with the AI Bill of Rights and the UK AI Safety Summit, where the balance of power is skewed towards industry giants – many of them

---

[7] It should be understood that AI ethics is a sub-discipline of the ethics of AI (Ruttkamp-Bloem, 2020) which is in its turn a sub-discipline of ethics or moral philosophy, which is concerned with systematising, defending, and recommending concepts of right and wrong behaviour (and of good/bad character).

transnational – that control these technologies. Currently, the research, design, development, and deployment of AI primarily take place in the private sector, while governments worldwide are increasingly contracting out these powerful technologies (Gesser, 2017). As a result, the topic of AI ethics and governance is increasingly imperative for the attention of the board of directors of private companies. Thus, the importance of the debate on how boards oversee and monitor AI technologies is evident and is a discussion that is relevant to all organisations, whether designing and developing AI systems or buying AI-powered software and applications.

Corporate governance theories highlight the principles of ethical leadership, transparency, accountability, fairness, and responsibility to achieve an ethical culture, good performance, value creation, legitimacy, and effective control of the organisation (King IV Report, 2017). As the custodians of corporate governance in the organisation, boards are responsible for the strategic direction in responding to the ethical concerns that AI technologies may generate. Moreover, these technologies are becoming more and more embedded into the corporate processes; thus, there is an increasing need for the board to make complex decisions about ethical and governance issues, which may have legal and financial implications (Brand, 2020).

The qualitative study conducted by Torre, Teigland, and Engstam (2019) found that boards tend not to be aware of AI technologies in the organisation, especially when it comes to off-the-shelf AI-powered software and applications. A recent report commissioned by Microsoft, titled 'AI in the Middle East and Africa – South Africa outlook for 2019[8], also found that only 37% of non-executive directors considered AI technologies and their governance to be important topics at a board level. The report further indicated that only 38% of the organisations surveyed are driving their AI agenda from the top down.

To date, most studies on AI ethics generally focus either on the work of software developers or on regulating AI technologies in general, with little focus on the board's role or the board's responsibility in overseeing and monitoring these technological advancements in the private sector companies (Kambil and Lucas, 2002). Despite this lack of focus on the board's role in AI ethics, recent research has highlighted the importance of corporate governance in

---

[8] Artificial Intelligence in Middle East and Africa South Africa Outlook for 2019 and Beyond
https://info.microsoft.com/rs/157-GQE-382/images/MicrosoftSouthAfricanreportSRGCM1070.pdf

ensuring responsible AI practices within private sector companies (Kambil and Lucas, 2002; Chung, 2017). Additionally, there have been growing calls for greater transparency and accountability in AI decision-making processes, which can only be achieved through effective board oversight (Cheong, 2024). As such, there is a need for further research on the role of boards in promoting ethical AI practices and ensuring responsible use of AI technologies in the private sector. Boards have an important role in ensuring an ethical AI culture and effectively managing both the potential of AI technologies and their organisational risks, including ethical, programming, and reputational risks.

This research seeks firstly to explore how boards perform their functions over artificial intelligence ethics and governance in JSE-listed companies by means of an inter-disciplinary study driven by a phenomenological, selective literature review and philosophical research methodologies. This investigation, strengthened by a thorough literature review completed in the philosophical tradition, will inform the realisation of the final objective of the research, which is to formulate recommendations on the role of boards in private sector companies in terms of AI ethics regulation actualisation and governance from within an interdisciplinary context. This final objective will be informed by semi-structured interviews conducted in JSE Listed Companies, and the results will be analysed using a phenomenological approach against the background of the literature review.

1.3. PURPOSE STATEMENT AND OBJECTIVES

The purpose of the research is to investigate how boards perform their functions regarding ethics and governance of artificial intelligence by means of a case study of JSE-listed companies. The study will explore the board's role as a focal point of Corporate Governance, leading and setting an ethical tone relating to ethically responsible governance of AI technologies and ensuring that emerging AI technologies support in setting strategic objectives and goals that are ethically acceptable. This investigation on the empirical role of the board in managing AI ethics and AI technology governance concerns, strengthened by a thorough literature review, will inform the further objective to formulate recommendations on the role of boards in private sector companies in terms of AI ethics and governance.

1.4. RESEARCH QUESTION

The primary research question is: What is ***the role of the board in governing artificial intelligence ethics – a case for JSE Listed Companies***? The following thematic research questions are relevant to the study:

- *How does the board discharge its fiduciary responsibility in overseeing AI technologies and its potential impact on the organisation?*
- *What is the board's role in ensuring that AI technologies are aligned with the organisation's ethical culture?*
- *What is the board's role in ensuring that the research or design or development or use of AI technologies does not negatively impact on society?*
- *What is the board's role in ensuring that adequate AI governance structures are in place?*

1.5. RESEACH METHODOLOGY

To realise the objectives of the study, which is to explore how boards perform their functions over governance of AI technologies ethics, the researcher used a qualitative research approach. The qualitative interpretation was executed within the discussion of corporate governance and AI ethics. As such, corporate governance and AI ethics research requires more qualitative investigation as qualitative research can assist policy-makers and practitioners to develop more efficient governance mechanisms, by shedding light on the efficacy of policy prescription. The design uses phenomenological methodology (Creswell, 2003). The study focuses on the understanding and experience of the board members as it attempts to unfold meanings of/in lived experiences of board members. In this context, the phenomenological approach matches the assumptions and aspirations of the study. The study is fundamental and provides a window to what can be 'directly', phenomenologically known, because all knowledge is captured from the board members' and specialists' experiences and perspectives.

A purposive sampling was used for the purposes of the study. The researcher selected twenty (20) JSE listed companies board members and five (5) specialists in the area of artificial intelligence who interact or report at the board level i.e., Chief Information Officers, for semi-structured interviews. The researcher selected a semi-structured interview approach for data collection mainly because it allows the participants an opportunity to express their experiences, views, and opinions (Lauer, 2006). Moreover, this allowed the

researcher to have an open conversation with board members and AI specialists in order to obtain their understanding of the subject matter. The researcher also utilised secondary data to compare the study with readily available information. Secondary data refers to already collected data such as surveys, observations, experiments, questionnaires, personal interviews, websites, books, journal articles, internal records, etc. Analysis and discussion of this work are mainly captured in the literature review. For analysis, the researcher conducted thematic analysis through transcription of verbal data retrieved in interviews with the participants, then developed significant statements, which were then grouped into a theme to be critically reflected on in order to formulate findings of the study and to inform final recommendations.

1.6.   SIGNIFICANCE OF THE STUDY

The rapid increase in the significance of AI ethics and governance due to cases of algorithmic bias, privacy and safety concerns, and human rights, have highlighted the need for algorithms which are ethical, accountable, explainable, auditable, and transparent (Whittaker et al., 2018). The research provides a theoretical foundation for understanding the intricate dynamics between corporate boards and the ethical implementation of AI technologies. This investigation is situated at the intersection of philosophy and business management and is driven by research into the ethics of AI.

From a philosophical perspective, the study engages with ethical theories and principles, exploring how they intersect with the practical decision-making processes within boards. These ethical issues and social concerns have made society shift from thinking purely about the functional capabilities of AI technologies to the ethics behind creating such powerful and potentially life-consequential technologies. The infusion of philosophy into the study adds depth to the examination of AI ethics, considering questions related to moral responsibility, accountability, and the societal impact of AI technologies.

From a business perspective, the board, as the focal point of governance, is responsible for strategic decision-making, risk management, and ensuring that AI applications align with ethical standards, legal frameworks, and stakeholder expectations. The board's role in navigating AI's ethical implications is pivotal for promoting responsible and sustainable business practices, safeguarding corporate reputation, and fostering trust among investors, customers, and the broader society. Thus, the study contributes to the ongoing scholarly

discourse on how corporate governance structures can adapt to the challenges posed by rapid technological advancements.

This interdisciplinary approach enriches the academic landscape by integrating insights from both fields to provide a comprehensive understanding of the ethical implications of AI development, deployment, and overall adoption. This interdisciplinary lens contributes to the academic significance of the research, offering a unique and comprehensive exploration of the ethical dimensions inherent in the intersection of artificial intelligence, corporate governance, and philosophical inquiry. Finally, the interdisciplinary nature of the study enables the significant novel contribution of analysing the role of the board in AI ethics adoption and governance in private sector companies from the viewpoint of AI ethics literature as well as a corporate governance perspective.

1.7. STUDY LAYOUT

The remaining chapters of the study are briefly outline below:

```
Chapter 1:          →   Chapter 2: Artificial   →   Chapter 3:              →   Chapter 4: Artificial
Introduction            Intelligence                Corporate                   Intelligence Ethics
                                                    Governance and
                                                    Artificial Intelligence
                                                                                        │
                                                                                        ↓
Chapter 5: Artificial  →  Chapter 6: Research  →  Chapter 7: Data   →   Chapter 8: Summary
Intelligence              Approach and            Collection and        and
Governance                Methodology             Analysis              Recommendations
        │
        ↓
Chapter 9:
Conclusion
```
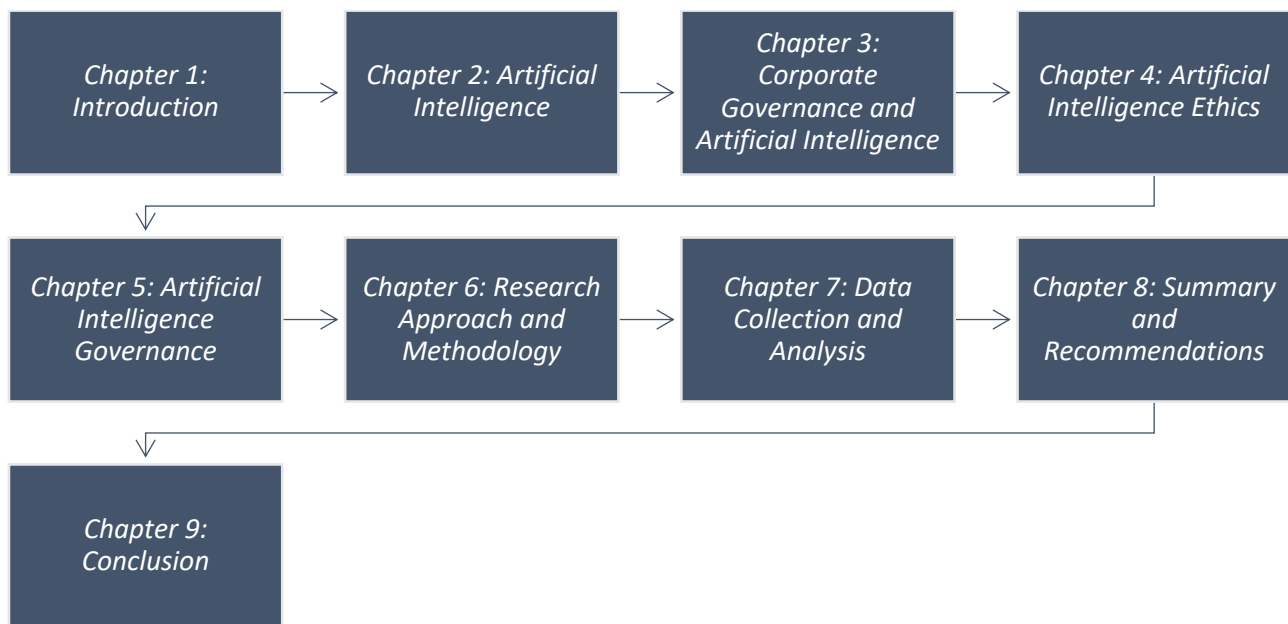
Chapter 2: Artificial Intelligence
The chapter explores an overview of definitions and history of artificial intelligence. The chapter also focuses on the key concepts of AI, such as explanation of machine learning, deep learning and generative AI. The chapter concludes with the importance of AI in business, global and local contexts in implementing AI.

Chapter 3: Corporate Governance and Artificial Intelligence

This chapter dives into the theoretical framework of corporate governance, which includes a discussion on the agency theory, stewardship theory, resource dependence theory, resource-based view theory and stakeholder approach theory. The chapter further details the different board models, the fiduciary duty of the board in terms of artificial intelligence technology and the conceptual framework of the role of the board in governing AI.

Chapter 4: Artificial Intelligence Ethics

The chapter focuses on broad discussion of ethics (metaethics, normative ethics and applied ethics), AI ethics, and the role of the board in AI ethics. The chapter focuses on the most common AI Ethics principles, i.e. accountability and responsibility, fairness, transparency, explainability, and privacy. It further analyses organisational ethics in the era of AI technologies and reflects on AI technologies' alignment to corporate ethical culture (values).

Chapter 5: Artificial Intelligence Governance

This chapter is an in-depth discussion on AI governance as a subset of corporate and IT governance. The chapter further discusses the different IT frameworks and dives into AI governance as part of organisational corporate governance structure, including AI governance frameworks that boards can utilise to govern AI technologies.

Chapter 6: Research Approach and Methodology

The chapter details the discussion of the research design used for the study. The theoretical framework for the study is described, followed by the paradigms applied. The chapter also outlines the methods used for research population, sampling and analysis. Lastly, the chapter concludes by discussing the study's validity, reliability and ethical considerations.

Chapter 7: Data Collection and Analysis

The chapter first presents the data collected from the semi-structured interviews conducted with the participants. The chapter also discusses the detailed analysis of the data collected from participant interviews.

Chapter 8: Summary and Recommendations

The chapter provides an overall summary of the thematic analysis and provides recommendations.

Chapter 9: Conclusion

The chapter draws conclusions from the research results and indicates how the research objectives were met. It also provides an overview of the novelty of the study, limitations and future research recommendations.

1.8.   CONCLUSION

This chapter gave an overview of the business reliance on artificial intelligence (AI) technologies for efficient operations and support functions to achieve competitive advantage, its achievements in different industries, ethics concerns raised, and the spotlight on the board's role in governing AI ethics. The chapter further outlines the research problem, which focuses on the boards' lack of AI awareness in general, and the lack of awareness of the board's role in overseeing and monitoring AI technologies in particular. Thus, the research aims to investigate how boards perform their functions regarding the governance of AI technologies through a case study of JSE-listed companies. The chapter further highlights the qualitative research design using phenomenological methodology to understand the lived experiences of board members in governing AI ethics. Furthermore, the chapter details an overview of the significance of the study due to the rapid increase in the need for AI ethics reflection and responsible governance of AI technology due to cases of algorithm bias, privacy, and safety concerns.

CHAPTER 2: ARTIFICIAL INTELLIGENCE

2.1. INTRODUCTION

The chapter explores an overview of definitions and history of artificial intelligence. The chapter also focuses on the key concepts of AI, such as explanation of machine learning, deep learning and generative AI. The chapter further discusses the importance of AI, especially in health, agriculture and education. At the conclusion, the chapter details the global spending on AI by corporations, projected future spending, and South Africa's context in developing and adopting AI technologies.

2.2. AI DEFINITION

The point of departure in defining artificial intelligence should start by answering the question, "What is intelligence?". This complex question has puzzled philosophers for centuries. For instance, Coppin (2004) argues that intelligence could be defined by its characteristics, i.e., the ability to deal with new situations, solve problems, answer questions, devise plans, and so on. It is more challenging to define the difference between the intelligence exhibited by humans and that exhibited by animals (Coppin, 2004), which raises questions on the distinction between human and machine intelligence. While it is true that defining the difference between human and animal intelligence is difficult, it is important to note that there are still significant differences between the two. For instance, humans possess the ability to reason abstractly, while animals do not. Additionally, humans have the capacity for self-awareness and consciousness, which is not present in animals. Therefore, while there may be similarities between human and machine intelligence, it is important to recognise that machines lack the essential qualities that make human intelligence unique.

The introduction of natural language processing models has accelerated the debate on human vs. machine intelligence. Recently, the debate was brought to the attention of the public by the introduction of Generative AI technologies, i.e., Generative Pretrained Transformer (GPT) developed by OpenAI, the Bard tool developed by Google, DALL·E – a generative AI model for image synthesis, the AlphaCode transformer-based language model with 41.4 billion parameters, etc. Many argue that generative AI might redefine intelligence as it has elements of what Coppin (2004) argue as being characteristics of intelligence i.e., ability to deal with new situations, solve problems, answer questions, devise plans.

Generative AI technologies are highly advanced language-processing AI models that can generate text similar to human language. It has many practical uses, including language translation, modelling, and creating text for chatbots. For example, GPT-4 is currently one of the largest and most sophisticated language AI models available, with a total of 1.76 trillion parameters parameters (OpenAI, 2023). However, like all AI technologies, it is not capable of independent thought or consciousness. It is an AI technology tool designed to perform certain tasks and functions based purely on the data and algorithms that have been programmed. In the case of GPT-4, the training data information is up until April 2023. However, it is certainly a highly advanced and sophisticated tool that is capable of performing many complex language-processing tasks with a high degree of accuracy.

According to Russell & Norvig (2002), the field of artificial intelligence attempts to understand "intelligent entities". The academic domain of AI is regarded as a branch of computer science that deals with the development of algorithms and the study of computation activities that can simulate or even recreate the cognitive capabilities of human beings (Hu, 2008; Vassilopoulos & Georgopoulos, 2010). The earlier concept of AI can be traced back to Alan Turing in the 1950s (Sidorova & Rafiee, 2019). Alan Turing first considered the question, "Can machines think?" in his seminal paper, 'Computing Machinery and Intelligence', published in 1950. To answer the question "Can machines think?" Turing proposed a game known as the Turing test. He argued that if a machine could imitate humans through knowledge processing, representation, and reasoning and could mislead observers to believe it is human, then certainly it should be considered intelligent (McCarthy, 2007). To 'pass' the Turing test, a computer would need at least the following skills (Kok et al., 2009):

- Natural language processing should communicate in a natural language such as English.
- Knowledge representation: it should have the knowledge and the ability to store it.
- Automated reasoning: it should be able to reason based on the stored knowledge.
- Machine learning: the ability to adapt to new circumstances and to detect patterns.

Critics argue that the Turing test is a one-sided test by considering the machine that passes the test as "certainly intelligent". McCarthy (2007) argued that a machine can still be considered intelligent without knowing enough about humans to imitate a human. In their turn, many philosophers have argued that a machine that passes the Turing Test would still

not actually be thinking but would perform only a simulation of thinking (Stuart & Peter, 2016), as there would not necessarily be semantic understanding.

The term "artificial intelligence" was introduced into academic debates in 1956 at the Dartmouth conference by John McCarthy and other colleagues such as Simon Newell and Marvin Minsky (Skilton & Hovsepian, 2017). The actual definition of AI was coined in 1973 at the Lighthill Debate by John McCarthy (Lighthill Debate, 1973):

"Artificial Intelligence (AI) is a science; it is the study of problem solving and goal achieving processes in complex situations… AI is a basic science, like mathematics or physics, and has problems distinct from applications and distinct from the study of how human and animal brains work. It requires experiment to carry it out, it involves a very large number of parts to it …".

The Encyclopedia Britannica (1991) defines AI as the ability of digital computers or computer-controlled robots to solve problems that are normally associated with the higher intellectual processing capabilities of humans". Ertel (2018) argued that this definition has some limitations as it suggests that a computer with large memory and the ability to save long text and retrieve it on demand displays intelligent capabilities. Thus, according to this definition, every computer is an AI system as long as it has a higher intellectual processing capability than humans.

Other definitions by different scholars are shown in Table 1 below.

Table 1: Table 1: Some definitions of AI (Russell & Norvig, 2002)

AI definition	Author, year
"The exciting new effort to make computers think . . . *machines with minds,* in the full and literal sense"	(Haugeland, 1985)
"A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes"	(Schalkoff, 1990)
"The branch of computer science that is concerned with the automation of intelligent behavior"	(Luger and Stubblefield, 1993)
"The study of the computations that make it possible to perceive, reason, and act"	(Winston, 1992)
"The study of how to make computers do things at which, at the moment, people are better"	(Rich and Knight, 1991)

"The study of mental faculties through the use of computational models"	(Charniak and McDermott, 1985)

According to Russel & Norvig (2002), these definitions are organised into four categories or approaches: AI that thinks like humans, AI that acts like humans, AI that thinks rationally and AI that acts rationally. Historically, all four approaches have been followed. However, there has been an existing tension between humans and rationality centred approaches. Russel & Norvig (2002) argued that a human-centred approach should be an empirical science involving hypothesis and experimental confirmation, while a rationalist approach involves a combination of mathematics and engineering. People in each group sometimes cast aspersions on work done in the other groups, but the truth is that each direction has yielded valuable insights.

Over the years, the definitions of artificial intelligence have begun to shift based upon the goals set for an AI system. The European Commission (2018) refers to AI technologies as systems that display intelligent conduct by analysing their environment and can perform various tasks independently to achieve specific goals.

Today, the AI technologies at issue are data-driven and governed by machine learning and probability algorithms. For decades, however, logic-based AI dominated AI, which focuses on the formal representation of knowledge processes such as reasoning and decision-making. This branch of AI continues, and many predict that at some stage, it will become necessary to incorporate it into data-driven AI – for many reasons, but most of them relate to semantic and related ethical concerns. This study focuses on the data-driven AI approach because this is the main approach at present and it is about improving data quality and data governance in businesses and aims to improve the performance of a specific problem (Xu & Sun, 2018). Data-driven AI generates technologies (AI systems) that have the ability to act, sense, reason and adapt almost entirely independently from human interference. This is mainly because data-driven AI heavily relies on vast data to train their algorithms.

In contrast, the logic-driven AI approach refers to the process of using a model constructed based on the objective, physical mechanism and domain knowledge for a specific task. A prominent feature of the model-driven approach is that when the model is sufficiently accurate, the solution can be generally expected to be optimal, and the minimisation algorithm is commonly deterministic. Xu & Sun (2018) argued that a flaw of the model-driven

approach lies in accurately modelling for a specific task in real applications, and sometimes the pursuit of accurate modelling is a luxury expectation.

Marwala (2020) simplifies the definition of data-driven AI as "the technology that brings intelligence to machines". Data-driven AI technologies have the ability to emulate some of the cognitive functions that we attribute to the human mind (Syam, Sharma, 2018). This implies that currently AI is mainly about the development of algorithms that can engage in a human-like thought process, i.e. learning, reasoning, problem-solving and self-correction. Thus, for the purpose of the study, one concludes that AI can be defined as the science that attempts to have computers perform and execute tasks that humans would perform, in part or full, and in some instances, AI systems perform even better than humans.

2.3. HISTORY OF ARTIFICIAL INTELLIGENCE (TIMELINES)
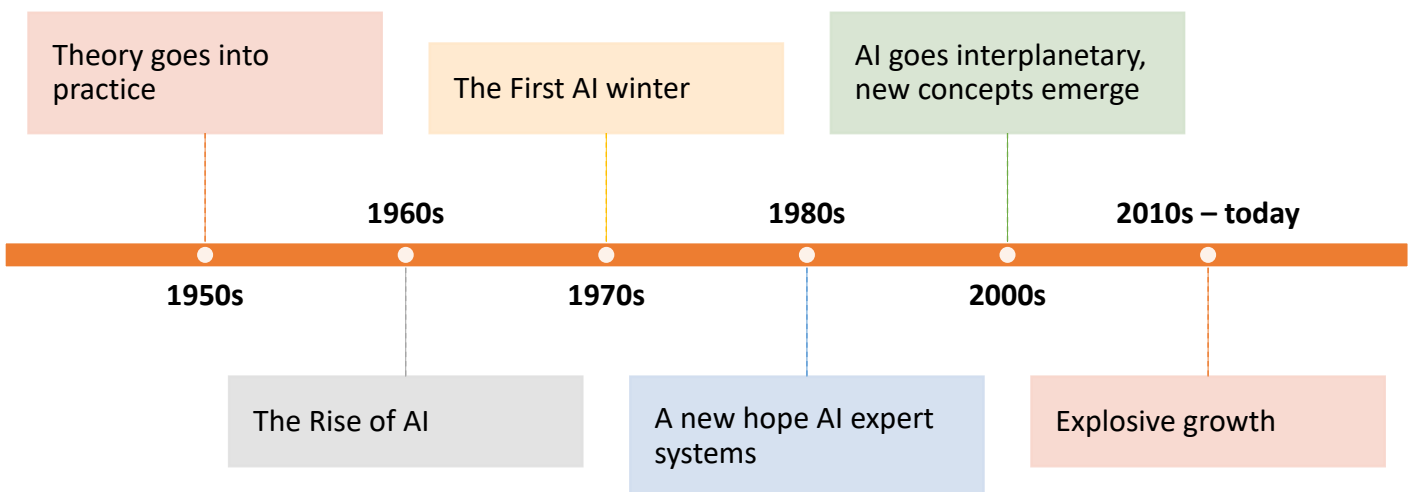
Figure 1: The history of AI from the 1950s to the 2000s.

- *1950s: Theory goes into practice*
  In the 1950s, the field of AI transitioned from theory into practice with figures such as Alan Turing, John McCarthy, Marvin Minsky, Simon Newell, and others laid the foundation for AI by proposing key concepts and developing early algorithms (Lighthill Debate, 1973; Skilton & Hovsepian, 2017). At the time, the practical exploration sought to create a machine that could mimic human intelligence. Artificial intelligence was officially introduced in 1956 at a conference at Dartmouth College, which included a session titled Artificial Intelligence for the first time (Cannataro et al., 2022).

Despite initial optimism, computational power limited AI progress until 1970, after this first period.

- *1960s: The Rise of AI*

  AI gained the attention of researchers with demonstrations such as Newell and Simon's General Problem Solver (Newell, Shaw & Simon, 1959) and Joseph Weizenbaum's ELIZA, which showed promising results toward the goals of problem-solving and the interpretation of spoken language, respectively. In particular, ELIZA created Weizenbaum (1966), a natural language processing computer program, demonstrating the potential of AI in understanding and generating human language. During that period, another important milestone was the introduction of the first industrial robot arm, Unimate, by General Motors factor (Gasparetto & Scalera, 2019).

- *1970s: The First AI winter*

  Despite the initial optimism, the lack of significant progress led to a period known as the "AI Winter". The AI Winter was characterised by reduced funding and interest in AI research due to its failure to achieve its ambitious goals. The hype and high expectations caused by the media and the false predictions and exaggerations by the experts in the field about their outcome led to significant funding cuts in AI research (Toosi, Bottino, Saboury, Siegel & Rahmim, 2021). During the AI winter, commercial and scientific activities in AI declined dramatically.

- *The 1980s: A New Hope AI Expert systems*

  After the *1970s* AI winter period, the 1980s marked a period of renewed optimism in the field characterised by the rise of expert systems. AI researchers began developing systems that sought to emulate human expertise in specific domains, employing rule-based reasoning to solve complex problems. Examples of such applications were DENDRAL (designed for chemical analysis), XCON (a computer hardware configuration system), MYCIN (a system for diagnosing bacterial infections), and ACE (AT&T's cable maintenance system). At that time, AI expert systems were considered revolutionary solutions because of their capability to solve problems in any area of human activity and were also perceived as a direct threat to humans (Kurzweil, 2005). During this period, over half of the Fortune 500 companies were involved in developing expert systems (Singh, 1999), and the usage grew at a rate of 30% a year (Kurzweil, 2005).

- *The 2000s: AI goes interplanetary, and new concepts emerge*

  In the 2000s, the field witnessed a notable expansion of its applications, marked by advancements in interplanetary exploration and the emergence of new conceptual paradigms. New AI concepts such as reinforcement learning, deep learning, and neural networks emerged, gained prominence, and demonstrated exceptional performance in various domains. For example, machine learning algorithms, which incorporated autonomous navigation and decision-making capabilities, were used to analyse vast amounts of datasets for NASA Mars exploration (Maimone, Biesiadecki, Tunstel, Cheng & Leger, 2006). Deep learning architectures, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), created basic building blocks in constructing complex deep learning solutions for image, speech recognition, language translation, and other complex tasks (Kamath et al., 2019). The synergy of advanced algorithms, increased computational power, and large datasets paved the way for transformative applications of AI in diverse fields, shaping the trajectory of artificial intelligence into the subsequent decades.

- *2010s – today (2024): Explosive growth*

  From 2010 to 2021, the global investment in AI-based start-up companies grew from $1.3 billion to $40 billion (Zhang et al., 2021). In 2023, more than 25% of all investment has been channelled into AI-related US start-ups, a significant increase from previous years, which averaged about 12% between 2018 and 2022 (Mittal, 2023). The current period has witnessed explosive growth in AI development, deployment, and investment, marked by breakthroughs in various domains and the proliferation of AI applications in everyday life. One significant development is the widespread adoption of deep learning techniques, mainly fuelled by the success of deep neural networks. Generative Pre-trained Transformer (GTP) models, including GPT-3 (OpenAI's), have garnered attention for their ability to generate coherent and contextually relevant human-like text. For example, ChatGPT, developed by OpenAI, Google's PaLM[9] and Together's GPT-JT[10], notably exemplifies the power of large language models for conversational AI, enabling natural and context-aware interactions.

---

[9] Google's PaLM https://ai.google/discover/palm2/

[10] Together's GPT-JT https://www.together.ai/blog/releasing-v1-of-gpt-jt-powered-by-open-source-ai

2.4. KEY CONCEPTS OF AI

There are key AI concepts such machine learning, deep learning, generative AI, algorithm, neural networks, computer vision, natural language processing (NLP), robotics and autonomous systems, reinforcement Learning, datasets which collectively contribute to the multifaceted and dynamic field of artificial intelligence. Most current AI technologies incorporate machine learning, as intelligent behaviour needs extensive information or knowledge (Sharma, Sharma, & Jindal, 2021). However, machine learning, deep learning, robotics and generative AI are layers of data-driven AI (Janiesch, Zschech, & Heinrich, 2021). Below is the detailed explanation of machine learning, deep learning and generative AI.
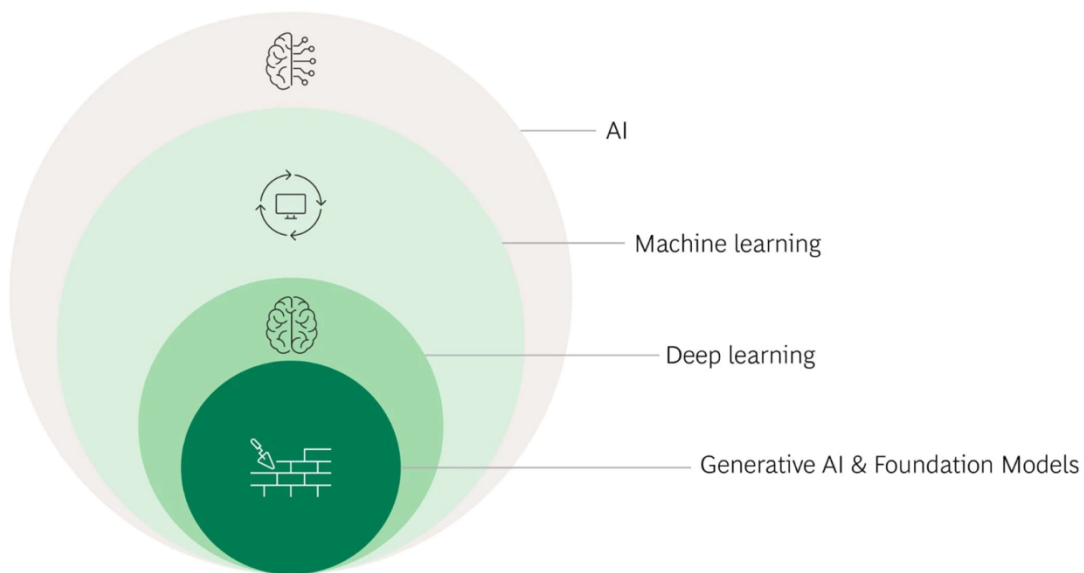
Figure 2: AI, ML, DL, and Generative AI Source (Purohit, 2023).

2.4.1. Machine learning (ML)

Machine learning (ML), as duplicated in Fig.3 below, is a sub-layer of AI. ML is a discipline that addresses the question of how to build AI technologies that improve automatically through experience. Jordan & Mitchell (2015) argue that ML is one of today's most rapidly growing technical fields, at the intersection of computer science and statistics and at the core of artificial intelligence and data science. Machine learning is a subset of AI that enables systems to learn from data and improve performance over time without being explicitly programmed. It involves the development of algorithms and models that use statistics and

pattern recognition to learn automatically, similar to how humans use sensory and other data, make predictions, and adapt to new information. The learning process allows AI systems to automatically adjust their behaviour and improve their decision-making abilities based on the data they encounter.

ML algorithms have the capability to gather a vast amount of data, analyse it and learn ways to improve efficiency and effectiveness (Marwala, 2020) through enhanced classification abilities. ML algorithms make decisions based on the data input from programmers, and their performance improves automatically through more exposure to data feeds and experience. In recent years, ML has emerged as a method of choice to develop practical solutions such as virtual personal assistants, speech recognition, image processing, video surveillance prediction, classification, regression, etc. (Jordan & Mitchell, 2015).

2.4.2. Deep learning (DL)

Deep learning (DL) is a sub-layer of data-driven AI, which is modelled similarly to the neurons of human brains. DL has been widely recognised as the representative advance of ML or AI, and this has been attributed to recent breakthroughs in a series of challenging applications (Gregor & LeCun, 2010). Some of these breakthroughs include an improved accuracy rate of face recognition higher than 99% at the human level (Schroff, Kalenichenko, & Philbin, 2015), reaching the optimal performance level of a simultaneous interpreter on speech recognition and translation (WU et al., 2016), and matching the level of senior physicians in medical diagnosis (Gulshan et al., 2016).

DL technologies do not require programmers to tell them what to do with data (Marr, 2018); it learns from its experience but needs a large database or large information provided at the input stage (Lee, 2018). DL technologies apply multiple layers to learn and make decisions, making them suitable for analysing complex data such as language, images, and sound.

Lee (2018) argues that DL technologies can do better than humans at "identifying faces, recognising speech, and issuing loans" (p.9). This is a contentious remark as we know that data-driven AI systems cannot recognise all faces of all races and genders, lose nuances of languages in Natural Language Processing contexts, and are often biased in terms of the grounds on which they decide loans and financial risk. This, in fact, points very clearly to the urgency of engaging with the ethics of AI, as the remark by Lee (2018) quoted above was

made solely in terms of the computational ability of the systems at issue, completely separate from human, social, and ethical considerations. Such remarks lead to false trust in these systems, which is highly problematic. This is why interdisciplinarity becomes crucial in the study of AI systems and their behaviour. This is also why the role of the board in AI technology governance is important in a business context, as it is very easy to fall into variants of technological solutionism in the business context and to forget about ethical or social concerns about the technology so deeply admired.

2.4.3. Generative AI

The introduction of applications such as ChatGPT, DALL·E, or Midjourney, which make Large Language Models (Yang et al., 2023) accessible to end-users, has set a milestone in the application of artificial intelligence to content generation, enabling wide audiences to effortlessly engage in the creation realistic content such as images, text, or even music. Thus, the introduction of the term Generative AI. However, the concept 'Generative Artificial Intelligence' lacks a universally accepted definition which can result in misunderstanding and misinterpretation (García-Peñalvo & Vázquez-Ingelmo, 2023). Aydın & Karaarslan (2023) argue that Generative AI is an artificial intelligence field that concentrates on generating new and original information by machine learning on massive databases of experiences. Purohit (2023) regards Generative AI as a branch of artificial intelligence and a subset of Deep Learning, focusing on creating models capable of generating new content that resemble existing data. These models aim to generate content that is indistinguishable from what might be created by humans.

2.5. IMPORTANCE OF AI IN BUSINESS AND SOCIETY

AI technologies have been essential tools in addressing some societal and business challenges; they enable human capabilities to be undertaken by algorithms increasingly effectively, efficiently, fast, and at low cost. The importance of AI technology comes down to its ability to rapidly analyse large amounts of data. Different scholars have argued the importance of AI technologies for health care, education, insurance, law and compliance, climate change (Liu et al., 2022), etc. Below is a high-level summary of the importance of AI technologies in five sectors (health, education, business, environment and agriculture) but these technologies have beneficial applications in many additional ones. However, it is important to note that AI has beneficial applications beyond these five sectors. The study

focused on these sectors due to their substantial and immediate impact, but AI's potential extends to numerous other fields such as transportation, finance, entertainment, and public safety.

2.5.1. Healthcare. AI technologies have the potential to improve the healthcare sector. Some companies have developed AI applications to screen medical images such as mammograms, tools that scan patient health records to predict future health risks, apps that help people monitor their health, and systems that help track disease outbreaks.

On drug discovery, AI technologies accelerate time to market and reduce uncertainty. Since the outbreak of Covid-19, different AI technologies have been used to predict potential epitopes to design and develop vaccines (Park et al., 2011; Yang and Leibowitz, 2015; Ton et al., 2020).

DL technologies can replace complex, human-coded sets of probabilistic rules and identify subtle correlations between vast, multi-variate data sets to deliver scalable, efficient, automated diagnosis (MMC Ventures, 2019).

2.5.2. *Education.* The use of AI technologies in education is rapidly expanding (Roos, 2018), and one of the most popular AI technologies used to support teaching and learning is the Chatbot system. Chatbots are online chat conversations via text or text-to-speech (Mabunda and Ade-Ibijola, 2019) that provide an instant response to the user. They are increasingly used to improve students' [learning]. Chatbots can instantly provide students with standardised support, such as practice questions and answers, evaluation criteria, assignment due dates, campus path direction, and generating study materials (Mabunda and Ade-Ibijola, 2019; Cunningham-Nelso et al., 2019). GPT-4 also plays a significant role in ensuring equal access to information by learners, especially those from disadvantaged backgrounds.

Furthermore, AI technologies can accelerate personalised learning to support students with special needs. The algorithms can also help deliver predictive analysis at the system level to reduce dropout and assess new skills (Vincent-Lancrin, Van der Vlies, 2020). AI technologies can also predict if a school teacher will have the most excellent value-added system (Rockoff, Jacob, Kane, & Staiger, 2010). AI

technologies, thus, may assist in making the online learning environment efficient and optimal in general.

2.5.3. ***Business.*** *The h*ave numerous tangible use cases of AI technologies enabling corporate revenue growth and cost savings in existing sectors to date. Many of the AI technologies are geared towards increasing productivity and efficiency in the workplace. Due to its ability to process big data, AI technologies can test different production and outcome possibilities models and be more precise in the analysis. Most businesses now use AI technologies for predictive maintenance to improve operational efficiency, impacting profitability.

On the revenue growth, AI technologies have the ability to detect very weak signals, thus helping companies develop, refine, and generate numerous forecasts. The processing speed at which AI technologies work allows companies to analyse large amounts of data to make decisions in real time (Day & Schoemaker, 2005). By improving the accuracy of predictions and by enabling real-time decisions, these technologies help companies generate more revenue.

2.5.4. ***Environment.*** ML supports power generation and distribution efficiencies, from autonomous maintenance and leak monitoring to route optimisation and fleet management. AI technologies such as Google's Deepmind have the ability to predict wind patterns up to 36 hours in advance to optimise wind farms (Marr, 2021). The study by Khatry et al. (2021) found that AI could positively enable 93% of the environmental targets, including the creation of smart and low-carbon cities; Internet-of-Things devices and appliances that can modulate their consumption of electricity, better integration of renewable energy through smart grids; the identification of desertification trends via satellite imagery; and combating marine pollution.

2.5.5. ***Agriculture and food production.*** In a developing country such as South Africa, where agriculture is among the key contributors to the GDP, AI technologies can help in precision agriculture by improving the overall harvest and accuracy. AI sensors can provide farmers with real-time insight to detect and target weeds, identify areas for irrigation, fertilisation, and pesticide treatment and then decide which herbicide to apply within different regions (Saleem, Potgieter & Arif, 2021; Jin, et al., 2022). Innovative farming practices such as vertical farming, which uses data-driven AI

technologies, may help increase food production while minimising the use of resources and ensuring food security.

As with every emerging technology, the importance and benefits of AI technologies are also accompanied by challenges and risks. The challenges and risks relate to AI technologies' security, privacy, biases, accuracy, etc. AI (ML & DL) algorithms learn based on the data they are trained from; thus, the integrity of data is paramount as, for instance, if data collected in the past have underlying biases, those biases will be present in future predictions. Hence, the importance of the boards in ensuring responsible governance and ethical culture in managing both the potential of AI technologies and their organisational risks, at least in terms of privacy, security and robustness, fairness and non-discrimination, transparency, accountability and responsibility.

2.6. GLOBAL TRENDS AND LOCAL CONTEXT

2.6.1. Global trends

The McKinsey Global Survey on Artificial Intelligence (AI) found that more organisations are using AI technology as a tool for generating value and revenue (McKinsey, 2023). Bughin et al. (2018) estimate the potential value of AI technologies for the global economy by 13 trillion U.S. dollars by 2030. Accenture's report (2016) indicates that by 2035, AI could double annual growth rates of gross value added (a close approximation of GDP) in 12 developed countries that made up 50 per cent of global GDP in 2016.
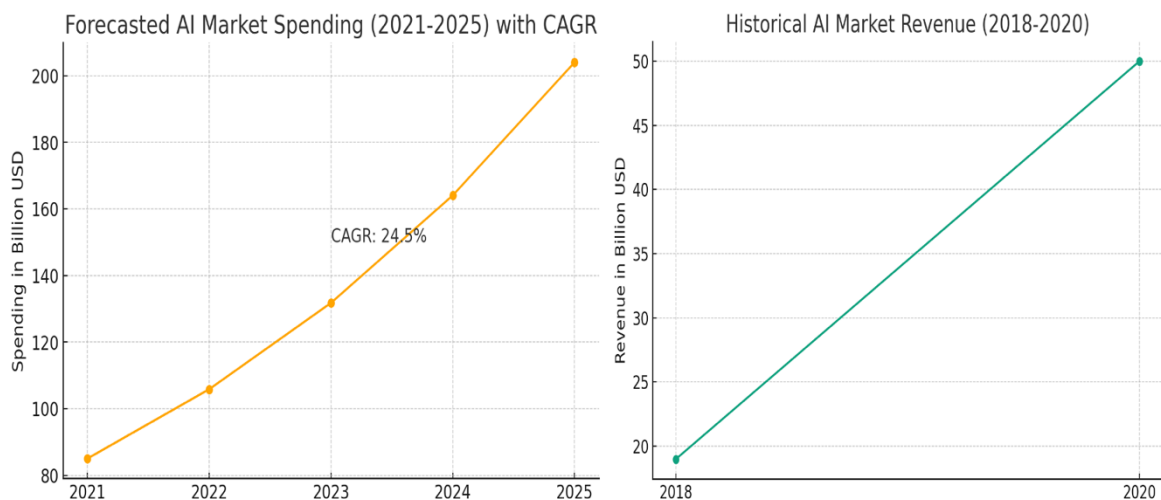
Figure 3: Forecasted AI Market and AI Historic Revenue. Source International Data Corporation (IDC)

The global revenue for the AI market, which includes software, hardware, and services segments, was around 281.4 billion U.S. dollars in 2020. It is estimated that in that same year, the AI software segment brought in 88 per cent of the total AI market revenues with 247.7 billion U.S. dollars (Liu, 2021). According to the International Data Corporation (IDC), global spending on AI has increased from 19 billion U.S. dollars in 2018 to 50 billion U.S. dollars in 2020. IDC further forecasts that the spending will double over the next three years, from 85 billion U.S. dollars in 2021 to more than 204 billion U.S. dollars in 2025 (*See Figure 3 above)*, meaning that the compound annual growth rate (CAGR) for the 2021-2025 period will be 24.5%.

The drivers for AI adoption are economic, such as reduction in cost and time, increased performance and customer experience, and more accurate predictions and decision-making (Cubric, 2020). Automated customer service agents, sales process recommendation and automation, automated threat intelligence and prevention, and IT automation are the key use cases of AI technology. These four use cases represent nearly a third of all AI spending this year. Some of the fastest-growing use cases are automated human resources, IT automation, and pharmaceutical research and discovery (IDC, 2020).

2.6.2. South African context

Many businesses in South Africa (SA) have been driving towards incorporating artificial intelligence technology into businesses and products to streamline operations, analyse user behaviour, and predict potential consumer behaviour (Berkowitz, 2020). Microsoft SA (2019) found that at least 46 per cent of SA companies are actively piloting AI technology within their organisations. Over the past ten years, SA companies have invested more than R24 billion in AI. According to the report, the bulk of this investment went towards IoT and social media, followed by planning, scheduling and optimisation, as well as smart mobiles. The research found that businesses are experimenting with a range of different technologies, including chatbots, robotic process automation and advanced analytics. About 67 per cent of SA businesses pinpointed machine learning as the most useful AI technology, followed by smart robotics and biometrics (Microsoft SA, 2019).

The insurance and banking industry in SA is increasingly using AI technologies to improve customer experience by creating insurance and banking models to address customer needs in order to better predict and analyse customer behaviour. The industry also uses these AI

technologies to source data from other online platforms linked to a customer, i.e., social media accounts, to build customer risk profiles and exposure levels for short and long-term insurance and to better understand customer interests in products and services (Berkowitz, 2020).

South Africa (SA) has also witnessed many emergent start-ups which make use of different AI technologies, such as machine learning. For example, Aerobotics provides AI-enabled pest detection, disease detection, drone imagery services, and orchard and yield management to farmers. Jumo AI-powered technology is a stack start-up that connects banks with traditionally inaccessible customers and enables banks to deploy loans, savings, and insurance services. Another example is InsurTech, which is a start-up that uses AI with image recognition to analyse photographs of items that end-users wish to insure and have submitted via the app. The app can identify the item from the photograph and offers the end-user insurance for the item (Berkowitz, 2020). This is a clear sign that SA businesses have generally embraced new technologies and AI technology in particular. However, the ever-changing nature of the AI technological landscape requires endless resources to regulate and ensure the responsible governance of these technologies unless the operator and/or creator remains responsible and accountable (Mulamula, Lushaba, 2020).

The SA government finalised the ICT and Digital Economy Masterplan for South Africa in 2021. The master plan includes, amongst other things, action to regulate the ICT environment in the next five years. From a governance and compliance perspective, SA has introduced the Protection of Personal Information Act, 2013, also known as POPIA, which was enacted on 1 July 2020. Thus, corporations need to ensure that AI technologies or systems comply with the POPIA Act, especially section 71(1), which governs automated decision-making. The section protects data subjects from being subjected to a decision which is based solely on automated decision-making. For example, if the AI technology system has the ability to profile clients seeking a life cover or car insurance and determine their risk profile based on previous accidents, payments, income, indebtedness, etc., Section 71(1) prohibits the insurance companies from making a decision to grant or reject the insurance application based solely on the profile created by the AI system (Mostert & Tembedza, 2020). However, sections 71(2) & 71(3) of POPI provide for an exception to the above preclusion and include an explanation concerning whether such a decision has been taken. This means the insurance company which is receiving the application would

need to determine whether it can rely on one of the exceptions to the prohibition on automated decision-making, which are set out in section 71(2) of POPIA.

The POPIA act alone is not sufficient to address the legal and ethical concerns arising from AI technologies. For example, the POPI act does not regulate the biases such as gender or race inherited in AI technologies. Thus, there is a greater need to review the existing laws and factor in the accountability of key players in the unfolding machine and technology-driven environment that is facilitated by AI technologies. The current legal framework may not adequately cater to the responsibilities and accountability issues that arise from AI. Corporate governance needs to develop best practices and principles to guide boards to govern ethical and other risks that arise from the evolving new world of AI (Mulamula, Lushaba, 2020).

2.7. CONCLUSION

In this chapter, the researcher explored different definitions from the early figures of AI, such as Alan Turing, John McCarthy, Marvin Minsky, and Simon Newell, who laid the foundation to recent scholars. The chapter detailed the history of AI from the 1950s, transitioned from theory into practice, the rise of AI in the 60s, the first AI winter in the 70s, the age of renewed optimism in the 80s, the expansion of AI applications such as deep learning architectures in 2000s and explosive grown to date which include wide adoption of deep learning techniques and generative pre-trained transformer models. The chapter further clarifies key concepts used interchangeably, such as AI technologies such as machine learning, deep learning, and generative AI, which are layers of data-driven AI technologies. Furthermore, the chapter highlighted the importance of AI in health, education, business, environment and agriculture while acknowledging the risks and ethical concerns. Lastly, the chapter discussed the global spending on AI and South Africa's incorporation of AI technologies into business operations and product development.

CHAPTER 3: CORPORATE GOVERNANCE AND ARTIFICIAL INTELLIGENCE

3.1. INTRODUCTION

This chapter discusses in detail different corporate governance models within the governance literature and context. The chapter details the theoretical framework of corporate governance and provides insights into the role of the boards (board of directors). The chapter also outlines the conceptual framework for the role of the board in governing artificial intelligence ethics. The main rationale is to paint the broader foundational understanding of the role and responsibilities of the board in implementing effective oversight mechanisms, strategic decision-making, and ethical considerations to ensure responsible development, deployment and adoption of AI technologies.

3.2. CORPORATE GOVERNANCE

3.2.1. Corporate governance definition

Since the dawn of the millennium, numerous corporate governance scandals have dominated the headlines, from Enron (2001)[11], Lehman Brothers (2008)[12], Volkswagen (2015)[13], Steinhoff (2017)[14], KPMG South Africa (2017)[15], Tongaat Hulett (2018)[16], VBS Bank (2018)[17], amongst others. As a result, corporate governance has become critical over the last decade, both in South Africa and internationally (Malherbe & Segal, 2001). The series of corporate governance failures has reignited the debate regarding the board's role in directing and controlling these corporations. Its importance arises in modern corporations due to the separation of management and ownership control wherein the shareholder's interest conflicts with the interest of the board and management in the organisation.

---

[11] Corporate Governance Failure: The Case of Enron and Parmalat https://core.ac.uk/download/pdf/236418354.pdf
[12] The Failure of Lehman Brothers: Causes, Preventive Measures and Recommendations https://core.ac.uk/download/pdf/234629813.pdf
[13] Volkswagen's Diesel Emissions Scandal https://www.researchgate.net/publication/308021860_Volkswagen%27s_Diesel_Emissions_Scandal
[14] Inside the Steinhoff saga, one of the biggest cases of corporate fraud in South African business history https://www.cnbcafrica.com/2018/steinhoff-rise-fall/
[15] KPMG rocked by South African corruption scandal https://www.accountancyage.com/2017/09/27/kpmg-rocked-south-african-corruption-scandal/
[16] The case against the Tongaat executives https://mg.co.za/news/2022-02-11-former-tongaat-hulett-bosses-in-court-for-fraud/#:~:text=Straude%20and%20his%20co%2Daccused,which%20appointed%20a%20forensic%20investigation.
[17] The Great Bank Heist https://www.corruptionwatch.org.za/wp-content/uploads/2018/10/VBS-Mutual-Bank-The-Great-Bank-Heist.pdf

The purpose of this section is to highlight the debate on the definition of corporate governance. There is no universally accepted definition of corporate governance (Rwegasira, 2000; Mallin, 2007; Solomon, 2007). Some definitions have been criticised for being either too narrow or too broad. According to Ntim (2009), a definition is considered to be broad if it prioritises the protection of all stakeholders without specifying their identities, while it is considered to be narrow if it prioritises the maximisation of shareholders' interests while disregarding or neglecting the interests of other stakeholders.

The term corporate governance has developed as an independent field of study (Keasey et al., 1997; Denis, 2001), and its scope has been integrated into different disciplines, i.e., economics, ethics, finance, law, management, politics, etc. As a result, there are many definitions (e.g., Cadbury Report, 1992; Shleifer and Vishny, 1997; Zingales, 1998; Denis and McConnell, 2003; OECD, 2004; Solomon, 2007; L'Huillier, 2014; amongst others). According to the agency theorists, 'corporate governance' is deemed a systemic provision of measures to control and manage agents' actions, such as the board and management (L'Huillier, 2014). Garvey and Swan (1994) assert that "governance determines how the firm's top decision-makers (executives) actually administer such contracts" (p. 139). Shleifer and Vishny (1997) define corporate governance as "the ways in which suppliers of finance to corporations assure themselves of getting a return on their investment" (p.737).

The famous definition of corporate governance (Report on Financial Aspects of Corporate Governance in the United Kingdom, 1992), namely that "[c]orporate governance is the system by which companies are directed and controlled" (p.15), was established by Rayman-Bacchus (2003). They argued that corporate governance is primarily the process in which corporations are controlled, supervised and managed. Similarly, the concept has been defined as a system whereby directors are entrusted with responsibilities and duties in relation to the direction of a company's affairs (Sheikh and Chatterjee, 1995), or ways of ensuring that corporate actions, agents and assets are directed at achieving the corporate objective established by the corporation's shareholders (Sternberg, 2004). This implies that corporate governance is concerned with processes, structures or mechanisms through which the board and executive management are held accountable (Bell, Filatotchev and Aguilera, 2014). This definitions highlights the need for good management and good control without getting into the dilemma of being pro maximisation of shareholders' interests or pro protection of stakeholders.

The OECD (2004) defines corporate governance as;

"a set of relationships between a company's management, board, shareholders and other stakeholders. Corporate governance also provides the structure through which the company's objectives are set, and the means of attaining those objectives and monitoring performance are determined" (p.1-66).

These definitions by the OECD (2004), Thomas (2012) and Oman (2001) do not only focus on the internal structure of the company, but also appreciate and recognise the other material stakeholders of the corporation. Typically, and as has been explained above, internal corporate governance structures may include shareholders, the board of directors, and the executive management. By contrast, the "other stakeholders" may consist of those who have a material interest in corporations, such as government regulators, local communities, political, social and economic policy-makers, and institutions within which corporations operate. In this case, the corporation would be regarded as a corporate social citizen accountable to different or multiple stakeholders (Freeman and Reed, 1983; West, 2006; Mallin, 2007), rather than to shareholders only.

It is evident from the above that there is considerable disagreement among researchers and institutions regarding the definition of corporate governance. This diversity of opinion is largely due to the expanding scope of the subject of Corporate Governance itself and the differing views on whether executive management is responsible for maximising owners' equity or for also protecting and maximising the interests of other stakeholders. However, Freeman & Reed (1983) and McVea (2001) contend that firms have a responsibility to all stakeholders, not just shareholders. Thus, McVea's (2001) argument emphasises the need to align companies' objectives with those of shareholders in order to balance shareholders' wealth maximisation with their responsibility to other stakeholders. Therefore, according to this view, the definition of corporate governance should incorporate both shareholders' wealth maximisation and the firm's responsibility to other stakeholders.

The King IV report on corporate governance has elevated the definition by focusing its version of it not only on "directing and controlling" the company, but also the stakeholder-inclusive approach, ethical leadership, transparency, accountability, fairness and responsibility to achieve ethical culture, good performance, value creation, legitimacy and effective control of the organisation (King IV Report, 2017). Therefore, this research adopts

the corporate governance definition provided in the King IV report, which focuses on the principles of good governance and the creation of sustainable value for a wide range of stakeholders, including society as a whole. This perspective is consistent with stakeholder theory, a philosophical approach that highlights the interconnectedness of various stakeholders involved in or affected by a corporation's operations (Freeman & Reed, 1983; and McVea, 2001). The study recognises the board role in balancing the diverse interests of stakeholders, ranging from shareholders and employees to customers, suppliers, and the wider community. This inclusive approach by the board is crucial for achieving long-term success and ethical business practices, ensuring that all material stakeholders are taken into account in strategic decision-making processes.

3.2.2. The theoretical framework of corporate governance

3.2.2.1. Corporate governance and the agency theory

Corporate governance is located within the theoretical frameworks of agency theory. According to Jensen & Meckling (1976) the agency relationship is defined as "one in which one or more persons (the principal) engage another person (the agent) to perform some service on their behalf which involves delegating some decision-making authority to the agent". Corporate governance has been widely grounded on the agency theory perspective, which is concerned with creating ways to motivate the agent to act on behalf of the principal. The theoretical framework of agency theory can be located back to Berle & Means (1932), who describe the agency problem due to the "separation of ownership and control" of corporations. The separation of powers in corporations at the time prompted a debate to effectively align the interest of the principal (shareholder) and the agent (board of directors & management). Thus, agency theory is defined as relating to the relationship between the principals, such as shareholders, and agents such as the company executives and managers, which implies the principal or shareholders delegate the corporation's running to the directors (Clarke, 2004).

According to agency theory then, the separation between shareholders and those who control the company (board of directors) gives rise to agency problems since the shareholders are not in a position to directly oversee boards in order to ensure that they are making decisions based on the shareholders' interests (Ugowe, 2016). From the agency theory perspective, managers are expected to act and make decisions in the shareholder's interest. Thus, the board exercises oversight, supervision, and monitoring to ensure that

management performs in the shareholder's best interest (Hillman & Dalziel, 2003). However, in reality, the board may make decisions contrary to the shareholder's best interest (Padilla, 2000). In terms of Section 75 and 76 of the Companies Act of 2008, the board of directors is expected to act in the corporation's best interest, not of the shareholders. This then creates what is known as the agency/principal problem.

The theory is based on the premised presumption that humans are rationally bounded, self-interested and opportunistic, resulting in differences between the objectives of the shareholder and board-management interest. The agency problem arises because the agent (managers) has more information about the corporation's operations than the principal (shareholder). The separation of ownership has allowed managers to gain more knowledge and expertise about the corporation's operation over the shareholders who are absent from the day-to-day operation (Madhani, 2017). Because of this information asymmetry, the shareholder cannot always ensure that the managers perform in accordance with the shareholder compact. Kamalnath (2019) argues that the agent has an incentive to not perform at the expected standard or to divert to him or herself what was promised to the principal.

One of the main purposes of corporate governance (CG) is establishing governance and monitoring mechanisms to resolve these differences in order to protect the interests of shareholders (Jensen and Meckling, 1976) and to maximise shareholder wealth by reducing agency loss (Adegbite et al., 2012; Bonazzi and Sardar, 2007; Cohen et al., 2002). According to Madhani (2017), CG acts as a mechanism wherein the board (board of directors) is a monitoring tool to reduce/deal with the principal-agent relationship. In this context, agents are the managers, and principals are the owners, and the boards act as the monitoring mechanism (Mallin, 2004). As a result of the board monitoring, managers' opportunistic pursuit of self-interest will be minimised, thus lowering the agency costs, which may result in greater returns or increased profits for shareholders.

The limitation of agency theory is the conjecture that CG is more about reward and consequence management, which need to be put in place to encourage the board-management to treat corporate governance principles as a simple checklist of recommended practices (da Silveira, 2015). As a result, a complex subject of corporate governance has been reduced to reward and control mechanisms to motivate board management to make decisions in the principal's best interest (shareholder). This theoretical

basis and its basic principles have influenced conceptualisation models and implementation techniques of corporate governance and, by extension, that of information and technology governance in companies.

The agency theory critics have argued that it has a narrow focus on maximising shareholder value and responsibility, while neglecting the rights of other stakeholders in the corporation such as employees, suppliers, government, communities, etc. (Donaldson and Preston, 1995). They argue for a non-economic based theory, the stakeholder theory, which is "conceptualised as the system for arbitration between a range of claimants and the processes associated with joint decision-making" (Zeitoun and Osterloh 2012). The stakeholder theory is mainly about the interconnected relationships between the corporation and those who have a legitimate stake in the organisation, i.e., customers, suppliers, employees, investors, communities, the society, the environment, etc. (Wagner et al., 2012; Fontaine et al., 2006). The theory argues that corporations should balance the interests and create value for all legitimate stakeholders, not just shareholders (Fontaine et al., 2006).

On the other hand, the stakeholder-inclusive approach has gained traction in the corporate world in recent years. It acknowledges that the corporation's best interest is not necessarily the best interest of the shareholder and that shareholders do not necessarily have first preference over all other stakeholders. The shareholder approach, on the other hand, means that the board does not consider other stakeholders as an instrument to serve the shareholder's interest, but rather that they have intrinsic value for decision-making in the best interest of the corporation. Thus, according to the stakeholder approach the board should identify material stakeholders and their interests and expectations relevant to the corporation's strategic objectives and long-term sustainability.

Accordingly, the relationship with stakeholders should be based on normative and ethical responsibility. Berman et al. (1999) argue that corporations should establish a fundamental ethical principle (or set of principles) to guide stakeholders' behaviour and decision-making. Thus, it supports the idea that CG should not only be a simple checklist of recommended practices, but rather there should be morally principled (reasoning and) decision-making by the boards. Incorporating stakeholder theory into the corporation operating in the fourth industrial revolution and, by extension, AI environment, would not only result in collaboration amongst various stakeholders during the planning, design, development, and implementation of AI technologies, but may also help address global challenges related to

AI ethics. The study adoption of a stakeholder-inclusive approach is grounded in agency theory, which primarily examines the relationship between principals and agents. Traditionally, agency theory has faced criticism for its narrow focus on shareholder primacy, but by integrating stakeholder inclusivity, this approach broadens the scope of governance to encompass the interests of all stakeholders, thereby mitigating agency conflicts. This is achieved by fostering a governance framework that promotes transparency, accountability, and ethical conduct. Research conducted by Mhlanga and Molio (2020) showed that corporations that adopt a stakeholder approach are more effective in serving society as a whole.

Other theories developed after the agency theory include stewardship theory, resource dependence theory, and resource-based view theory. Stewardship theory regards executive managers as inherently trustworthy individuals (Nicholson and Kiel, 2003). From this perspective managers should be given full authority to run corporations as they are good stewards of the resources entrusted to them (Letza et al., 2004). The theory assumes that since top managers usually spend their entire working lives in the corporations they govern, they are more likely to understand the business than outside directors (Donaldson and Davis, 1991; Mtim, 2009). Secondly, executive managers possess exclusive formal and informal information and knowledge about the corporations they manage, and this can aid better decision-making (Donaldson and Davis, 1994) as compared to having non-executive members making decisions. As a result, advocates of stewardship theory contend that better financial performance is likely to be associated with internal corporate governance practices that grant managers greater powers, such as combining the positions of company chairman and CEO (Donaldson and Davis, 1991, 1994; Mtim, 2009).

Resource dependence theory suggests that the institutionalised internal governance structures, such as the board, are necessary for ensuring effectively monitored executives and serve as an essential link between the corporation and the critical resources that it needs in order to maximise the profits (Pfeffer, 1973). Resource dependence theory indicates that internal corporate governance structures, like the board of directors, help to link the firm to critical business inputs needed for higher financial performance. From the theory's perspective, the board and non-executive directors, in particular, offer critical resources such as expert advice, experience, independence, knowledge (Haniffa and Cooke, 2002), reputation and critical business contacts (Haniffa and Hudaib, 2006). Nicholson and Kiel (2003) argue that the board can facilitate access to business/political

elite, information, and capital. Finally, the board provides a critical link to corporations' external environment and significant stakeholders, such as creditors, suppliers, customers, and competitors. As a result, it has been argued that a greater level of links to the external environment is associated with better access to resources (Nicholson and Kiel, 2003). This can impact positively on the firm's financial performance.

Although stewardship theory, resource dependence theory, and the resource-based view theory all make significant contributions to the field of corporate governance, this study explicitly favours agency theory and stakeholder theory. This decision was made because these theories provide more robust frameworks for examining internal governance mechanisms, conflicts of interest, and the ethical considerations and inclusive stakeholder approach in decision-making processes.

3.3.   THE BOARD (BOARD OF DIRECTORS)

3.3.1. The Board (Board of Directors)

Around the early seventies, Mace (1971) provided insights into the role of boards (board of directors). Since then, a wide range of roles boards play in the corporation has been placed at the centre of the corporate governance debate (Johnson, Daily, and Ellstrand,1996); Zahra & Pearce, 1989). This has also been exacerbated by a high number of corporate governance scandals and the financial crisis. The call for a special focus on the role of the board and their accountability toward a diverse range of stakeholders has become more pronounced, reinitiating the debate on which roles boards should perform (Pugliese, Minichilli, & Zattoni, 2014). These advancements in corporate governance (CG) suggest that the board (board of directors) is one of the most important governance structures of a corporation (Fama and Jensen, 1983; Williamson, 1984; Madhani, 2017).

According to agency theory, the board is a monitoring tool to reduce/deal with the principal-agent relationship. As such, boards have an oversight role to control, oversee strategy implementation, and monitor decisions of the executive management in the corporation (Harrison, 1987; Johnson et al.,1996; and Withers and Hillman, 2008). Hillman and Dalziel (2003) further submit that boards serve as a critical function for providing valuable resources to the corporation, such as advice, expertise, and legitimacy.

Over the last several decades, however, oversight, supervision, and monitoring have become the core functions of boards instead of managing and controlling, as boards delegate the role of running the corporation to the executive management. Boards are part-time, intermittent decision-making structures. This implies that their members only meet periodically, in most instances, four times per year, and the majority of the board members are non-executive members who are not employees of the corporation. Petrin (2019) argued that in practice, this setup makes it impossible for boards to comprehensively manage and control a corporation on a daily basis. To ensure oversight, supervision, and monitoring of executive management in a corporation, there are two types of board systems: a one-tier board and a two-tier board.

3.3.1.1. One-Tier Board Model[18]
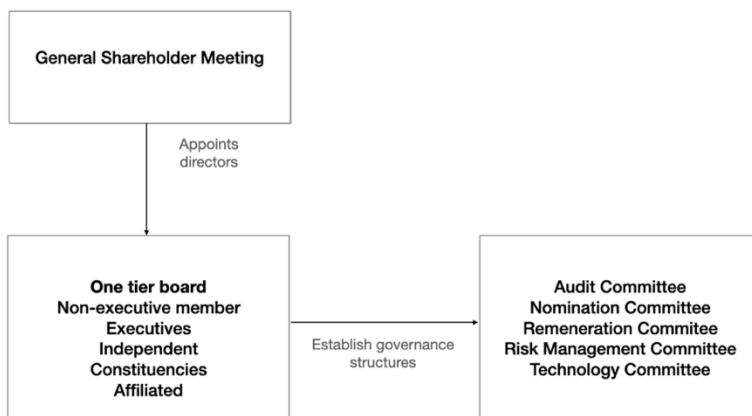
Figure 4: One tier board model

In a one-tier model, the board of directors is composed of executive directors/management (e.g., C-Suites Executives such as Chief Executive Officer, Chief Financial Officer), who are engaged in the daily operations of the corporations, and an independent non-executive director, who are expected to fulfil an oversight and monitoring role within the board (du Plessis et al., 2017). The one-tier board model is widely adopted in European countries, the USA, the United Kingdom and South Africa.

For South Africa, the Companies Act provides for a one-tier board structure. Section 66(1) of the Companies Act provides that the business and affairs of a company must be managed by or under the direction of its board, which has the authority to exercise all the powers and

---

[18] For the purposes of this study, the focus will be on a one-tier board model used by South Africa and adopted by the JSE listed companies.

perform any functions of the company, except to the extent that the Act or the company's Memorandum of Incorporation provides otherwise. Accordingly, the board is a focal point of governance in the corporation. As depicted in Figure 4 above, it gives strategic direction to the company. Muswaka (2014) argues that while the board may reserve specific powers to it and delegate some authority or powers to the executive management, any such delegations by the board must have due regard for the director's statutory and common law duties to the company, while taking into account strategic and operational effectiveness and efficiencies.

In the one-tier board model, the board is mainly composed of the independent non-executive director and executive directors, and are appointed by the shareholder to work together with executive management for the long-term sustainable value of the corporation (Idowu, 2013). Both the non-executive and executive members are part of the board subcommittees, such as the Audit Committee, Technology Governance Committee, Remuneration Committee, etc.

The one-tier model is an effective corporate governance board structure that makes decision-making faster at the strategic level. In a single-tiered board, a clear division of roles and responsibilities of the non-executive and executive directors are defined, resulting in effective leadership (Khoza, 2015). This is simply because non-executive and executive directors are part of the same board; thus, no separate approvals are required. The one-tier board allows for more integration of the business strategy, and decision-making is witnessed between the executive and non-executive directors (Block, 2016). Thus, the board of directors functions as a collectively appointed corporate body. However, the one-tier model is usually criticised for its propensity to compromise the independence of the non-executive director. Thus, this may result in diluting their independence and oversight role.

3.3.1.2. *Two-Tier Board Model*

On the other hand, a two-tier system is a two separate board structure. It is composed of the executive management board responsible for day-to-day operations and the supervisory board responsible for oversight and monitoring the executive management board. The supervisory board is also responsible for the appointment and removal of management board members (du Plessis et al., 2017).
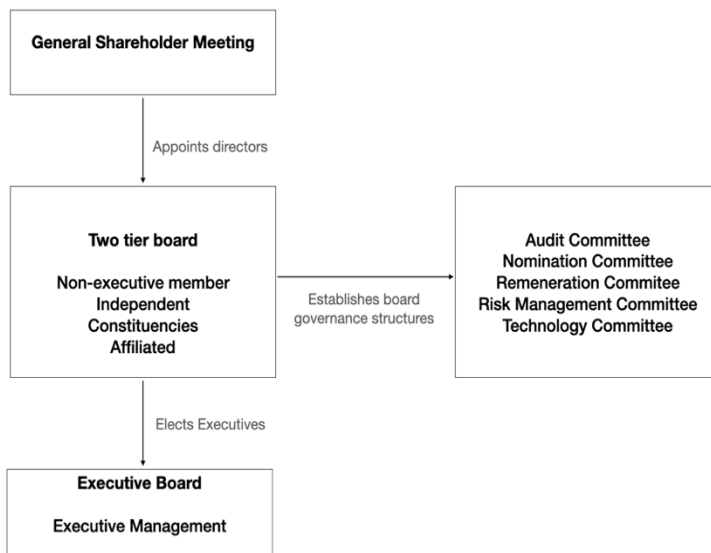
Figure 5: Two tier board model

The countries such as Germany, China and Norway use this board model. Both the executives and the non-executive have separate boards, and their mandates are kept separately, which means that members of the supervisory board cannot be the management board members at the same time (Goergen, 2012; Idowu, 2013). The model recognises the board's role as setting the strategy, and overseeing and monitoring the executives in a corporation (Khoza, 2005). Those who favour the two-tier approach argue that the model provides a clear mandate of responsibilities, as the non-executive is responsible for oversight while management is concerned with strategy implementation. They further argue that the supervisory board acts as an arbitrator between the executives and the shareholders. The two-tier system board is composed of stakeholder representatives such as employees, who are a part of the decision-making processes of the corporation (Corporate Strategy Board, 1998).

Moreover, members of the supervisory board are recommended for appointment by the management board and only formally elected in the annual general meeting. This raises the question of "who" gets to be pre-selected or recommended by the management. Du Plessis et al. (2017) argues that the approach of pre-selection by the management may result in "only persons that are 'adequate' in the opinion of the management board becoming elected as members of the supervisory board, and only those who 'adequately' control in the eyes of the management board remain in office".

3.3.2. Role of the board

The conceptualisation of the board's role has become complex over the years mainly because of the ongoing evolution of the board legal duties and societal expectations (Baxt, 2005). The early attempts by Eisenberg (1969) and Conard (1976) at defining board roles point out the differential conceptualisation of boards by different disciplines, and often by scholars within the same discipline. Pfeffer and Salancik (1978) identified three key roles: as a co-optive mechanism accessing key resources, as boundary spanners, and as enhancing corporation legitimacy (controlling or overseeing management).

In contrast, scholars have generally agued four key roles and responsibilities of the board: (1) the control role (Mizruchi 1983; Baysinger & Hoskisson, 1990; Stiles & Taylor, 2001); (2) the strategic role (Ingley & Van der Walt, 2001; Schmidt & Brauer, 2006); (3) the stakeholder role (Arena, Bozzolan, & Michelon, 2015; Ayuso, & Argandoña, 2009); (4) the advisory and counsel role (Nicholson and Kiel, 2004; and Bonn and Pettigrew, 2009) and (5) resource provision role. The control role of the board suggests the fiduciary duty of monitoring and supervising the operation and performance of the corporation. The board's strategic role relates to setting and steering strategy and ensuring the realisation of the corporation's vision and mission. The stakeholder role is primarily concerned with ensuring good relations among material stakeholders of the firm, i.e., shareholders, suppliers, government agencies, creditors, employees, etc. Lastly, the resource provision role of the board focuses on the set of resources that each member brings to the board, such as the use of knowledge, skills, and experience (Madhani, 2017). Resource provision includes approval of annual budget and ensuring that the corporation has sufficient financial resources to meet its current and future needs. This may involve raising additional funds through activities such as issuing debt or equity, or by securing grants or other forms of funding.

In South Africa, the role of the board of directors has been codified under the Companies Act 71 of 2008. Prior to the introduction of the companies act, the role and responsibility of the directors were overseen under the SA Common Law. However, it must always be kept in mind that in the event that the common law duty has not been codified, the duty still exists and is upheld by the courts, which is confirmed in the Act. Therefore, one should not assume the common law duties have fallen away in place of the codification of directors' duties in the Act.

The Companies Act extends the role and responsibility of the directors and increases accountability. The Companies Act, 71 of 2008 states that the board has a fiduciary duty to "act in good faith and the company's best interest" by exercising "care, skill and diligence" to promote the company output through independent judgment. Section 66 of the Act provides (as stated above already) that "the business and affairs of a company must be managed by or under the direction of its board, which has the authority to exercise all of the powers and perform any of the functions of the company, except to the extent that the Act or the company's Memorandum of Incorporation ('MOI') provides otherwise" (p.132). Therefore, it is clear that both the Act and the MOI can curtail the powers, and thus the duties, of the board of directors. The Act directs the role of the directors as follows: A director:

- Should act in the best interests of the corporation.
- Should not act beyond or exceed the limitations of powers or capacity, which means a director should always act within the ambit of their authority.
- Should act within their powers and for a proper purpose. This means that they should act in the company's best interests and in a manner that benefits the company as a whole and bona fides towards the company interests.
- Should exercise unfettered and independent discretion and judgement. This requires that a director exercise independent and unbiased judgement when reaching and making decisions for the company.
- Should not be in a position or should avoid a position of conflict of interests. Their personal interests should not conflict with those of the company, and this also includes the duty to disclose any potential conflicts of interest.
- Should not make any secret profits or possible incidental profits at the expense of the company.
- Should disclose any personal interests in any contracts with the company.
- Should not misappropriate corporate opportunities due to the company as well as improperly competing with the company.
- Is to display and demonstrate the same care as a reasonable person in their personal affairs as well as the degree of skill which one would reasonably expect from a reasonable person with his/her knowledge and experience.

Apart from the core legal duties, most jurisdictions also have additional layers of regulations and "best practices" that company boards need to follow. In South Africa, the King IV Report

on Corporate Governance for South Africa – 2016 ("King IV") is a set of voluntary principles in the area of corporate governance. Companies listed on the Johannesburg Stock Exchange are required to comply with King IV by law. The King IV Report (2017) states that; "the board should serve as the focal point and custodian of corporate governance in the organisation". This implies that the board's role is to steer and set strategic directions, provide ethics leadership, monitor performance, approve policies, ensure accountability, and direct the management in the corporation's best interest (Mulamula, 2018). The report details the role of the board, under 17 principles, as follows:

- Principle 1: The board should lead ethically and effectively.[19]
- Principle 2: The board should govern the organisation's ethics in a way that supports the establishment of an ethical culture.
- Principle 3: The board should ensure that the organisation is and is seen to be a responsible corporate citizen.
- Principle 4: The board should appreciate that the organisation's core purpose, risks and opportunities, strategy, business model, performance and sustainable development are all inseparable elements of the value creation process.
- Principle 5: The board should ensure that reports issued by the organisation enable stakeholders to assess the organisation's performance and its short, medium, and long-term prospects.
- Principle 6: The board should serve as the focal point and custodian of corporate governance in the organisation.
- Principle 7: The board should comprise the appropriate balance of knowledge, skills, experience, diversity, and independence for it to discharge its governance role and responsibilities objectively and effectively.
- Principle 8: The board should ensure that its arrangements for delegation within its structures promote independent judgement and assist with the balance of power and the effective discharge of its duties.
- Principle 9: The board should ensure that the evaluation of its performance and its committees, chair, and members support continued improvement in its performance and effectiveness.

---

[19] The role of the board as per the King IV principle 1 to 3 above will be discussed in detail in Chapter 4 of the study.

- Principle 10: The board should ensure that the appointment of, and delegation to, management contribute to role clarity and the effective exercise of authority and responsibilities.

- Principle 11: The board should govern risk to support the organisation in setting and achieving its strategic objectives.

- Principle 12: The board should govern technology and information to support the organisation setting and achieving its strategic objectives[20].

- Principle 13: The board should govern compliance with applicable laws and adopt non-binding rules, codes, and standards to support the organisation being ethical and a good corporate citizen.

- Principle 14: The board should ensure that the organisation remunerates fairly, responsibly, and transparently to promote the achievement of strategic objectives and positive outcomes in the short, medium, and long term.

- Principle 15: The board should ensure that assurance services and functions enable an effective control environment. These support the integrity of information for internal decision-making and the organisation's external reports.

- Principle 16: In the execution of its governance role and responsibilities, the board should adopt a stakeholder-inclusive approach that balances the needs, interests, and expectations of material stakeholders in the organisation's best interests over time.

- Principle 17: The board of an institutional investor organisation should ensure that the organisation practices responsible investment to promote good governance and value creation by the companies in which it invests (King IV Report, 2017).

According to the G20/OECD Principles of Corporate Governance:

- The board provides the structure through which the objectives of the company are set and the means of attaining those objectives and monitoring performance are determined.

- The board has a key role in setting the ethical tone of a company, not only by its own actions, but also in appointing and overseeing key executives and consequently the management in general.

---

[20] The role of the board in technology and AI governance is expanded more in chapter 5 of the research paper.

- The board should fulfil certain key functions, including reviewing and guiding corporate strategy, major plans of action, [and] risk management policies and procedures.

The OECD (2005) also argues similar roles to those of the King IV Report above. As the focal point of corporate governance in corporations, the board is by inference responsible for responding to ethical and risk issues that AI technologies generate. As AI-driven technologies improve and become more and more embedded in the business processes, there will be increased pressure on the board to make complex decisions about AI ethics issues as well as legal and financial implications. Brand (2021) notes that the need for boards to pay more attention to AI technologies has begun to receive more coverage from scholars, professionals, and the press, pointing to the significance of AI in society (Brand, 2021). This has resulted from AI's ability to make independent decisions (e.g., in terms of allocation of resources) that impact individuals in society as a whole. Good ethical governance by the board is essential to building trustworthy AI technologies in society (Langford, 2019). Poor governance of AI raises the risk of eroding trust and incurring reputation risks and liability of boards and corporations (Brand, 2021).

3.3.2.1. The fiduciary duty of the board

The word 'fiduciary' is associated with a relationship of trust "where one party has an implicit dependency upon and peculiar vulnerability to another within defined parameters" (Rotman, 2017). Fiduciary duty means to provide the dependent with ways to protect their trust. Common Law dictates that directors owe to shareholders two basic fiduciary duties: the duty of loyalty and the duty of care (Kamalnath, 2019). The duty of loyalty is concerned with conflict of interest. In essence, the aim is to ensure that the board acts in the best interests of the corporation and its shareholders, rather than in their own interest (Lafferty, Schmidt, & Wolfe Jr, 2011).

The duty of loyalty is codified in the SA companies Act, which dictates that directors must act in the corporation's best interest. The duty discourages the directors from engaging in a business transaction where they have a conflict of interest. Failure to prevent, mitigate, manage, or respond to an incident may amount to a breach of directors' duties under the Companies Act. Conflicted transactions are not strictly prohibited; instead, they are

regulated; thus, another crucial fiduciary responsibility is that of disclosure of actual or preserved conflict of interests by the directors.

The second fiduciary duty in terms of Common Law is the duty of care, and it requires that the directors be attentive to details when making decisions on behalf of the corporation (Veasey, 1997). The duty requires directors to act with care, skills and diligence when performing their responsibilities. In other words, directors are required to take reasonable steps to inform themselves before making business decisions. The business judgment rule is related to the duty of care, which provides that directors are not liable for losses arising from mere errors of judgment if they acted in good faith. The purpose of this protection from liability is that business decisions are often made in a specific context, and it is not proper for courts to, in hindsight, "substitute their judgment for that of the directors" (Kamalnath, 2019).

The board's overall fiduciary responsibility is to promote the value of the corporation. In fulfilling that duty, directors must exercise their business judgment in considering and reconciling the interests of various stakeholders, i.e., shareholders, employees, customers, suppliers, the environment and communities, and the attendant risks and opportunities for the corporation (Mickels, 2009). Accordingly, a breach of fiduciary duty may lead to the board of directors' liability. Positioning boards to effectively exercise their fiduciary responsibilities for AI technologies innovation is a most important governance imperative that demands boards' immediate and ongoing attention (Broccolo & Peregrine, 2021). The fiduciary board duties do not change when the corporation plans, designs, and implements AI technologies. In this context, boards should still be guided by their basic duties of care and loyalty.

However, AI poses emerging risks and opportunities for the board relating to legal and fiduciary duties (Kaplan and Haenlein 2020). Boards have a responsibility to take reasonable steps to gain deeper insights into the impact of AI technology in the corporation. Exercising care, skill, and diligence means that the board, individually and collectively, need to constantly obtain and uphold adequate information and understanding to execute their oversight and control mandate effectively. They should look at new approaches for risk oversight which are now increasing with the use of AI technologies. This should include how the company's enterprise risk management system informs that oversight, revisiting the governance model and skillsets to focus on their risk oversight efforts. There is no argument that AI technology advancements are transforming businesses to operate more effectively

and efficiently. However, despite these advancements, ethical issues related to AI technology are less understood, especially at the board level (Wright and Schultz 2018).

Verbeke and Hutzschenreuter (2020) argue that boards have not previously had to deal with new risks, including the need to balance AI technology risks with those not acting fast enough to reap the benefits for the corporation in the extractives sector. The board needs to implement effective oversight of AI technologies. This implies that the board should be knowledgeable enough about AI technologies and has a point of view on a critical issue such as the ethics of AI. The board should keep abreast of the impacts (and potential) of AI technologies to avoid shocks to the corporation and to be open to the impacts that may reasonably be expected to arise during the course of executing their duties as directors. Therefore, the board should ensure that it effectively manages potential AI technology and corporation risks, including ethical, programming, and reputational risks. It is thus part of a board's fiduciary responsibility to oversee AI technologies and their potential impact on the corporation. Even if AI technology is being used in boardrooms themselves, it does not replace the board's responsibility to exercise due care and loyalty, to be fully informed and make its own decisions in the best interests of shareholders and stakeholders.

3.3.2.2.   *The strategic role of the board*

Over the years, the traditional forms of corporate governance have been challenged, and boards are now required to develop a broader mindset to deal with the uncertainty of higher-level issues such as direction-giving and implementation of the strategy. The control and advisory role of the board has been broadly researched, but the studies on the strategic role have mainly focused on overseeing and monitoring, while largely ignoring the more active role of strategy formulation (O'Neal & Thomas, 1995). Ingley and Van de Walt (2001) contend that setting strategic direction has traditionally been the role of executive management, as "texts on strategic management commonly affirm" (Ingley & Van de Walt, 2001). However, increasingly the call for the board to take responsibility beyond just approving strategic direction set by executive management and monitoring implementation thereof, and to adopting a leadership role in setting the corporation's strategic direction and accountability for its performance (Davies, 1999; Helmer, 1996) is heard.

The board plays a key role in assisting the executives in comprehending and focusing on the risks associated with corporate strategies, the drivers of corporate performance, and the

business, determining the company's risk appetite and devoting appropriate resources to risk identification and management activities (Gregory, 2022). The strategic role of the board relates to setting and steering strategy and ensuring the realisation of the corporation's vision and mission (Ingley & Van der Walt, 2001; Schmidt & Brauer, 2006). The board has a responsibility to steer the strategic direction in line with its core purpose and values by approving and monitoring short, medium, and long-term strategy while taking care of risks and opportunities (King IV, 2016). Cossin and Metayer (2015) argue that the particular role a board plays shapes its participation in the strategy debate in several distinct ways that each has advantages and disadvantages. The board's involvement in strategy also depends on the nature and complexity of the context the corporation competes in (Cossin & Metayer, 2015), for example, the board may play an advisory role in a stable market, but may need to take a more robust, hands-on approach to strategy development in a highly competitive market.

The strategic board involvement should now cover mission development, strategy formulation and strategy implementation. The strategic board involvement is viewed to be important from a practical viewpoint as it offers board reform activists a basis to delineate the sphere of directors' potential role in strategy. As it relates to artificial intelligence, the board needs to ensure that technology governance aligns with the corporate strategy (Weill and Woodham, 2002). Van Grembergen, De Haes & Guldentops (2004) argue that alignment between technology strategy and organisational strategy is crucial for technology to deliver business value. Technology strategic alignment seems to be a critical element of ensuring that technology investment adds value to the organisation. Thus, it is imperative that technology supporting business objectives is not only aligned but integrated into the strategy. One of the main reasons for the lack of return on investment is the lack of alignment between technology strategy and business strategy (Ritter, 2007). Furthermore, lack of alignment may also lead to information silos, bad business management, and failure to gain business value from technological investments.

3.3.2.3. Effective control

The control role of the board suggests the fiduciary duty of monitoring and supervising the operation and performance of the corporation. The board is the ultimate centre of control in a corporation. The control role "refers directly to the responsibility of directors to monitor managers on behalf of shareholders" (Hillman & Dalziel, 2003). This role, referred to as

'bottom-line control, includes, e.g., hiring and firing of CEOs, managing performance, monitoring budgets and operational decisions of the corporation. Muzruchi (1983) argues that the act of "hiring and firing" provides an empirical illustration of board power of control as it sets limits or parameters within which management operates. Another key responsibility of the board is to control the executives and to ensure full compliance with the law, accounting codes, and the company's statutory rules, particularly concerning the company's finances and risk management.

The agency theory argues that the separation of ownership and board control is regarded as an effective tool to safeguard the shareholder's interest. Placing the responsibility of control with the board precludes executive management from evaluating the corporation performance and their performance. Effective control gives boards high powers of inference in that directors are able to scrutinise both the financial and non-financial performance of the corporation, including the performance of executive directors. The board's evaluation and reward allocation methods are effective only to the extent that executives are responsible for performance outcomes. Therefore, through adequate board controls, executives are "protected either from the adverse personal consequences of outcomes over which they have no control or from short-term losses resulting from investments in future cash flows" (Mizruchi, 1983).

The board is responsible for ensuring that AI investments are directed and executed in a manner that achieves business value and performance. Thus, as stated, the board have a fiduciary duty to monitor and supervise operations, performance and value delivery of AI technologies in the corporation. The board needs to ensure that AI investments deliver on promised returns against the strategy, optimise costs and prove the intrinsic value in the corporation. Lunardi, Maçada, & Becker (2014) argue that delivering value requires strategic evaluation, assessment of risk and impact of AI activities on business processes, defining roles and responsibilities for delivering AI capabilities, management through an investment's economic lifecycle, and defining and monitoring performance metrics. However, the board should ensure that AI performance metrics are linked to measuring strategic alignment, value delivery, and risk management results.

Monitoring and performance measurement provides the board with information about what is currently taking place in terms of AI and value delivery in the corporation. This information also enables the board to become comprehensive in terms of its own and management's

responsibilities to keep AI aligned with the business goals (Mulamula, 2020). Thus, the concept of task programmability, discussed in agency theory, also plays a role as it enables behaviour control and crystallised goals (i.e., measurable outcomes), and outcome control (Posthumus & von Solms, 2008).

3.3.2.4. *Duty of oversight*

Silverman (2020) suggests that we can defer the debate whether artificial intelligence will eventually take over the board's functions. However, we cannot defer the discussion about how boards will oversee AI. With the rapid increase in technology use, every board needs to develop a proactive approach for overseeing how AI operates within the context of a corporation's overall mission, risk management (Silverman, 2020) and impact on society. McKinsey's 2019 global AI survey report found that although AI adoption is increasing rapidly, overseeing and mitigating its risks remain unresolved and urgent tasks. As demonstrated in the previous sections, the ultimate responsibility for AI oversight in a company resides with the board and individual directors.

Under corporate governance, the board is responsible for strategic direction and management, and it typically delegates authority for the day-to-day operations of the corporation to the executive management. Once the board has delegated broad authority, its primary responsibility is to oversee management on strategic initiatives, financial performance and the integrity of financial statements and accounting and financial reporting processes, risk management, and compliance with laws and regulations. The board oversight role is continual inquiry into whether the board has received sufficient and accurate information from management to take an informed decision (Gregory, 2022).

The board is responsible for identifying and monitoring risk and compliance on an ongoing basis, including risks presented by the adoption or development of AI technologies. This requires that the board understands the risks associated with corporate strategy and business operations, the risk management and compliance systems in place, and the information and control systems designed to bring risk and compliance issues to management's and the board's attention. Boards need to be prepared to act on risk and compliance issues as they arise and should attend to the board's own structure and processes for oversight of risk and compliance. Markel, Morduchowitz, & Catalano (2022) suggest the following oversight measures for the board of directors.

- Board members should keep themselves informed on the significance or high-risk issues to the corporation. Directors should receive regular updates from in-house or outside experts regarding risks presented by adopting or developing AI technologies.
- The board should establish reporting frameworks requiring executive management to keep the board apprised of compliance practices, incidents, red flags and AI technology risks in a timely manner.
- The board should insist that there be a reasonable system to assure updates and information flow on important issues and risks.
- The board should delegate to an existing or new committee additional oversight of major risk areas.

Boards should also evaluate whether they are appropriately structured or appoint sub-committees for risk and compliance oversight and should also periodically review the information and control systems designed to ensure that relevant information is brought to the board's attention in a timely manner.

2.2.2. Composition of the board

Most studies around board composition are centred around board member background, knowledge and expertise (Minichilli & Hansen, 2007; Minichilli, Zattoni, & Zona, 2009; Zhang et al., 2011), diversity in the board, i.e., female board representation (Post & Byron, 2015; Cumming et al., 2015), board members' dependence/independence (Gabrielsson & Winlund, 2000; Hillman & Dalziel, 2003, Avci et al., 2018; Jain and Zaman, 2020), as well as well-functioning board committees (Agrawal and Chadha, 2005; Beasley et al., 2000). To date, there have been inconclusive findings regarding the relationship between board composition and firm performance (Finegold et al., 2007; Bermig & Frick, 2010; Rashid, De Zoysa, Lodh & Rudkin, 2010). However, a positive result suggests that independent board composition can result in enhanced decision-making through increased information flows (Sanda et al., 2011).

Fama (1980) argues that board composition is an essential element in effective one-tiered boards because the non-executive directors represent a means of monitoring the actions of the executive management. In the context of South Africa, the King IV report (2016) on corporate governance stresses that the board should include a balance of executives and non-executive directors, and the majority should be independent directors. Diversity of

membership must be considered to ensure effective decision making, better utilisation of the talent pool, and enhancement of corporate reputation and investor relations by establishing the corporation as a responsible corporate citizen.

In today's fast-paced world, board directors have to ensure they are in sync with the fast-changing realities of business. In reviewing over 200 board of director compositions on the Fortune 1000, Gordon (2021) found that many of them do not have sufficient technology depth and knowledge expertise. The majority of the board still have traditional knowledge, particularly in finance and law and have often held a CEO or high profiled leadership role in a prior company. Even though most of these corporations have a technology strategy and risk committees working with their Chief Information Officers (CIO) and Cybersecurity or Risk officers, the study found that there is a lack of skill and knowledge on the depth of AI technology and digital transformation leadership and solid execution experiences of AI technologies at the board level.

The starting point to address these concerns of lack of sufficient technology depth and knowledge expertise in the boards lays with the composition of the board. Does the board have a sufficient understanding of AI and its impact on strategy and organisational culture? Few expect boards to recruit AI specialists, but all directors should be able to consider AI in the context of their corporation, its strategy, value, and ethics policy. Some scholars do however suggest that the board of directors should add a technical expert at the board level (Brennan et al., 2019).

2.2.3. Board committees

The board role ensures the establishment of structures such as a steering committee and strategy committee to support technology governance. The technology governance subcommittee should be a board-level structure that is concerned with ensuring that the board technology strategy is implemented and aligned to business objectives (De Haes and Van Grembergen, 2006). Posthumus, Von Solms & King (2010) argue that the technology committee ensures that the board receives all the information in detail to make an insightful decision on all IT related matters. To ensure that the committee advises the board appropriately on technology matters, its composition is critical. Nolan & McFarlan (2005) concur that the technology committee should be composed of a majority of non-executive directors who are experienced and skilled to understand IT's (including AI technologies) impact on the business. Although the board may delegate authority, it should not result in

the abdication of powers by the board. This means that delegating authority to board committees does not in any way mitigate or dissipate the fiduciary duties of the board. Board committees are an aid to assist the board and its members in discharging their duties and responsibilities, and boards cannot shield behind these committees.

As mentioned already, some scholars propose that governance oversight at the board level is crucial for AI deployment (Silverman, 2020; Gesser, Regner & Gressel, 2022). Thus, companies should on this view have effective board committees to oversee the deployment of innovative technologies and oversee the development and deployment of new technologies in a responsible manner, working with compliance, ethics and responsibility departments. The board committees should have in-depth knowledge and experience of AI technology. The board committee role is to support the management board when it comes to strategic decisions or even the minor ones that relate to the adoption and 'execution' of AI strategy. The board or committee should have relevant knowledge, experience and – what is quite important – shall be diverse and independent. In most cases however in reality, the board committees are mainly engaged in the typical board duties, such as monitoring and strategy, while they also play a minor role in the exploitation of AI potentialities and, above all, in the management of AI-related risks.

2.2.4.   Artificial Intelligence systems as members of the board

Corporate governance ensures that structures and processes are put in place to mitigate the agency problem with the corporation. Studies on corporate governance mechanisms for AI technologies are limited; therefore, we draw inference from existing corporate governance agency theory. One can draw inference by considering existing agency theory literature as above, which seeks to understand how the principal ensures the agent acts in the principal's best interest. Sidorova & Rafiee (2019) argue that corporate governance mechanisms such as pre-defined key performance indicators could help reduce the agency problem in one of two ways: increasing goal alignment between the principal and the agent or, reducing information asymmetry.

Moloi & Marwala (2020) posed an interesting question: What is happening to the agency theory in the era dominated by AI technologies? For example, Deep Knowledge Ventures (Hong Kong-based firm) has an AI board member called 'Vital' that 'advises' against potentially catastrophic decisions. The company uses 'Vital' as a board member on observer

status, and they have agreed not to make critical investment decisions without collaboration by 'Vital' (Kyrou, 2015). Another example is Salesforce Inc. CEO Marc Benioff, who uses an AI product called 'Einstein' to help him make "better decisions" (Taulli, 2020). The product provides Benioff with a data-driven analysis of quarter performance, strengths, weaknesses, and specific executives that may need close attention. The same data will be used in the board for decision making. Hence, the importance of the debate on the principal-agent problem, as AI technologies can act differently than the developer (the principal) intended or expected, as we have seen in the past (AI ethical scandals), which in some cases end up costing the shareholders their return on investment.

The current corporate governance model of "controlling, monitoring and supervising" to resolve the principal-agent problem has been based on the agent and principal being human. In terms of AI-driven technologies, it may be that most of the incentives and monitoring mechanisms used to resolve the agent-principal problem and motivate the agent to deliver shareholder value, no longer work when the agent is artificial. Van Rijmenam, Schweitzer & Williams (2019) assert that when dealing with an artificial agent, existing corporate governance models need to be adjusted to fit the characteristics of the artificial agent. They further argued that solving the principal-agent problem when dealing with artificial agents could help corporations "develop responsible conversational agents, where organisations exercise strict control, supervision and monitoring on the performance of the artificial agent" (van Rijmenam, Schweitzer & Williams, 2019).

In the context of South Africa, the viability of AI as a board member may present a challenge, unlike Vital's case in Hong Kong, particularly concerning the fiduciary duties under the Common Law and Companies Act 71 of 2008. The Companies Act outlines the responsibilities of directors to act in good faith and independently with a degree of care, diligence, and skill. Thus, while AI technologies can certainly contribute significantly to data-driven decision-making, the core fiduciary duties, such as loyalty and care, traditionally require a human understanding of ethical considerations, especially with the ethics of AI still under question. Integrating AI technologies into the board would require careful consideration of AI ethics, transparency, and accountability. Until then, AI technologies can be used to augment and support the decision-making processes for the board. The human judgement in fulfilling fiduciary duties and ensuring a balance between technological advancements and ethical governance should remain.

3.4.   CONCEPTUAL FRAMEWORK ON THE ROLE OF THE BOARD IN AI
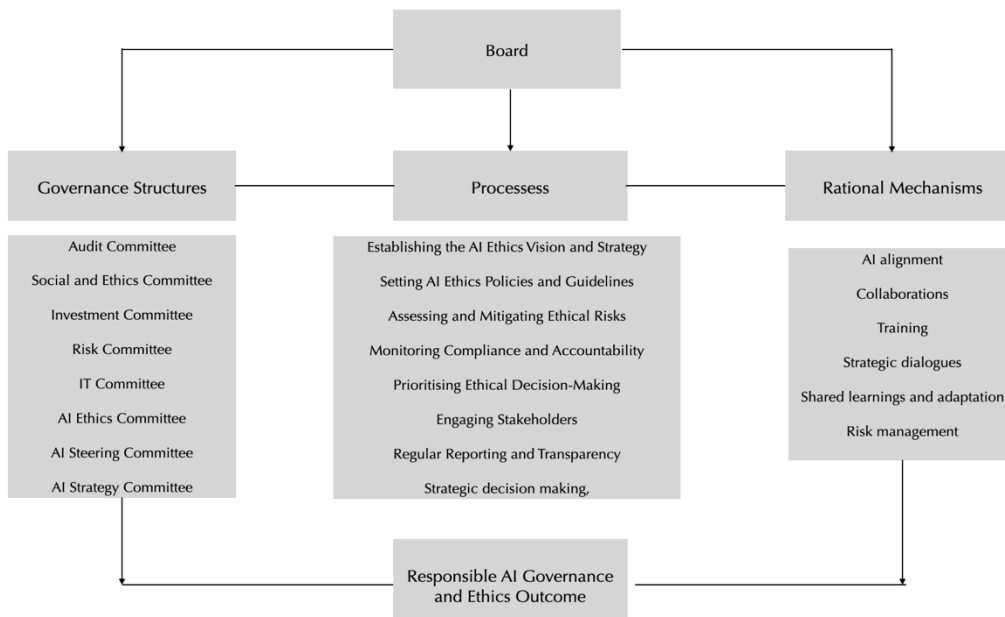
Figure 6: Conceptual Framework of the role of the board in AI

As already argued in this chapter, one of the main purposes of corporate governance (CG) is establishing governance and monitoring mechanisms. CG acts as a mechanism wherein the board (board of directors) is a monitoring tool to reduce/deal with the principal-agent relationship (Madhani, 2017). Over the last few years, there has been a substantial growth of CG research and scholars and have broadened the understanding of what constitutes proper CG practices (Huse, Hoskisson, Zattoni, & Viganò, 2011). The research on a effective formal CG mechanism has been amongst the top priorities for many scholars in this field. However, some scholars are now considering the broader GC structure and key players, such as the board's role in implementing AI governance and ethics. It is against this background that the study aims to contribute to CG theory and practices by exploring the role played by the board governance of AI ethics. Figure 6 above presents the overall conceptual framework of the study.

The conceptual framework primarily focuses on the role of the board as a focal point of CG, which means monitoring, accountability, and oversight of AI governance. The board is the overarching structure responsible for governance and oversight of the organisation's AI ethics. The board plays a key role in establishing the vision, setting policies, monitoring compliance, and ensuring accountability. AI could be governed through a mix of processes, structures, and rational mechanisms, as depicted in Figure 6 above.

Structural governance mechanisms define reporting structures, governance bodies, and accountability (Borgman et al., 2016). They comprise roles and responsibilities and the allocation of decision-making authority (Schneider, Abraham, Meske & Brocke, 2020). Accordingly, the board's role ensures the establishment of structures such as an audit committee, AI ethics committee, AI strategy committee, social and ethics committee, etc. These structures are responsible for overseeing and managing AI deployment, development, and implementation, including managing AI ethics-related activities, and ensuring AI strategies are implemented and aligned with business objectives. At the executive management level, the AI steering committee's role is to manage AI projects, costs, allocations, and risk management and to ensure that ethics is cascaded down to the lower levels of the organisation. Furthermore, management is responsible for ensuring the appointment of the right people who have experience, skills and expertise, such as the Chief Information Officer and AI Ethics Officer, who oversee AI-related activities, ensure ethical practices and monitor, assess, and address AI-related ethical concerns within the organisation.

Accordingly, mix of processes refer to "strategic decision making, strategic information system planning and monitoring". Processes aim to ensure that AI technologies operate correctly and efficiently, they are held securely, and their operation meets legal and company internal requirements and satisfies ethical values and principles concerning explainability, fairness, accountability, security, and safety. Processes comprise strategy, policies, standards, procedures, performance and monitoring mechanisms (Schneider, Abraham, Meske & Brocke, 2020). The processes is essential for the board to leverage issues such as information economics, AI service level agreements and the control of AI governance frameworks.

The rational mechanisms are critical to ensure business and AI alignment, and these include business units' participation in AI, collaborations, training, strategic dialogues and shared learnings (De Haes and Van Grembergen, 2006). Serban et al. (2020) express the need for interdisciplinary teams and collaborative development platforms of AI technologies, which means that business units should act as partners and share learning regarding the development, deployment, and adoption of AI technologies. The board should ensure active participation of key stakeholders in strategic dialogues, risk management and shared learnings.

3.5.   CONCLUSION

In this chapter, the researcher reviewed the definition of corporate governance from Cadbury Report (1992), Shleifer and Vishny (1997), Zingales (1998), Denis and McConnell (2003), OECD (2004), Solomon (2007), L'Huillier (2014), amongst others. In the end, the study aligned with the King IV report on corporate governance, which elevated the definition by focusing it not only on "directing and controlling" the company, but also highlighting the importance of stakeholder, ethical leadership, transparency, accountability, fairness and responsibility to achieve ethical culture, good performance, value creation, legitimacy and effective control of the organisation which is more relevant to the study. Although stewardship theory, resource dependence theory, and the resource-based view theory all make significant contributions to the field of corporate governance, this study explicitly favours agency theory and stakeholder theory. This decision was made because these theories provide more robust frameworks for examining internal governance mechanisms, conflicts of interest, and the ethical considerations and inclusive stakeholder approach in decision-making processes.

The chapter also discussed the different board models, i.e., the tie and two-tie board model and the control role, the strategic role, the stakeholder role, and the advisory and fiduciary duty of the board. The chapter further discussed AI technologies as board members and argued that AI technologies can contribute significantly to data-driven decision-making and argue that the core fiduciary duties, such as loyalty and care, and human understanding of ethical considerations, especially with the ethics of AI, would be required.

The chapter concluded with a conceptual framework which focuses on the role of the board as a focal point of CG, meaning the board is the overarching structure responsible for governance and oversight through a mix of processes, structures, and rational mechanisms of the organisation's AI ethics.

CHAPTER 4: ARTIFICIAL INTELLIGENCE ETHICS

4.1. INTRODUCTION

The integration of AI technologies has raised several ethical issues, including concerns about privacy (such as compromising personal information), decision-making biases (stemming from flawed historical data), opacity (unclear decision-making processes), and the potential displacement of human jobs by automation. These issues pose risks not only to society at large but can also erode trust in businesses that implement AI without sufficient oversight. Consequently, if not properly regulated, AI can damage the reputation of a business. Thus, the importance of the role of the board in governing AI ethics to protect the society, business and the environment.

This chapter focuses on reviewing literature related to definitions of ethics, ethics sub-disciplines (metaethics, normative ethics and applied ethics), and AI ethics. The chapter will also discuss the concepts of fairness, accountability, safety, and explainability of AI. The chapter will also discuss in more detail than in the previous chapter the role of the board in AI ethics and organisational ethics. The literature review on ethics in general and AI ethics in particular is of utmost importance to provide the researcher with more information about the subject. It further helps the researcher define the study, demonstrate the research need, and derive its theoretical background from the existing literature (Bryman, 2012).

4.2. ETHICS DEFINITION

Ethics has been studied from different disciplines over centuries and scholars agree that ethics is a complex, complicated, and convoluted concept (Siau & Wang, 2020). Ethics is concerned with questions of how people ought to act (Velasquez, Andre, Shanks & Meyer, 2010), and the search for a definition of right conduct (Gilman, 2005). While ethics is a critical aspect of human society, it is important to recognise that different cultures and individuals may have varying perspectives on what constitutes ethical behavior. Therefore, any definition of ethics must consider the complexities of human experience and the diverse perspectives that shape moral decision-making.

Mle (2012) defines ethics as "what is right and wrong, what is acceptable or unacceptable and is intertwined with the value system of people" (p.26). Mle's (2012) definition of ethics

highlights the intricate relationship between moral standards and personal values. Ethics encompasses not only the distinction between right and wrong actions but also the assessment of what is considered acceptable or unacceptable within a given context. This intertwining of ethics and value systems underscores the complex nature of ethical decision-making, as individuals must navigate the differences of their own moral frameworks when determining the appropriateness of certain actions or behaviors.

On the other hand, Gildenhuys (2004) refers to ethics as "principles or standards of human conduct…ethics and morality…are two philosophical concepts that deal with what is good and what is bad; what is right and what is wrong; and what is acceptable and not acceptable". The term 'ethics' is used by philosophers to refer to the study of morality, which is understood as a set of social rules, principles, norms that guide or are intended to guide the conduct of people in a society, and as beliefs about right and wrong conduct as well as good or bad character (Gyekye, 2011). In other words, ethics is a system of principles or rules (and related reasoning models) that help a particular society or group determine what is right versus wrong, and the moral obligations and duties of entities i.e., Western theories such as consequentialism, deontology, and virtue ethics or African ethical theories that are communitarian in nature, such as ubuntu[21].

Gyekye (2011) argues that the ethics of a society is embedded in the ideas and beliefs about what is "right or wrong", what is a "good or bad character"; it is also embedded in the conceptions of satisfactory social relations and attitudes held by the members of the society; it is embedded, furthermore, in the forms or patterns of behaviour that are considered by the members of the society to bring about social harmony and cooperative living, justice, and fairness. Ethics primarily deals with the philosophical aspect of human action and trying to assess the question of right and wrong and also judge the moral consequences of such human action. Essentially, ethics is a discipline that serves to find a balance in uncertain situations which may arise when processes are interpreted differently. Ethics is more focused on processes and reasoning than on judgements. The study of ethics is broken down into three sub-disciplines, namely, metaethics, normative ethics and applied ethics.

---

[21] African Ethics refers to values associated with the largely black peoples residing in the sub-Saharan part of the continent, thereby excluding Islamic Arabs in North Africa and white Afrikaners in South Africa, among others (Metz, 2007).

Metaethics is primarily concerned with the status and nature of ethical claims society makes. It considers the meaning of ethical judgments, and seeks to understand the nature of ethical statements, attitudes, and judgments, and how they may be supported or defended.

Normative ethics is primarily concerned with establishing how things should or ought to be, how to value them, what is regarded as good or bad, and which actions are right or wrong. Normative ethics attempts to regulate behaviour through developing sets of rules governing human conduct, or a set of norms for action. For example, the Golden Rule is an example of normative ethics: treat others like you want to be treated (Gavrilova, 2021). Kumar & Choudhury (2022) argue that AI technologies may be trained and coded with normative theories to teach AI technologies how to understand and predict everything such as good or bad, and which actions are right or wrong. This area of research is known as machine ethics as it incorporates insights from AI technologies and normative theories and seeks to embody ethical reasoning in machines in order to ensure machine ethical behaviour as well as to explore the nature and computability of human ethics (Brundage, 2014).

As robots take on more and more responsibility, they must be programmed with moral decision-making abilities. Some authors have proposed that software agents and robots inculcate explicit ethical principles to govern their behaviour (Allen & Wallach, 2010). Asaro (2020) argues that the best approach to machine ethics is by ensuring (i) ethics systems built into AI robots, (iii) the ethics of people who design and use AI robots, (iii) and the ethics of how people treat AI robots. He argues that through this approach it is "possible to think of a continuum of agency that lies between amoral and fully autonomous moral agents" (Asaro, 2020). The normative theories in question might either be translated as a set of universal rules that are always valid, or they could also be translated into a set of rules (and, as a result, an algorithm) that can be applied exclusively to specific instances.

Applied Ethics is a discipline of philosophy that attempts to apply ethical theory to real-life situations (Battersby, 1989). Applied ethics is a subdiscipline of philosophy. Philosophy as a discipline discusses problems such as the ideal or the purpose of life, the nature of knowledge and reality, justice, and the norms of right actions. Applied ethics not only studies ideas and theories from the perspective of general studies of 'good and bad' or 'right and wrong' actions but also attempts to find solutions to these problems in a particular field. On the other hand, philosophy analyses and clarifies moral problems in an abstract manner but is not focused on solving domain-specific problems. Applied ethics is a principle-based

ethical approach which searches for solutions to specific problems that are not necessarily universally acceptable or impossible to implement. The word 'applied' is a technical term describing a variety of new philosophical enterprises (Kopelman, 1990). Applied ethics focuses more closely on ethical discussion and focuses more on the ethical obligations of a moral agent in a specific situation (Bagnoli, 2013; Copp, 2006; Horgan & Timmons, 2006).

AI ethics is a sub discipline of the field of applied ethics. It is concerned with two things: the moral behavior of humans that create AI technologies and the impact of AI technologies on society, which both relate to governance of these technologies. The main differences between ethics in general, and AI ethics in particular, are automation and scale. AI ethics challenges are automated by the very nature of AI technology.

The development of AI technologies is largely concentrated in North America and Western Europe, which has resulted in its standards and thought processes being predominantly influenced by these cultures. Consequently, the discourse on AI ethics has primarily been discussed from an Anglo-American (Western) perspective, which tends to overlook the value that relational theories might bring to reflection on the ethical implications of AI on human relationships and social structures. Relational ethics, as highlighted by scholars such as Jecker, Jonsen & Pearlman (1997), Metz (2007), Held (2006) and Coeckelbergh (2010), emphasises the significance of relationships and context in moral decision-making. Jecker, Jonsen, and Pearlman (1997) argue that moral decision-making should consider the intricate network of relationships and contextual factors that shape our actions and interactions with others.

Metz (2021) critiques the utilitarianism approach in AI as opposed to a relational ethics based on African values. This approach emphasises the significance of relationships and interconnectedness in moral reasoning (Metz, 2021), thereby offering a framework for ethical decision-making that is cognizant of human connections and the impact of our actions on others. In the context of AI ethics, it is imperative to recognise and acknowledge historical injustices and the present-day consequences of development, deployment and overall adoption of AI technologies on marginalised communities (Mhlambi, 2020). Therefore, integrating relational ethics into AI ethics highlights the need for boards to understand the social, cultural, and relational dynamics of the communities affected by AI technologies (Held, 2006).

This perspective aligns well with stakeholder theory (adopted by this study; see chapter 3), as this theory focuses on the impact of decisions on various stakeholders and the importance of maintaining trust and cooperation for successful leadership. In an AI ethics context, stakeholder theory would guide the development, deployment and overall adoption of AI technologies that consider the welfare and rights of all affected parties, ensuring that AI technologies serves the interests of diverse groups rather than a select few.

4.3. AI ETHICS

Due to the challenges that arise from the research, design, development, deployment and overall adoption of AI systems, the consensus has been that these AI technologies may amplify and contribute to existing structural bias, and lack accountability and transparency, among other concerns. This state of affairs is exacerbated by the fact that identifying shared values to address these concerns globally is not an easy task (Marwala, 2020). Human bias has been cited as the most significant factor that has in most cases been inherited by AI technologies (Wang & Siau, 2018), such as the gender bias in recruitment being replicated through the use of machine learning (Larson, 2017; Raso et al. 2018) or race bias perpetuated through machine learning in probation processes (Koolen and Cranenburgh, 2017). AI systems are trained by a human on datasets containing human data, therefore, existing biases in training data sets may be learned in the process by the algorithm and be displayed and amplified in real applications. For instance, a software used to predict future criminals showed bias against a certain race (Bossmann, 2016).

Another most cited AI ethical issue is that of privacy and data protection (Stahl, 2021). The key privacy concern is informational privacy, and data protection can be understood as a means to safeguard informational privacy (Stahl, 2021). Data collected by means of AI technologies raises privacy issues such as informed consent freely given, opt out options, limiting data collection, describing the nature of AI processing, being able to revise or delete data on request, and third-party sharing practices. In some cases, data collected on human subjects through a "data repurposing"[22] or perhaps given a spillover[23] effect, may also result in people not knowing that data was collected about them, thus limit the ability to make any inquiries to corporations with respect to their own data or to request that it be deleted.

---

[22] Repurposing data – is when data is used for a completely different decision/task to what it was originally intended to be used for (Woodall, 2017)

[23] Data spillovers – data collected on people who are not the target of data collection

Safety is also a key ethical issue of AI technologies, in particular in systems that interact directly with the physical world, such as autonomous vehicles or systems governing critical healthcare provision. Other ethical issues have already been raised in Chapter 1 of this study, such as a "big brother state" through mass surveillance programmes, fears that these technologies might replace jobs for human workers, that autonomous weapons may be misused at the hands of malevolent characters and that their development eludes accountability. These fears have been exacerbated by cases such as algorithmic discrimination in facial recognition systems which have been found to have misidentified Africans and Asian people (Waelen, 2022) more often than white people and so on.

AI ethics covers a dominant global debate amongst academics and policymakers. Significant resources have been dedicated to initiatives aiming to define universal values or principles to guide ethical development and deployment of AI technologies globally. These initiatives are meant to bring consensus on a common set of issues and principles, and raise awareness on the ethical challenges that accompany AI technologies (Mittelstadt, 2019). The field of AI ethics has largely emerged as a response to the range of individual and societal harms that the misuse, abuse, weak governance, poor design, or negative unintended consequences of AI technologies may cause (Leslie, 2019).

AI ethics is regarded as a subfield of applied ethics, which seeks to understand how to address and manage risks and adverse outcomes of the design, development, deployment and use of AI systems for all stakeholders in specific domains. AI ethics is a study that examines the ethical principles, guidelines, and regulations surrounding the use and development of AI technologies. Essentially, it is the study of the moral obligations and duties of entities, such as humans and intelligent robots, in relation to AI technologies (Butvinik, 2022). According to Leslie (2019) AI ethics is defined as:

> *" A set of values, principles, and techniques that employ widely accepted standards of right and wrong to guide moral conduct in the development and use of AI technologies"* *(Leslie 2019, p. 3).*

Lawton & Wigmore (2021) defines AI ethics as a system of moral principles and techniques which informs the development and responsible use of AI technologies. Müller (2020) argues that AI ethics sets out moral principles that guide moral obligations and

responsibilities of individuals or groups in AI adoption. These principles serve as a guide for decision-making during ethical dilemmas. These principles focus on designing and building AI technologies that reflect awareness of the values and principles to be followed in the deployment scenarios to minimize ethical risks that can arise from the design, inappropriate application, or intentional misuse of the technology.

There is widespread agreement that AI technologies should be used for the common good, should not be used to harm society or undermine their rights, and should respect values such as fairness, privacy, and autonomy (Whittlestone, Nyrup, Alexandrova & Cave, 2019). Thus, the need for AI ethics to focus on core issues such as bias, fairness, safetey and privacy. Thus AI ethics refers to the guidelines, declarations, ethical codes, and normative rules-of-thumb for users and practical checklists for designers in the development and responsible use of AI technologies (Jobin et al., 2019; Mittelstadt, 2019, Hagendorff, 2020). The main purposes for development of these values, principles, and techniques is to support and improve the ethical design, development, and deployment of AI technologies (Hallamaa, 2022). Secondly, AI ethics provides a general basis with respect to theory, principles, values, etc., for the designers to apply in the design and development of AI technologies thereby improving the ethical sustainability of AI development (Morley et al., 2019; Hagendorff, 2020).

In the last few years advances in the capabilities of AI technologies have led to growing ethical concerns around these developments. As a result, there has been a spike in publications of AI ethics principles; over 200 sets and increasing (Algorithm Watch, 2024); from civil society organizations, research centers, private companies, and governmental agencies. These ethical principles statements in AI contribute to the formation of a moral background for analysing the lifecycle of AI systems as they make the connection between values, ethics, and technologies. AI ethics principles statements explore how to conceptualise, analyse, and assess ethically relevant features of AI technologies design, development, and deployment/application as well as how to determine methods of directing, regulating, and governing AI technologies in an ethically sustainable manner (Hallamaa & Kalliokoski, 2022).

These principles are also treated as the guides on how the policymakers and professionals should prioritize and structure future legislation, standards, and governance models. Many envision these high-level contributions to be translated into technical fixes, governance

frameworks, developer codes of ethics and actionable design requirements at lower levels (Mittelstadt, 2019; Jobin, Ienca, & Vayena, 2019). On the whole, the concern is that these publications are too abstract, sometimes contain vaguely formulated principles, have not proven to be helpful in guiding practice (Canca, 2020; Hickok, 2021), and fail to address fundamental normative and political tensions embedded in key concepts i.e., fairness, privacy (Mittelstadt, 2019).

In some cases, these principles differ in approach, but the similarities between the suggested formulations of AI ethics indicate that there exists a common will to develop AI technologies in a morally responsible way (Hallamaa & Kalliokoski, 2022). Despite the differences in AI ethics formulations, they propose to offer practical ethical guidance and support for those who develop, design and deploy AI technologies to ensure that these technologies perform the task in an ethical and responsible way.

In the quest for actionable AI ethics principles, AI ethics should be approached through the stakeholder theory lenses; to ensure inclusivity of the various stakeholders and society at large to counter commercial values driving advancement in AI technologies (Ruttkamp-Bloem, 2021). To prevent such abuses, companies working on ethical guidelines for AI technologies need to engage users from around the world to help create appropriate standards to govern these systems. Inclusivity and geographic representation, particularly the marginalised in AI ethics, are important. The aim is to work towards an agreed consensus on a set of international definitions and concepts for ethical AI. Gupta (2020) argues that more inclusive geographic representation will produce a global vision for AI ethics that currently reflects only the perspectives of people in a few regions of the world, particularly America and Europe.

Furthermore, critics argue that "there is a blind spot in AI ethics" (Crawford & Calo, 2016) which refers to the ways in which AI technologies perpetuates and amplifies existing biases and discrimination. They identified several sources of this "blind spot," including: (i) the data that is used to train AI technologies; (ii) the algorithms that are used to build AI technologies; (iii) lack of diversity in the field of AI; and (iv) lack of transparency in AI technologies (Crawford & Calo, 2016). Researchers argue for a more fine-grained technology assessment of AI technologies, rather than repeatedly discussing the same small set of abstract ethical issues. AI ethics often fails to ask the question whether some of these

systems should be developed or applied in the first place i.e., development of autonomous weapons (Hagendorff, 2021).

Canca (2020) argues that AI principles should be operationalised so that we help computer developers, and designers to (1) spot and think through ethical issues; and (2) recognize when a complex ethical issue requires in-depth expert analysis. It should be recognised that by their nature, AI ethics principles are not complete systems for ethical decision-making and not suitable for solving complex ethical problems. But once operationalised, they provide a valuable tool for detecting, conceptualizing, and devising solutions for ethical issues.

However, the AI ethics debate have changed significantly along the way, with new theories, frameworks and guidelines being developed as a form of checks and balances in the design, development, and implementation of AI systems, and much more emphasis on policy actions as opposed to listing abstract values and principles in isolation of implementation discussions. The aims of AI ethics can be achieved in the corporate context by studying and proposing technical and nontechnical solutions to AI ethics concerns, such as in the case of this study, reflecting on the board's role in ensuring AI technologies that are data responsible, respect privacy, and are fair, value-aligned, accountable, and transparent.

4.3.1. AI Ethics Principles

A review of AI ethics published guidelines found that most of the publications focus on similar values and priorities, however, with different definitions. In their 2019 meta-analysis of the "global" landscape of AI ethics guidelines, Jobin et al. (2019) reviewed 84 AI ethics guidelines publicly released. They found that that the most common principles are: "transparency, justice and fairness, nonmaleficence, responsibility and privacy". These overlap interestingly with the medical ethics principles of "nonmaleficence, beneficence, autonomy and justice" (ibid.). A similar study (Hickok, 2021) conducted on a different sample in 2020 found that while there are certainly points of convergence, by no means is there unanimity across the 36 principle documents reviewed. Fairness and non-discrimination principles were present in all documents, and the other key themes were privacy, accountability, safety and security, transparency and explainability, human control of technology, professional responsibility, and promotion of human values (Hickok, 2021).

The European Commission AI HLEG (2019) suggests that AI principles should be implemented throughout all stages of AI technologies' life cycle, and that both technical and non-technical approaches should be considered when implementing AI principles. Hagendorff (2020) argues that some of the typical AI ethics principles, such as "accountability, privacy, fairness, explainability and safety, are most easily operationalised mathematically and thus tend to be implemented in terms of technical methods". Other technical methods include continuous monitoring and rigorous testing and validation of the AI system, whereas the non-technical methods comprise standards, certification, governance frameworks, education and awareness, stakeholder participation, and diversity in AI design and development teams (AI HLEG, 2019).

This study focuses on the most common AI Ethics principles, i.e. accountability and responsibility, fairness, transparency, explainability, and privacy, in order to set the groundwork for defining the role of the board in private corporations in actualising these core AI ethics principles. Below follows a brief description of these principles.

4.3.1.1. Accountability and responsibility

Beu and Buckley's (2001) definition: "accountability is the perceived need to defend or justify behaviors to an audience with reward/sanction authority, where the rewards/ sanctions are perceived to be contingent upon the audience evaluation of such conduct". Accountability means the state of being responsible for AI technologies, their behaviour and their potential impacts. Buhmann et al. (2019) argue that with regard to 'algorithmic accountability,' there may be special, discrete accountability characteristics specific to AI technologies that render expectations of accountability as highly fluid.

Algorithms themselves cannot be held accountable as they are not moral or legal agents (Raji et al., 2020). AI technologies have the potential to significantly complicate traditional governance mechanisms of accountability by dispersing moral agency among potentially numerous agents who may be responsible for a grievous harm in the society (Tóth et al., 2022). However, the organisations designing and deploying algorithms can through governance structures identify who is responsible and accountable during the different stages of the AI lifecycle and ensuring human oversight and controls (Kompella, 2022). The burden of accountability and responsibility should be allocated among those who develop, deploy, and use these systems (Ryan 2020).

To ensure AI accountability, Fjeld et al. (2020) suggest the establishment of internal review committees to oversee the use and development of AI. Such committees should have the power to critically examine, approve or deny any use of AI technologies in terms of legal compliance, potential risks, and opportunities. The committee should be composed of different stakeholders from many business functions, including legal, operations, compliance, marketing, data science, and security (Kroll 2018). This is a role that I argue should be ascribed to the boards of private corporations.

Human control is also linked to accountability, and AI technologies should be built in such a way that humans can intervene in their actions or decision-making (Fjeld et al. 2020). The AI HLEG (2019) defines human control as human agency and oversight. A human control paradigm ensures that humans have the power to influence or direct the course of events as well as the ability to manage AI technologies. Using this paradigm to design AI technologies can ensure accountability. The capability for human intervention during both the design and implementation, including the possibility to override a decision made by the AI technologies when it violates the law (Kumar et al.,2020), would ensure greater accountability of AI technologies.

4.3.1.2. Explainability (XAI)

In recent years, the concept of eXplainable AI (XAI) has come to dominate global debates on AI strategies for research institutions (Braunschweig, 2016), countries (Holdren et Smith, 2016; Villani et al., 2018) and in the private sector (Mars, 2019; Guggiola et al., 2018), including large corporations such as Microsoft, IBM and Google. In simple terms, XAI is about providing explanations regarding AI processes to stakeholders (Mars, Dès, & Boussard, 2020). XAI is concerned with implementing the transparency and tractability of AI and machine learning (ML) methods (Holzinger et al., 2022) to unveil the underlying explanatory factors of AI technologies decision-making processes (Pearl, 2019). It contrasts with practices related to the concept of the "black box" in ML in the sense that even their designers cannot explain the AI technologies decision-making processes, and as a result, simply not addressing explainability issues. XAI on the other hand is an implementation of the social right to explanation (Lundberg & Lee, 2017).

Wang et al. (2020) propose three desirable factors that XAI should satisfy: (i) improve people's understanding of the AI model, (ii) help people recognise the model uncertainty, and (iii) support people's calibrated trust in the model. Barredo Arrieta et al (2020) suggest that XAI have the potential to unlock other AI principles, such as fairness, transparency, accountability, safety, and privacy.

XAI is crucial for corporations to build trust and confidence and to provide greater clarity on their AI governance. Corporations developing AI technologies should be able to explain the data sources and decision-making processes up to inherent mathematical opacity. Society needs a degree of traceability to ensure that if harm arise, it can be confidently traced back to the cause. Such transparency will also help these organisations adopt a responsible approach to AI technologies design, development, and implementation. Explainable AI (XAI) technologies support an AI governance narrative as it ensures the ethical use of AI.

4.3.1.3. Fairness

Fairness is commonly defined as the state of being fair, especially fair or impartial treatment; it can also mean different things in different contexts to different people (Mulligan, Kroll, Kohli & Wong, 2019). In philosophy, fairness rest on a sense that what is fair is also what is morally right and just, while in law, fairness includes protecting individuals and groups from discrimination or mistreatment with a focus on prohibiting certain behaviors and biases. While there is no universally agreed upon definition for fairness, we can broadly define fairness in our current context as the absence of prejudice or preference for an individual or group based on their characteristics. Fairness is closely linked to non-discrimination, inclusivity, and the prevention of bias (Fjeld et al. 2020).

The most commonly discussed fairness-related objection that has been raised in connection with algorithmic decision-making is that predictive factors can be unfairly biased (Chiao, 2019) if data sets contain inherent structural bias (identity prejudices). In recent years, concerns around algorithm fairness have made headlines. For example, in the United States criminal justice system AI technologies had falsely predicted future criminality among African-Americans at twice the rate as it predicted for white people (Waelen, 2022). In another case of a hiring application, it was recently exposed that Amazon discovered that their AI hiring system was discriminating against female candidates (Andrews & Bucher, 2022). Lawton & Wigmore (2021) argue that in data sets involving personally identifiable

information, it is crucial to ensure that there are no biases in terms of race, gender or ethnicity. It is suggested that these issues can be addressed with high-quality and representative datasets, which is why these should be measured and monitored for accuracy, consistency, and validity (Fjeld et al. 2020).

Kroll (2018) suggests introduction of data governance to manage fairness issues. Structures such as a review committees composed of diverse and cross-functional committees to analyse ethics, legal compliance, risks, and impacts of AI technologies, processes in place to investigate foreseeable issues and risks, and how to mitigate them, and continuous monitoring of correctness (e.g., review modelling errors, concept drifts, and bias). There are also a number of technical resources such as Fairness Flow available to identify fairness issues, but again, we need more than just a technical fix.

4.3.1.4. Transparency

In the age of pervasive AI, and the ability of AI technologies to make automated and self-learned decisions, a call for transparency of how these technologies reach decisions has echoed within academic and policy circles (Felzmann, Fosch-Villaronga, Lutz & Tamò-Larrieux, 2020). Transparency is critical to justify extensive data collection and its benefits to users (Kumar et al.,2020). Transparency can be understood in two ways: (i) the transparency of the AI technologies itself and (ii) the transparency of the corporations(s) developing and implementing AI technologies (Ryan and Stahl 2020).

Transparency of AI technologies relates to understanding the design of the AI system and how it reaches a decision (Barredo Arrieta et al., 2020). Corporations' transparency refers to the understanding of 'what, why, and by whom' the decisions were made during the development and design processes (Vakkuri, Kemell, and Abrahamsson 2019). There is a third aspect to transparency which is the right of members of society to know that AI systems are taking decisions that impact on individual persons' lives. Transparency leads to trust in AI technologies, when humans are able to understand, interpret and explain the "why" of AI decisions-making process. Such ability of providing interpretable explanations of AI decisions making processes is also linked to the accountability of AI technologies, as well as their explainability.

Transparency in AI plays a very important role in the overall strive to develop more trustworthy AI technologies. Transparency in AI, consequently, can best be seen as a balancing of interests and a governance challenge demanding multidisciplinary development (Larsson & Heintz, 2020).

4.3.1.5.   Privacy

Privacy, is one of ten important principles of the ethics of AI mentioned in the UNESCO Recommendation on the Ethics of AI, but also in most other mainstream documents. Privacy is a normative concept deeply rooted in philosophical, legal, political and economic literature. Privacy is "the right to be let alone", as a component of personhood, control over personal information, and the right to secrecy (Zhang et al., 2021). A well-researched understanding of privacy harms has been developed in a vast body of literature and addressed by different legal frameworks (Solove, 2006). Data privacy and control over personal information arguably becomes increasingly important aspects of privacy protection, as AI technologies bring further threats to privacy protection (Tian, 2016). Privacy concerns regarding the collection, storage and use of personal information are a recurring topic of public discourse, thus, the protection of human dignity, human autonomy and human agency must be respected, protected and promoted throughout the life cycle of AI technologies (UNESCO, 2021).

AI technologies processing makes use of vast amounts of data, and some of it could be sensitive personal information. Through analyses some of the data could be used for the intended purposes, but there is a number of cases where data has been abused or used for targeted advertising. AI technologies expands the ability to use personal information in ways that can infringe on privacy interests by bringing personal data analysis to new levels of power and speed (Kerry, 2020). There has been calls for greater oversight of the big data in the health sector for instance, to ensure appropriate safeguards are in place to maintain privacy and patient agency. The biggest concern is the external risk of privacy breaches as some AI technologies have the ability to reidentify data that have been anonymised or deidentified (Murdoch, 2021). AI technologies have made it easier to gather personal data through technologies such as smartphone, internet and surveillance cameras (Bartneck, Lütge, Wagner & Welsh, 2021). Key to privacy is to allow people to make their own, non-coerced decisions, to better calculate their behaviour, and also to take decisions and actions that do not conform to certain social norms.

UNESCO Recommendation on the Ethics of AI (2021) recommends an establishment of a multi-stakeholder approach for data protection frameworks and governance mechanisms throughout the life cycle of AI technologies. It further recommends that an adequate privacy impact assessment should be performed, such assessment to include societal and ethical consideration of privacy on the design and development of AI technologies. Moreover, those responsible for AI technologies should ensure accountability for their design and implementation and ensure that personal information is protected throughout the life cycle of these technologies (UNESCO, 2021).

4.3.1.6. Safety and security

Safety is key to the protection of human life, property, whereas security refers to protecting information confidentiality, integrity, availability, etc. "Safety is caused by reasonably foreseeable misuse, equipment malfunction, whereas security is caused by the intended attack". Safety issues can easily be detected as they appear as an accident, whereas security concerns, such as surveillance and intrusion, are difficult to detect (Kaneko, Yoshioka, & Sasaki, 2020). The main concern with safety and security is safety risks (unwanted harms) and security risks (vulnerable attacks). The corporations should always put systems in place to prevent and eliminate such risks throughout the life cycle of AI technologies to ensure human, environmental, and ecosystem safety and security. According to UNESCO (2021), safe and secure AI technologies will be "enabled by the development of sustainable, privacy-protective data access frameworks that foster better training and validation of AI models utilizing quality data".

4.3.1.7. AI principles conclusion

Mittelstadt (2019) argues that principles alone cannot guarantee ethical AI, and for principles to be useful in practice, they must be able to guide actions (Whittlestone et al., 2019). AI ethics principles tend to be highly generalised and too broad to be action guiding (Whittlestone et al., 2019). However, principles can be used to guide the development of governance practices, international standards, and further regulation (Whittlestone et al. 2019). Indeed, Fjeld et al. (2020) suggest that "principles are a starting place for governance," which has led many national and international organizations to develop expert committees on AI (Jobin et al. 2019).

AI ethics principles provide valuable inputs for a corporation's AI strategy. These principles give corporation(s) a handle on the acceptable use of AI and even to determine whether an AI technology is fit for purposes. AI ethics principles also provide clarity on design, data, documentation, testing and monitoring requirements. These principles are relevant throughout the entire AI lifecycle (Kompella, 2022).

In order to successfully action such principles, AI ethics should be integrated into the broader organisation ethics, culture and values. This implies that AI principles should be incorporated into corporate values, policies, codes of ethics, and guiding principles applied to AI technologies' product life cycle, from research and design to implementation or deployment. Häußermann and Lütge (2022) argue that AI ethics discourse is "too narrow a focus on technical systems," resulting in a tendency to ignore the context of using AI technologies integrated into business practices. Thus, the board should ensure a clear alignment of AI technologies to ethical culture and corporate objectives.

4.4.   ORGANISATIONAL ETHICS IN THE ERA OF AI TECHNOLOGIES

As discussed at the beginning of this chapter, ethics is derived from a branch of philosophy that aspires to analyse and recommend moral principles that govern 'right and wrong' in the society. Organisational ethics, also known as business ethics, sets out doctrines and standards within the business world so that corporations can operate with integrity and virtue. Organisational ethics focuses on the study of organisational situations, activities, and decisions where issues of right and wrong are addressed (Crane, Matten, Glozer & Spence, 2019). Organisational ethics is often limited to philosophical theories related to 'right or wrong' decisions (Ferrell, Harrison, Ferrell, & Hair, 2019). Corporate failures and misconduct have substantially contributed to the growth of business ethics as a discipline (Rockness & Rockness, 2005; Towell, Thompson & McFadden, 2004). Over the past decades, business ethics has built a profound literature body dealing with various corporate governance failures and implementing positive organizational ethics (Michalos, & Poff, 2012).

Organisational ethics can be evaluated from both a normative or descriptive perspective (Ferrell, Crittenden, Ferrell & Crittenden, 2013). A normative perspective in this case relates to principles, values and norms for organisation decision making while a descriptive perspective refers to codes, standards of conduct, and compliance systems; i.e., decisions

that can be judged right or wrong by stakeholders such as customers (Ferrell, Harrison, Ferrell, & Hair, 2019). Therefore, business ethical decision-making is closely linked with internal organisational decisions by individuals or groups about appropriate business conduct and such decisions can impact internal stakeholders and external stakeholders.

In organisational ethics in particular, the rise in AI technologies has opened up debates on how to ensure that AI technologies act in line with organisational core values. Establishing an organisational culture for AI technologies' research, design, development, and deployment is of central importance to effective AI ethics and governance. Not every decision that AI technology developers face can be prescribed, and where formal rules do not dictate specific action, organisational culture, values and norms of conduct can provide overarching guidance to fill in the gaps (Seger, 2022). For example, think of the use of general values to guide medical decision-making in pre-clinical emergency settings where time and access to more explicit guidance is limited (Torabi et al., 2018).

Thus, AI ethical principles, i.e. accountability and responsibility, fairness, transparency, explainability, and privacy, should be closely linked into the overall organisational ethics. It has been well researched that without AI ethics governance, corporations can break customer trust, perpetuate bias, open themselves up to safety concerns, etc. – all of which generate risk to the organisations. Thomaz et al., (2021) argues that AI ethics should be a logical extension of a corporate strategy and objectives. An AI ethics statement should serve as a blanket declaration that the adoption, use, and design of AI technologies are consistent with both AI ethics principles and broader business ethics principles. For the purpose of enhancing transparency, the statement should include the hierarchy of ethical values of the corporation.

4.5.   AI TECHNOLOGIES' ALIGNMENT TO ETHICAL CULTURE (VALUES)

The alignment of organisational cultural norms with AI technology development is key to the uptake and efficacy of explicit rules and regulations. This is so as individuals are more likely to adhere to norms they believe in (Seger, 2022). Explicit rules and requirements often generate extra work, and time-consuming requirements like checklists, self-evaluations, and impact statements may lead to resentment by AI developers if seen as unnecessarily restricting (Baum, 2017). Such resentment may, in turn, lead to rejection of the very principles being promoted (Pettit, 2002).

However, rules and regulations are effective if they appeal to the cultural norms and values that are held in the organisations to which they are being applied. Organisational culture motivates individual behaviour if internalised, through internally generated rewards and punishment (e.g., shame or pride) (Coleman, 1994). In respect of AI technologies, we may similarly expect that any rules and requirements will be more effective and more willingly adhered to if those extrinsic measures align with organisational cultural ones (Seger, 2022).

The goal of AI technologies alignment is to ensure that powerful AI is properly aligned with human values (Russell, 2019). It's necessary to align the processes of AI technologies with corporate strategy, ethical values, and overall corporate objectives in order to explicitly translate relevant ideas into specific norms that influence concrete design and governance to shape the development and use of AI technologies in corporations. Organisations should create AI ethics frameworks that are actionable, with clear accountability and concrete methods of monitoring. To align AI technologies with corporate values and business ethics implies that explicability, transparency, accountability, fairness, responsible use of data, guarding against criminal misuse, aligning AI with values and being mindful of the societal implications of AI, are all essential pillars of ensuring responsible use of AI technologies.

The design, adoption, and implementation of AI technologies should thus be aligned to the broader organisational strategies (Eitel-Porter, 2021; Jöhnk et al., 2020). The organisational AI technologies strategy provides a general direction and manages expectations regarding the overall set of AI technologies that the organisation intends to use and what they are meant to achieve. The strategically aligned use of AI requires organisational resources, capabilities, and processes that can be ensured (Shneiderman, 2020).

Ethical culture requires boards to state the value base and AI ethics principles to which the organisation adheres, and adherence to these ethical values and principles across the organisation's AI systems must be ensured. (Whittlestone et al., 2019). Furthermore, the alignment should consider AI technologies' impact on stakeholder groups, as in algorithmic impact assessment (Kaminski & Malgieri, 2020; Schiff et al., 2021). The inclusion of strategic and value alignments as separate themes highlights alignments at different levels. Most significantly, organisations need to align the use of AI with organisational strategies and objectives (Schneider et al., 2020) and with norms and ethical principles that emanate from the broader environmental layer (Floridi, 2018).

The board of directors set the tone for any ethics and compliance program in corporations. King IV report explains that the board must govern ethics to ensure that the ethical culture within a corporation aligns with the tone set by the board. This is done through the implementation of appropriate policies and practices. Therefore, it is the responsibility of the board and executive management to ensure that technological advancements in AI align with the organisation's ethical culture. King IV report also mentions how to align technological advancement such as AI with ethical principles (Mulamula, 2022):

- Organisational values, ethics, and culture: The board should govern the organisation's ethics in a way that supports the establishment of an ethical culture. The board must ensure that the Code of Ethics and Code of Conduct guide the responsible development of AI technologies features in products and services, including when to use and not use AI. The content and principles embodied in the Code of Ethics and Code of Conduct must be integrated into employee training.

- Responsible corporate citizenship: The board should oversee governance and activities that demonstrate the company's good corporate citizenship. It must ensure AI technologies' compliance with the national Constitution, laws, and standards as well as company policies and procedures, strategy, and the Codes of Conduct and Ethics. The board should also ensure, through engagement with stakeholders and communities, that AI technologies advancements improve the material well-being of the societies in which the organisation operates. This will ensure that racial discrimination, data biases, and privacy are addressed in all product design, development, and adoption stages, and further provide buy-in from society.

Boards must ensure that AI technologies are aligned with risk management of the organisation and they must take responsibility for ethical culture as the tone is set from the top (Barnett 2019; Halstead et al. 2019). The expectation from stakeholders, including government regulators, is that culture should be a board priority (Petschler 2019). To enable AI technologies to be best used while remaining anchored to a corporation's core values, corporations (boards) must ensure the right framework to ethically govern artificial intelligence technology. The framework should set out decision-making rights and accountability for the board and the management team in regard to how ethics standards are set and by whom they are enforced. The framework should also extend governance to the design and operation cycles of AI technologies. A strong culture of compliance should

be a goal to ensure that AI safety risks are identified, remedied, and reported in a timely manner. A framework for monitoring organisational culture and AI ethics is required if progress is to be made in further understanding the impact of culture on board behaviour and vice versa.

4.6. THE ROLE OF THE BOARD IN AI ETHICS

The role of a board in AI ethics is to establish governance structure, processes, and relational mechanisms and to oversee, monitor, and enforce ethical standards and guidelines for the development and deployment of AI technologies. Accordingly, the board plays a key role in setting the ethical tone, setting strategy to managing ethics, articulating policies and codes of ethics to ensure that AI technologies respect privacy, and are fair, value-aligned, accountable, and transparent. King IV Code (2017) principles 1 to 3 state that (1) the board should lead ethically and effectively; (2) govern ethics in such a way that it establishes an ethical culture; and (3) ensure that the corporation is seen as a responsible corporate citizen. The King IV code further indicates that the board of directors should individually and collectively cultivate the ICRAFT (integrity, competence, responsibility, accountability, fairness, and transparency) characteristics in their conduct. The boards can effectively judge whether the company's AI strategy, plans, and performance are in alignment with the core values and ethical standards of the corporation if those standards are well articulated. Ethical standards should guide boards to consistently act on a fully informed basis, in good faith, with due diligence and care.

AI technologies are increasingly critical tools in advancing and supporting corporate strategy, but they carry risk as well. Thus, AI ethics oversight by the board is necessary. The board's oversight of AI technologies can drive the right tone and set the guiding principles for an ethical framework for AI adoption that can be operationalised to achieve responsible outcomes. Carpenter (2021) argues that by implementing an ethical framework for AI, corporations create a common language which should articulate the values, management system, and the measures that will sustain trust and ensure data integrity among all internal and external stakeholders. A common AI ethical framework and governance and management of risks associated with AI consistently across the corporation can allow for faster and more consistent and trustworthy adoption of AI.

Free (n.d.) suggests the following duties for the board in AI ethics

- Establish policies and procedures regarding AI ethics
- Provide thought leadership regarding AI ethics
- Audit use and proposed use of AI and data
- Examine proposals for AI research
- Set up a training programme regarding AI ethics
- Handle any complaints regarding use of AI and/or associated data
- Make decisions about how AI is used/and or researched
- Maintain an inventory of AI used by or being researched.

The board should also have an AI technology code of ethics which spells out the principles and provide the motivation that drives appropriate behaviour. According to Lawton & Wigmore (2021) a good code of ethics should have the following tenets:

- An ethical AI technology should be inclusive, explainable, have a positive purpose and use data responsibly.
- An inclusive AI is unbiased and works equally well across all spectra of society.
- An explainable AI supports the governance to ensure the ethical use of AI.
- An AI system endowed with a positive purpose aims to, for example, reduce fraud, eliminate waste, reward people, slow climate change, cure disease, etc. Any technology can be used for doing harm, but it is imperative that we think of ways to safeguard AI from being exploited for bad purposes.
- An AI that uses data responsibly observes data privacy rights. Responsible collection, management and use of data is essential to creating an AI system that can be trusted.

Lastly, the board should ensure impact assessment and monitoring the use of AI technologies, a framework alone does not guarantee ethical AI practices. Impact and risk assessment is "a type of fact-finding and evaluation that precedes or accompanies research, or the production of artefacts and systems, according to specified criteria. Assessing the impact of some X upon some Y has been practiced for generations, and has engendered debates over methods, purpose, focus, policy relevance, terminology, and efficacy" (Raab, 2020). Renn (2017) argues that these "assessments are designed by notions of relevance (what is important to society and which phenomena are worthy of attention), evidence (identification of causes and effects), and normative claims (what is good, acceptable or tolerable)".

The UNESCO Recommendation on the Ethics of AI (2021), advocates for the introduction of AI ethics impact assessment frameworks, which should be used to identify and assess benefits and risk posed by of AI technologies. The framework should also prescribe mitigation and monitoring mechanisms. By proactively identifying and eliminating potential risks to public safety and mitigating against societal risks of AI technologies, developers and operators can build the public's confidence. AI ethics assessment must address the full diversity of individuals and groups that maybe impacted by AI technologies, particularly since developers and operators might be unintentionally blind to potential risks or impacts (Falco et al., 2021). Assessments also serve to identify design trade-offs, which reveal the strengths and weaknesses of alternatives (Fischer, 2017).

Accordingly, boards should develop due diligence and oversight mechanisms to identify, prevent, and mitigate the impact of AI technologies systems in society as a whole. There are already well-established governance methodologies for assessing and mitigating risks and impact of new technologies such as information privacy (Clarke, 2009), data protection i.e., POPI Act, and environment impact (Morgan, 2012), etc. which could be integrated into the impact assessment frameworks as a suggested by UNESCO.

An AI ethics assessment framework requires a multidisciplinary approach both for indicators and results analysis that implement quantitative and qualitive insights into AI technology design, development, and implementation. For example, think of a safety model which is used in the aviation industry to the extent to which collaboration across manufacturers, operators, employees, regulators, and researchers work together to identify and address potential safety risks. A multidisciplinary approach in assessment of AI ethics will also offer a combined assurance is achieved by board. Combined assurance relates to coordinating management of both internal and external assurance providers, increased collaboration, and developing a holistic view of the organisations ethical risks. Combined assurance incorporates and optimises all assurance services and functions which enable an effective control environment; supports the integrity of information used for internal decision-making by management, the board and its committees; and supports the integrity of the corporations' external reports. Thus, such an approach will ensure that AI ethics is assessed at all levels of the corporation and the board will be able to have a holistic view of the impact assessment of AI technologies.

4.7. CONCLUSION

The chapter focused on a broad discussion of ethics (metaethics, normative ethics and applied ethics) and concludes, based on the literature review, that AI ethics is a sub-discipline of the field of applied ethics, which is concerned with two things: the moral behaviour of humans that create AI technologies and the impact of AI technologies on society, which both relate to the governance of these technologies. The chapter further highlighted the most common AI Ethics principles, i.e., accountability and responsibility, fairness, transparency, explainability, and privacy, and agreed that principles can be used to guide the development of governance practices, international standards, and further regulation. The chapter further discussed organisational ethics and the alignment of organisational cultural norms with AI technologies. The chapter concludes by arguing that the board of directors sets the tone for any corporate ethics and compliance program. Therefore, it is the responsibility of the board and executive management to ensure that technological advancements in AI align with the organisation's ethical culture.

CHAPTER 5: ARTIFICIAL INTELLIGENCE GOVERNANCE

5.1.   INTRODUCTION

As corporations introduce AI technologies as part of their product offering and business operation to improve efficiency, the increase in corporate governance scandals, such as the Enrol case in the early 2000s, has spotlighted the board's role in governing corporations. However, this role in the context of AI technology has also resulted in new corporate governance scandals, i.e., self-driving car accidents, face ID hacks, deep fakes, police unfairly targeting certain neighbourhoods, bias in facial recognition systems, surveillance practices limiting privacy, etc. (Dilmegani, 2022). This situation has necessitated some corporations, governments, and non-governmental organisations to seriously reflect on the complex and often alarming ways in which the use of AI technologies affects organisational and social life.

Mikalef, Conboy, Lundström & Popovič (2022) argue that there has been growing, fractious tension between AI technological capabilities and the human social structures within which the technologies reside. This tension questions the true extent to which these AI technologies make our lives better and happier, while it's 'dark side' is almost certainly making humankind more unhealthy, sadder and exhausted (Conboy, 2019; Tarafdar et al., 2015). Thus, some governments (Singapore[24]), and public (WE forum[25]) and private sector organisations (Amazon[26], Google[27], IBM[28], and Microsoft[29]) have moved towards the adoption of 'responsible AI'. 'Responsible AI' is a term coined to signify AI technologies that are governed in such a way that they are as fair and non-biased as possible, transparent

---

[24] Singapore's Model AI Governance Framework https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI-Governance-Framework#:~:text=Model%20AI%20Governance%20Framework&text=By%20explaining%20how%20AI%20systems,understanding%20and%20trust%20in%20technologies.

[25] AI Governance A Holistic Approach to Implement Ethics into AI https://weforum.my.salesforce.com/sfc/p/#b0000000GycE/a/0X000000cPl1/i.8ZWL2HIR_kAnvckyqVA.nVVgrWIS4LCM1ueGy.gBc

[26] Building AI responsibly at AWS https://aws.amazon.com/machine-learning/responsible-machine-learning/

[27] Google Responsible AI practices https://ai.google/responsibilities/responsible-ai-practices/

[28] IBM watsonx governance™ toolkit for AI governance https://www.ibm.com/watson/trustworthy-ai

[29] Microsoft Empowering responsible AI practices https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimaryr6

and explainable, secure, safe, accountable, and to the benefit of society as a whole (de Laat, 2021).

This chapter is an in-depth discussion on AI governance as a subset of corporate and IT governance. The chapter further discusses the different IT frameworks and dives into AI governance as part of organisational corporate governance structure, including AI governance frameworks that boards can utilise to govern AI technologies.

5.2. AI GOVERNANCE

AI governance studies focus on how best society can navigate the transition to advance AI technologies with a specific focus on governance and ethical dimensions (Dafoe, 2018). AI governance is about oversight and monitoring algorithms for effectiveness, risk, bias, and return on investment (Taulli, 2020). Basis AI (2020) defines AI governance as a framework and process that guides the design, development, and deployment of AI technologies that are explainable, transparent, and ethical. AI governance is the process of defining policies and establishing accountability structures to guide the design, application, and deployment of AI technologies in an organisation (Hind, Tucker, & Vaidyanathan, 2021). Mäntymäki et al., (2022) define AI governance as follows:

AI governance is a system of rules, practices, processes, and technological tools that are employed to ensure an organisation's use of AI technologies aligns with the organisation's strategies, objectives, and values; fulfils legal requirements, and meets principles of ethical AI followed by the organisation.

The definition of AI governance by Mäntymäki et al. (2022) is normative in that the intention is action-oriented with a goal to guide organisations in implementing effective AI governance. In particular, the definition draws on that of AI governance for business while incorporating regulatory constraints specific to AI technology and ethical AI principles. On the other hand, Schneider et al (2020) define AI governance for business as "the structure of rules, practices, and processes used to ensure that the organisation's AI technology sustains and extends the organisation's strategies and objectives". Their definition conceptualises the scope of AI governance to include the ML model, the data used by the model, and the AI technologies that contain the ML model. Mäntymäki et al. (2022) contend that even though AI governance for businesses is a promising starting point, the concept

largely omits ethical and regulatory questions present in previous AI governance literature. Therefore, the concept contrasts with AI ethics literature and downplays established AI-specific ethical and regulatory issues stemming from the organisation's environment.

Nonetheless, there is a general consensus on AI governance being about AI technologies that are 'explainable, transparent, and ethical[30] as many scholars allude in their definitions. However, those three words (explainable, transparent, and ethical) mean different things to different corporations or functions within corporations, which results in slightly different definitions or descriptions of what AI governance is (Lapach, 2022). Lapach (2022) argues that there is a difference between a technical and legal definition of "transparency", wherein a technical definition includes details on the algorithm codes and individual data components, which a policy definition might not. Equally, "explainability" and "transparency" can include the ability to measure cost, value, auditability, and compliance. Some definitions, like the Singaporean government's version, make it explicit that AI technologies need to be human-centric, reducing the universe in which the AI governance serves.

According to Butcher & Beridze (2019), AI governance should be characterised as a "variety of tools, solutions, and levers that influence AI development and applications". They further argue that AI governance should include the promotion of norms, ethics and value frameworks; studies of the effects and societal implications of AI technologies and raising awareness for stakeholders; that algorithmic interpretability and explainability should be ensured; and compliance with regulatory and legislative frameworks which govern AI-related technologies and fields should be monitored (Butcher & Beridze, 2019). Hind, Tucker & Vaidyanathan (2021) argue in their turn, that AI governance empowers corporations to operate with agility and complete trust when implemented correctly. As corporations develop and deploy AI technologies to automate existing or new business workflows to gain a competitive advantage, AI governance enables them to trust AI-powered outcomes at every step of the way to create trustworthy business outcomes.

AI governance combines traditional governance paradigms (policy, accountability, transparency, etc.) with differential ones such as ethics review, bias testing, and surveillance. AI governance aims to close the gap between accountability and ethics in

---

[30] AI Ethics principles as explained in Chapter 3

technological advancement. As AI technologies are involved in making decisions that impact on society, AI governance becomes a necessity. The design, development, deployment and use of AI systems should be in accordance with the corporation's strategy, consumer expectations, and societal norms. AI governance should be aligned and integrated with other business units to ensure compliance and risk management. In reality, AI technologies are rarely coordinated across business units, and data science teams are often isolated from design, application and deployment (e.g., Damianides 2005, Torre, Teigland, & Engstam 2019). AI governance is the responsibility of the board to ensure that the corporation follows all the processes and structures required for incorporating AI technologies into its business operations (Thuraisingham, 2020) ensuring trustworthy processes.

5.3. WHY AI GOVERNANCE IS IMPORTANT

Understanding and managing the risks posed by AI technologies is critical to realise the benefits of the technology. Responsible AI governance is important to ensure that AI technologies are developed, deployed and implemented ethically, transparently, and in the best interest of society. Thus, it serves as a guardrail, ensuring that AI technologies adhere to ethical principles, are transparent, and prioritise the well-being of individuals and society. As AI technologies develop and increase in complexity, their risks and interconnectivity with other smart devices and systems will also increase, necessitating the creation of both specific governance mechanisms, such as for healthcare, transport and autonomous weapons, as well as a broader global governance framework for AI technologies (Butcher & Beridze, 2019).

AI technologies introduce a new issue or theme to the world of governance for corporations in terms of overseeing technologies that can learn and independently make complex decisions but are vulnerable to human weaknesses such as bias and criminality. Boards need to ensure processes and decision-making structures are at the heart of the life cycles of AI technologies and make decisions about transparency, fairness, and accountability. At the same time, AI technologies raise the importance of data governance because training machine-learning systems requires enormous datasets that should be integrated, representative, and unbiased. Therefore, it is important for boards also to consider governance processes that can address the questions AI brings to strategy, risk, and other control issues.

5.3.1. **AI governance as part of organisational corporate governance structure**

AI governance is a subset of corporate governance and IT governance (Mäntymäki et al., 2022). AI governance has not been explicitly connected with corporate governance and governance beyond adapting definitions from these established fields to cover AI (Schneider, Abraham, & Meske, 2020). Mäntymäki et al (2022) argue that locating AI governance in the corporate governance structure is important to ensure that AI technologies build on other governance areas and avoid duplicating processes. Figure 7 below highlights the relationship of AI governance (AIG) with corporate governance (CG) and IT governance (ITG). Corporate governance ensures that structures and processes are put in place to mitigate the agency problem with the corporation, and AI technologies are governed via mechanisms that fall under IT governance. More explanation of these relationships in the next section.
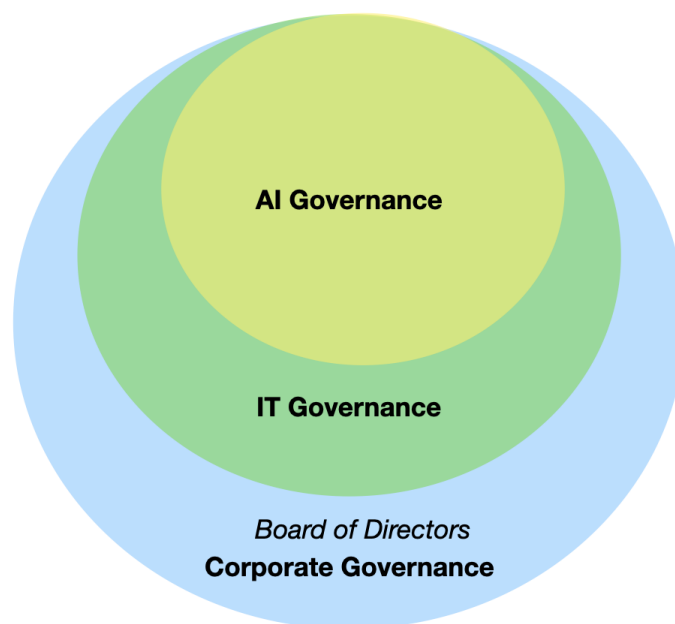
Figure 7: AI Governance as part of corporate governance structures (Mäntymäki et al (2022)

5.3.2. *Corporate Governance*

One of the most challenging issues confronting corporate governance, particularly the board of directors in the near future, will be its ability to exercise informed oversight over the application of AI technologies within corporations (Peregrine, 2022). The challenge will be across the board regardless of the industry sector in which the company operates and how it applies AI in that operation. The essence of the challenge is the rapidly emerging conflict

between the perceived societal and commercial benefits arising from AI implementation and the perceived societal and institutional risks arising from its use.

As already argued in Chapter 3 of this study, corporate governance[31] is a process in which corporations are controlled, supervised, and managed. The board, the custodian of corporate governance, is responsible for the strategic direction in responding to the ethical issues that AI technologies may generate. As depicted in figure 7 above, AI governance does not occur in isolation; it is part of a corporation's overall corporate governance structures. Corporate governance sets parameters for how the corporation operates with its internal and external stakeholders (Solomon, 2020; Sahut, Dana & Teulon, 2021), regulating the process of interactions with its environment (Schneider, Wickert, & Marti, 2017).

The main purpose of corporate governance (CG) is to establish governance and monitoring mechanisms to resolve agency problems to protect the interests of shareholders (Jensen and Meckling, 1976) and to maximise shareholder wealth by reducing agency loss (Adegbite et al., 2012; Bonazzi and Sardar, 2007; Cohen et al., 2002). As explained in chapter 2, CG act as a mechanism wherein the board (board of directors) is a monitoring tool to reduce the principal-agent relationship. This means that the board is required to establish governance structures through which the objectives of the organisation and the mechanisms for monitoring performance are set. In this context, CG allows the corporation to articulate rules and processes for governing the relationships between its management and shareholders, managing the potential tensions between its shareholders and other stakeholders, and controlling its environmental and social impact.

As a focal point of governance, CG is composed of several governance frameworks such as risk governance, management governance, financial governance, data governance, and IT governance. The alignment of all these sub-governance frameworks is essential, especially to close the gap where functions overlap (Wessels & Loggerenberg, 2006). For example, assets governance in the organisation will typically include computer hardware and software. Such assets (computer hardware and software) are also functions of IT governance and or financial governance wherein software, or algorithms, are used to

---

[31] The OECD (2004) defines corporate governance as; *"a set of relationships between a company's management, board, shareholders and other stakeholders. Corporate governance also provides the structure through which the company's objectives are set, and the means of attaining those objectives and monitoring performance are determined"*.

calculate or predict future investments. An alignment of these frameworks will ensure that all sub-governance frameworks are synced towards a common organisational goal. Therefore, AI governance would only be implemented successfully if there is an alignment with CG and IT governance, guided by precise, well defined, adaptable internal policies and procedures.

5.3.3. IT governance

As corporations increasingly depend on IT for operations and recognise it as a strategic partner, the need for IT Governance intensifies. The importance of IT governance and that it needs to be an integral part of overall CG came into light as a result of failures in IT infrastructure or systems. According to an ITGI study in 2008, the main two motivators for implementing ITG within an enterprise are risk management and value delivery. IT governance is an essential part of overall corporate governance that helps to ensure the alignment of IT decisions with business goals and ensure that IT investments deliver added value for the business (ITGI, 2009). Exler (2003) argues that IT governance should not be implemented outside corporate governance frameworks. The study conducted by Parker, Peterson & Ribbers (2002) found that a lack of business alignment to IT results in inconsistency and misinterpretations in IT governance implementation. Thus, it is crucial to alight business management and IT governance within the corporate governance frameworks.

In recent years, most corporations have come to process information through IT-related platforms (Anand, Manz, & Glick, 1998). IT governance is then a means to execute corporate governance and a key element of a corporation's governance structure. IT governance is a critical component of CG, and it becomes even more crucial when corporations implement emerging technologies, and AI is amongst the various emerging technologies that are currently transforming business processes (Witt, Wolanske & Merhout, 2021). AI technology is built through algorithms written in the form of software code; these algorithms can be stand-alone applications or integrated into a corporation's information systems (IS) to provide additional capabilities or increased performance. As a result, AI governance at the organisational level can be viewed as a subset of IT governance (Mäntymäki et al., 2022), which deals mainly with IT-related aspects within a CG framework (Kan, 2003; Carroll, Ridley & Young, 2004).

Corporations are spending huge investments on IT projects to access new markets and to gain a competitive advantage. Thus, most corporations have realised the importance of proper management of IT investments and operations along with its associated risks to make the best use of IT. These IT projects carry considerable risks related to data security, compliance, information security, IT system risk, financial capacity risk, cost risk, etc.

There are various definitions of IT governance, but there is still no generally accepted definition (Brown & Grant, 2005). The IT Governance Institute (2009) and De Haes & Van Grembergen (2015) define ITG as:

"IT governance is an integral part of enterprise governance and consists of the leadership and organisational structures and processes that ensure that the organisation's IT sustains and extends the organisation's strategies and objectives" (ITGI, 2009).

*"IT governance an integral part of CG, exercised by the Board, overseeing the definition and implementation of processes, structures and relational mechanism in the organisation that enables both business and IT people to execute their responsibilities in support of business/IT alignment and the creation of business value from IT-enabled business investment" (*De Haes & Van Grembergen, 2015).

Both the definition by ITGI (2009*)* and De Haes & Van Grembergen (2015) identify three main pillars of IT governance in corporations: "decision-making structures, processes and relational mechanisms" (Peterson 2004). The first pillar focuses on structures (leadership), which refers to the board's role in ensuring IT governance structures are in place and must be aligned with the corporation's strategies and objectives. The second pillar is processes, which point to the approach to decision-making, monitoring, and evaluation of AI activities. Lastly, the relational mechanism deals with the framework to promote all business units' communication, cooperation, and collaboration in planning, developing, and executing AI technologies to ensure desirable IT-related activities and outcomes.

Weill and Ross (2004) define IT governance as "specifying the decision rights and accountability framework to encourage desirable behaviour in using IT". Their definition adds a new aspect to ITG, specifying decision rights, roles, and responsibilities that seems to indicate they agree with Schwartz (2009) on the need to assess IT performance. On the other hand, Sambamurthy and Zmud (1999) define ITG as "a distribution authority for key

IT activities in business firms, including IT infrastructure, IT use, and project management". Sambamurthy and Zmud's definition does not include the strategy and IT business alignment aspects and concentrate more on the distribution of authority and decision making.

The definition by De Haes & Van Grembergen (2015) also acknowledges IT governance can only be achieved through strategic alignment. Strategic alignment refers to the state of congruence between business strategy and IT strategy to support the overall business purpose that influences the firm's business performance in a corporation (Luftman 2000 & 2005; Reich and Benbasat, 2000). Weill, P. & Ross, J.W. (2004) argue that effective IT governance is responsible for distributing the decision-making to the best-qualified individuals and to ensure that they are privy to the needs of the organisation and the implications of these needs. Another approach to IT governance is presented through several IT governance frameworks that have been suggested (such as CoBIT and ITIL) to help IT managers structure and formalise the IT management process for holistic management of IT governance processes (Van Grembergen, 2007; De Haes and Van Grembergen, 2004).

5.3.3.1.   IT governance frameworks

Accordingly, the IT governance framework assists the board in understanding the strategic importance of IT investment and assists corporations in sustaining operations and implementing strategies required to expand the business activities into the future. It provides assurance to the board that IT investment expectations are met, and IT risks are mitigated (Levstek, Hovelja, & Pucihar, 2018).

In recent years, a considerable number of IT governance frameworks have been developed, such as ISO 38500 (ISO/IEC, 2008), COBIT (IT Governance Institute, 2012), ITIL (GITIM, 1989), IT balance scorecard (Van Grembergen & De Haes, 2005), etc.

- **ISO 38500.** ISO 38500 is an international IT governance standard that provides guidance to boards on the effective, efficient and acceptable use of IT assets. This framework considers IT governance as an integral part of corporate governance and includes "strategies, policies, decision-making structures, and accountabilities through which the organisation's governance arrangements operate" (ISO 38502,

2017). The framework assists the boards in fulfilling their legal, regulatory, and ethical responsibility (ISO/IEC, 2008).
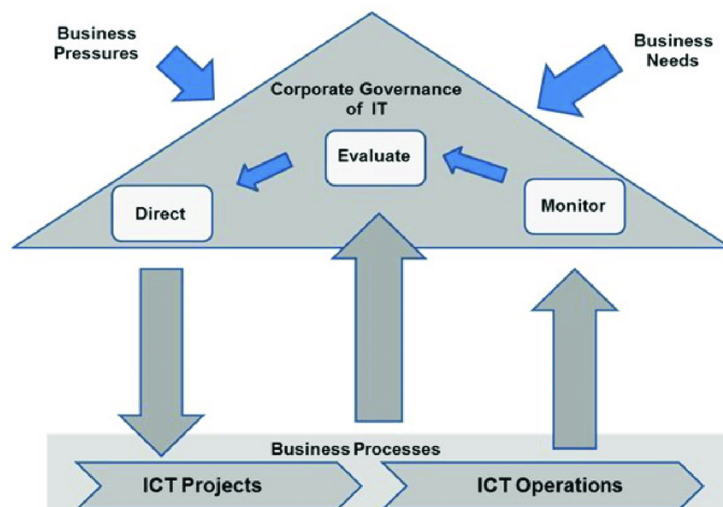
Figure 8: ISO 38500 Framework Source: ISO/IEC 38500

To ensure successful IT governance implementation, ISO recommends that the board manages IT through three main tasks; "(i) Evaluating the current and future use of IT. (ii) Directing the preparation and implementation of plans and policies to ensure that the use of IT is aligned with the organisational objectives and (iii) monitoring the conformance to policies and performance against the plans" (ISO/IEC, 2008).

- **COBIT**. COBIT is a widely used comprehensive framework which links business risks with IT Governance (Naidoo, 2009). The framework provides metrics and maturity models to assess how the above issues are managed and identify the associated responsibilities of business and IT process owners (Moeler, 2008; Marnewick & Labuschagne, 2011). COBIT provides a framework for board and management oversight, and control and monitoring of IT assets, with the view of aligning IT with organisation activities. COBIT considers IT governance to be an integral part of corporate governance and emphasises that the objective of governance is to create value (Mulamula, 2021). The COBIT framework allows for the creation of IT control policies as it constantly updates and harmonises other standards and guidance. Therefore, it is known to integrate good practices and frameworks encapsulating IT governance and helps comprehend and control the risks and benefits related to IT (Debreceny, 2006).

The newest version, COBIT 5, adopts five principles and places an increased focus on processes, enterprise goals, and enablers aligned with these principles. Principle 1 focuses on the stakeholder needs by maintaining a balance between the realisation of benefits and the optimisation of risk and use of resources. Principle 2 integrates governance of enterprise IT into enterprise governance, holistically covering all IT-related functions. Principle 3 encourages the application of a single, integrated framework, noting that in the modern era, there is a need not only to recognise, but also align with other international standards. Principle 4 advocates a holistic approach to IT by integrating principles, policies frameworks, processes, organisational structures, corporate culture, and the skills and competencies of people within the corporation. Finally, Principle 5 separates governance from management, distinguishing between the two (Theron & Koornhof, 2016). In this regard, it states that: "Governance ensures that stakeholder needs, conditions and options are evaluated to determine balanced, agreed-on enterprise objectives to be achieved; setting direction through prioritisation and decision making; and monitoring performance and compliance against agreed-on direction and objectives" (ISACA, 2012)

The framework ensures that the overall value creation is directly related to value delivery from digital transformation and mitigation of risks resulting from digital transformation. Specifically, the enterprise governance of IT is considered successful when it has three specific outcomes (ISACA, 2019): i) benefits realisation, ii) risk optimisation, and iii) resource optimisation. Boards use these three key outcomes to maintain an overall governance framework and manage stakeholders (ISACA, 2019). Per this framework, IT governance is primarily the board of directors' responsibility, under the leadership of a chairperson. The governance function evaluates strategic options, directs top management on the chosen strategic options, and monitors the achievement of the strategy (Mulamula, 2021). Using COBIT for AI development and monitoring means to maintain both guidelines and flexibility that encourage technological innovation and protect human values (Tuttle & Vandervelde, 2007; von Solms, 2005).

- **ITIL.** The ITIL framework was initially published by the Central Computer and Telecommunications Agency (CCTA), a UK government agency. According to

Hamidovic (2014), the framework has now been adopted and used across the globe as the de facto standard for best practices in providing IT services.

The ITL framework will help an organisation define its IT service, align services with customers' needs, and produce a clear model for ongoing success. It is used to improve core IT service processes and is flexible to each organisation's unique requirements. The framework is structured around IT assets' service lifecycle, including "service strategy, service design, service transition, service operation, and continuous service improvement" (Hamidovic, 2014).
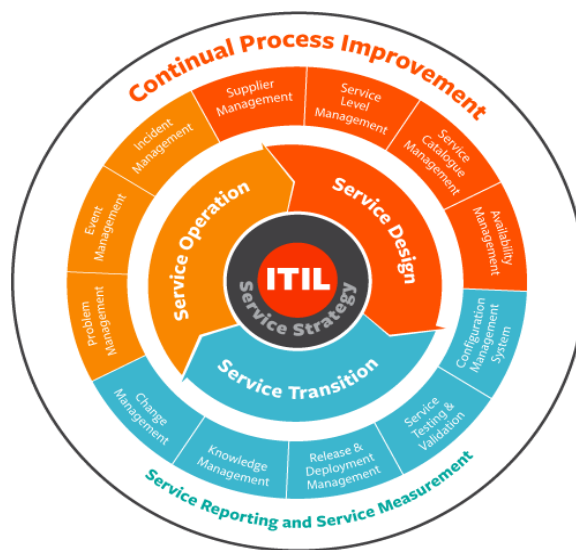
Figure 9: ITIL Framework version 4. Source (Ivanti, 2024)

ITIL version four (4) was introduced in 2019 and is the latest version of the ITIL Framework. ITIL version three (3) had streamlined IT service management processes and was aligned with international service management standard ISO 20000. The main element of the ITIL framework lifecycle, as depicted in figure 9, is the service strategy used to guide strategic objectives and expected value performance of IT assets or resources. Service design mainly ensures that the design of IT services is aligned to business goals. The scope of service design is not only to "design new IT services but also the processes of change and improvement of service quality, continuity of service or performance of service" (Suhairi & Gaol, 2013). The service transition phase of the lifecycle ensures that the organisation adapts to IT service design changes if there are changes in the operational environment.

ITIL version four has a renewed focus on customer experience and digital transformation. It engages with the reality of modern-day IT governance, such as cloud-based services or machine learning. Digital technology is rapidly changing, and ITIL 4 provides the tools and insight to react to and mitigate risk. The ITIL Framework helps organisations create an environment of efficiency and quality, delivering IT services at speed. Watts (2017) argues that the framework enables IT assets to be managed thought their lifecycle by ensuring IT alignment with business needs

Effective organisational IT governance comprises a single or combination of frameworks and standards. In actuality, each one is a formal set of practices that addresses specific objectives of IT governance (Othman, 2016). Thus, more scholars contend that the adoption of a formal framework produces measurable results towards achieving the organisations' strategic objectives or return on investment. The adopted frameworks should also be used to highlight an organisation's IT governance maturity levels (Tonelli, de Souza Bermejo, Dos Santos, Zuppo & Zambalde (2017).
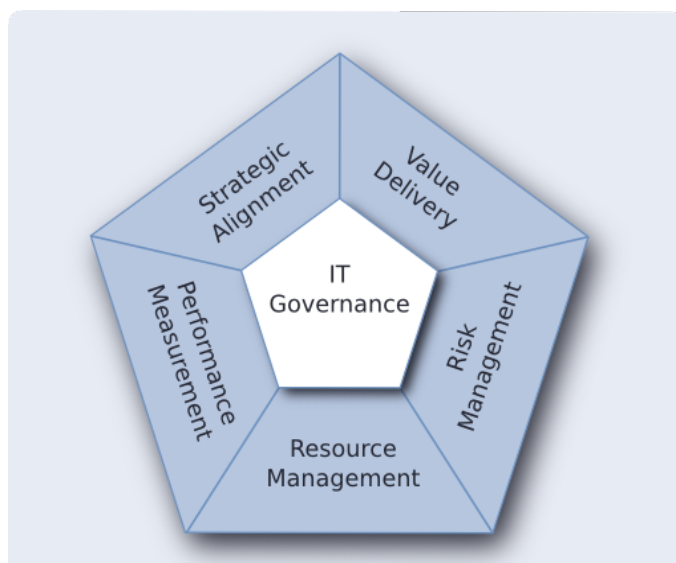
Figure 10: Five areas of IT governance. Source (ITGI, 2003)

At the board level, IT governance is concerned with ensuring that the right framework and processes support decision-making in the usage, investment, and security of IT assets. Efficient IT governance requires communication between the governance board and executive management. Scholars have identified five areas of IT governance that should be addressed by the board (Butler & Butler, 2010; Valentine & Stewart, 2015), as depicted in Figure 10.

- **Strategic alignment and value delivery.** The main objective of IT governance is to ensure alignment between the business units and IT. By creating the necessary structures and processes around IT investments, the board should ensure that IT projects are aligned with strategic business objectives to deliver value.
- **IT risk management and opportunities.** As corporations build on IT, risks associated with IT are often the same as risks to the business, thus, managing IT risk is vital. IT risks include cyber security, privacy risks, disaster recovery, business continuity risks resulting from systems outages, and the risks associated with project failures. However, the board should also leverage business opportunities resulting from IT investments.
- **IT resource management and oversight.** IT governance holds IT management accountable for the return on its investment in IT, and the credibility of IT's information and controls.
- **Monitoring the performance and delivery of IT investments.** The board should ensure the performance of IT through a balanced scorecard. The IT Balanced Scorecard consists of four perspectives: IT Value, User, Operational Excellence, and Future Orientation.
- **Value delivery focuses on optimising IT investment**. Delivering value to the business typically means growing revenues, improving customer satisfaction, increasing market share, reducing costs, and enabling new products and/or services.

IT governance is the board's ability to direct, evaluate, and monitor IT assets' utilisation in ensuring the achievement of strategic objectives (Haes, 2018). The board leadership, structures, and processes are used to leverage IT, which in this case include emerging technologies such as AI, to produce critical data and drive the alignment, delivery of value, management of risk, sustainability, and performance management (Butler & Butler, 2010; Valentine & Stewart, 2015). AI governance is the board's responsibility as the board has to ensure that corporations follow all the processes and methods required for incorporating AI into their business operations, including product development, manufacturing, and sales, to achieve corporations' strategic goals (Thuraisingham, 2020).

The King Code IV (2016) Corporate Governance for South Africa (Principle 12) deals specifically with IT governance and the board's role. It states that the board should govern IT to support organisations in setting and achieving strategic goals. IT governance offers valuable guides to operationalising AI governance. However, because of their learning

capabilities and dynamic evolution, AI technologies should initially be viewed as special cases of IT systems that require distinct governance tools compared with general IT governance. As AI governance matures, it may eventually be subsumed under IT governance.

5.4.   AI GOVERNANCE FRAMEWORKS

AI governance frameworks are sets of principles and guidelines designed to ensure responsible and ethical development, deployment, and use of AI technologies. These frameworks aim to address various aspects of AI, including fairness, transparency, accountability, safety, privacy, and the overall impact on society. These frameworks play a crucial role in mitigating the risks associated with deployment, development and use of AI technologies, while also maximising their benefits and opportunities, enabling safe implementation, and building stakeholder trust AI (Stogiannos N, Malik R, Kumar A, Barnes A, Pogose M, Harvey H, et al., 2023).

To date, there is insufficient direct scholarly work available on the development of holistic AI technologies governance frameworks that boards can utilise to govern AI technologies within corporate environments. However, much related work has been done in adjoining fields that can be leveraged as part of this study. For example, work done by the European Commission's high-level expert group on AI[32], the OECD expert group on AI in society[33], the Singapore advisory council on the ethical use of AI and Data[34], the UNESCO bio-ethics and ethics of science and technology team, and the United Kingdom select committee on AI. Companies such as Google and SAP have publicly released AI guidelines and principles. Non-profit organisations such as the Association of Computing Machinery (ACM), Access Now, and Amnesty International have issued declarations and recommendations on the responsible use of IA (Schiff et al., 2021). The Singaporean government released the Model Artificial Intelligence Governance Framework (2020). Furthermore, professional bodies, such as the IEEE, are playing core roles in this context. Unfortunately, these have not resulted in a global consensus on a single model for AI governance or regulation of the ethical and trust issues relevant to AI technologies.

---

[32] https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai
[33] https://www.oecd.org/innovation/oecd-creates-expert-group-to-foster-trust-in-artificial-intelligence.htm
[34] https://www.cms-holbornasia.law/en/sgh/publication/singapore-to-form-advisory-council-for-ethical-use-of-ai

The Singapore Model for Artificial Intelligence Governance Framework and Layered Model of AI Governance by Gasser & Almeida (2017) are important guidance frameworks toward responsible AI technologies deployment. The Singapore framework's objectives are to assist organisations in building stakeholder confidence in AI technologies through an organisation's responsible use of AI technologies to manage different risks in AI technologies' deployment as well as to demonstrate reasonable efforts to align internal policies, structures, and processes with relevant accountability-based practices in data management and protection.

The Layered Model of AI Governance (Gasser & Almeida, 2017) is focused on assisting governmental institutions in planning for a multi-year deployment of the governance of AI technologies. This model has three layers: (i) a technical layer (algorithms and data) focusing on data governance, algorithm accountability, and standards; (ii) an ethical layer deploying criteria and principles; and (iii) a social and legal layer focusing on norms, regulations, and legislation.

The World Economic Forum published a toolkit in January 2020 for boards of directors called "Empowering AI Leadership" (World Economic Forum, 2020). This toolkit focuses on strategy (including strategies for brand, competition, customers, operations, and technology), audit processes, cybersecurity, ethics, governance, people and culture, responsibility, risk, and sustainable development. Furthermore, there are institutes/government bodies focusing on setting up AI ethical frameworks or AI principles for organisations to follow.

AI governance frameworks provide oversight and supervision while entrusting the technical teams with core technical aspects, such as data engineering, building models, and choosing best-fit AI platforms. AI governance frameworks assist corporations to learn, govern, monitor, and mature AI adoption. AI governance ensures AI technologies' alignment with the corporate strategic objective.

5.5.   ROLE OF THE BOARD IN AI GOVERNANCE

AI technologies introduce a new issue to the world of governance, i.e., how to oversee systems that can learn and independently make complex decisions but are vulnerable to

human weaknesses such as bias. At the same time, AI technologies raise the importance of data governance because ML training requires enormous datasets and workforce governance (World Economic Forum, 2020). Boards must also consider whether today's governance processes and accountabilities are able to address the questions AI brings to strategy, risk, and other control issues.

The board needs to implement effective oversight of AI technologies. This implies that the board should be knowledgeable enough about AI technologies and have a perspective on critical issues such as AI ethics, risk, and governance. The board should keep abreast of the impacts (and potential) of AI technologies to avoid shocks to the corporations and to be open to the impacts that may reasonably be expected to arise during the course of executing their duties as directors. Therefore, boards have a responsibility to ensure that corporations have effective and efficient governance structures for setting objectives and monitoring performance. That structure must provide for how the board, management, shareholders, and other stakeholders participate in decision-making, and ensure the board has the information required to provide oversight, and that shareholders have information about the strategy, risks, and performance of the corporation.

Effectively, boards are also responsible for making sure executives' implementation of policies, actions, and plans involving the use of AI technologies are responsible and consistent with the corporations' ethical standards and legal obligations. At the strategic level, it is important for the board to prioritise AI ethics and allocate sufficient budget and resources. Corporations typically comprehend the budget appropriations, technology procurement and hiring, but so far, they are not proficient in translating AI ethics principles into action items (Kompella, 2022). Many boards are uncomfortable with the responsibility for overseeing AI risk because of their lack of expertise in this area. However, boards have fiduciary duties (also see chapter 2) to exercise their supervision obligations, even in technical areas, if those areas present enterprise risk, which is already true for AI at some corporations (Gesser, Regner & Gressel, 2022). That does not imply that board members must become AI technology experts or that they should be involved in day-to-day AI operations or risk management. But the board should consider how they will ensure effective oversight with respect to the growing opportunities and risks presented by AI technologies.

Principle 12 of the King IV report (2017), which the listed companies in South Africa have adopted, acknowledges that digital development such as robotics and artificial intelligence,

nano-technologies, genetics, and bio-technologies, have accelerated technology from being an operational enabler to leading the transformation of product, service, and business models. The report further indicates that technology governance and security have become critical issues and are now a source of many corporations' future opportunities and of potential disruption. Thus, the report recommends the following as the role of the board as it relates to information (data) and technology: (i) the board should assume responsibility for setting strategic direction on how technology and information should be governed; (ii) give approval of policies and frameworks that give effect to the set direction on the deployment of technology and information; (iii) delegate to management the responsibility to implement technology and information; and (iv) the board should exercise ongoing oversight of technology and information, in particular, oversee that it results in the following:

- "Integration of people, technology, information and processes across the organisation.
- Integration of technology and information risk into organisation-wide risk management.
- Arrangement to provide for business resilience.
- Proactive monitoring of intelligence to identify and respond to incidents, including cyber-attacks and adverse social media events.
- Management of the performance of, and the risk pertaining to, third-party and outsourced service providers.
- The assessment of value delivery to the organisation through significant investment in technology and information, including the evaluation of projects throughout their life cycle and of significant operational expenditure.
- Ethical and responsible use of the technology and information.
- Compliance with relevant laws.
- The protection of privacy of personal information" (King IV Report, 2017).

Eccles & Vogel (2022) suggest that the board should ensure responsible AI governance. Firstly, the board needs to establish an AI governance framework with clear timelines for long-term planning and testing algorithms. Secondly, the board has to identify a point of contact at the executive level and delegate responsibility to implement AI governance. Thirdly, the board has to designate (and communicate) the stages of the AI lifecycle when testing will be conducted (e.g., pre-design, design and development, deployment). Fourthly, the board should document relevant findings after each stage to promote consistency,

accountability, and transparency. Finally, the board should ensure the implementation of routine auditing.

The board thus has a responsibility to establish governance and monitoring structures to ensure effective AI governance. The World Economic Forum (2020) AI board toolkit on empowering leadership suggests that boards use the following to examine if adequate governance structures are in place for AI:

- **Board oversight** – clarifying which board committee(s) provide oversight, which AI activities they oversee, and whether AI necessitates establishing a new board committee.
- **Ethics board** – deciding on the responsibilities of the ethics board overseeing AI, how it will maintain its independence and selecting and confirming its members.
- **Risk assessment** – identifying AI-specific risks that should be considered while providing AI governance.
- **Monitoring, audit and response** – evaluating the effectiveness of compliance and ethics assurance programmes and the response to allegations of violations.
- **Training** – providing education for employees, contractors and other third parties about compliance risks and programmes.
- **Reporting to shareholders** – as required by law, ensuring reports to shareholders and filings to regulators include information about the risks of AI. They should also report on the use of AI in financial reporting and auditing to verify the accuracy of financial statements.

The board's responsibility in AI governance includes having AI as a periodic board agenda item. Gesser, Regner & Gressel (2022) suggest that, as with Environmental, social, and corporate governance (ESG) oversight over AI technologies should reside with the full board, an existing committee (*e.g.*, audit or technology), or a newly formed committee dedicated to AI technologies. Suppose the board is concerned that it does not have the necessary expertise to oversee AI opportunities and risks. In that case, it should consider adding one or more directors with that experience or have some board members receive AI training. Furthermore, they suggest that boards should ensure that there are effective management-level AI compliance and reporting structures to facilitate board oversight, which includes continuous AI technologies risk assessments and monitoring of high-risk AI systems and written AI policies, procedures, and training. Such policies may include

procedures for responding to a material AI-related incident, responding to AI-related whistle-blower complaints, and risk management for any vendors that supply the corporation with critical AI-related resources (Gesser, Regner & Gressel, 2022).

Board oversight must include requirements for corporate and use-case level AI policies. These policies should set out where AI technologies will be used, what checks and balances will be put in place, as well as setting out standards for robust and safe operation. Additionally, these policies must be underscored by practical processes, compliance structures and strong culture. Such should also include the publishing in the integrated report on how the company governs AI ethics openly and transparently. This is in line with the recommendations by the World Benchmarking Alliance (2021) that corporations (i) should be transparent about how they are using AI technologies and should ensure that their AI are explainable and accountable to both their stakeholders and the wider public; (ii) be proactive in promoting the responsible development and use of AI technologies, and should engage with relevant stakeholders and the wider community, to ensure that the impact of their AI systems is positive; (iii) support the development of ethical frameworks and guidelines for the development and use of AI, and should ensure that their AI systems are designed and operated in accordance with these frameworks.

5.6.   CONCLUSION

In this chapter, the researcher discussed the relationship between AI governance, corporate governance, and IT governance. By definition, AI governance closes the gap between accountability and ethics in technological advancement, and the chapter highlights the general consensus that AI governance is about AI technologies that are explainable, transparent, free from bias and ethical, as many scholars allude in their definitions. The chapter further highlighted the importance of AI governance in managing the risks posed by AI technologies. The chapter highlighted the different IT frameworks and AI governance approaches as part of organisational corporate governance structure, including AI governance frameworks that boards can utilise to govern AI technologies. The researcher concluded that the boards should ensure that there are effective management-level AI compliance and reporting structures to facilitate board oversight, which includes continuous AI technologies risk assessments and monitoring of high-risk AI systems and written AI policies, procedures, and training.

CHAPTER 6: RESEARCH APPROACH AND METHODOLOGY

6.1. CHAPTER INTRODUCTION

The methodology chapter presents the reader with the opportunity to discover the chosen methods and design which the researcher used in achieving the study's envisioned objectives. The research methodology chapter addresses the following aspects of the study:

- Paradigm and philosophical perspectives
- Qualitative research methods and rational for choosing a phenomenological research approach
- Research settings
- Primary and secondary data
- Sampling
- Data analysis
- Ethical considerations

The chapter further provides a detailed description of the research approach, tools and applications, primary and secondary data collection, and data analysis processes methods selected for this study. The sampling procedures, validity and reliability, ethical consideration, limitation, feasibility, and positionality of findings will further be discussed in the chapter.

6.2. PARADIGM AND PHILOSOPHICAL PERSPECTIVE

The philosophical perspective is based upon a consideration of ontology (concepts of what constitute reality), axiology (what we believe is truth) and epistemology (knowledge). Philosophically, reality is either portrayed as totally separate from, or as a construction of, the mind. Truth may be interpreted as reflections of reality based on evidence which is determined by an understanding of reality at any given point in time. Knowledge is a means of reflecting reality and truth, involving what can be proven using concrete data as well as abstract comprehensions of phenomena; and theory which provides ways of explaining or giving meaning to understandings extrapolated from data (Howell, 2013). Every researcher might have his or her own view on what constitutes reality, truth, and knowledge (Wagner,

Kawulich, Garner, 2012). These views guide our thinking, our beliefs, and our assumptions, and they frame how we view the world around us.

This research follows a constructivist / interpretivist paradigm of inquiry using a phenomenological research approach, which has its roots in pragmatism and symbolic interactionism. Constructivism and interpretivism are related concepts that address understanding the world as others experience it. The philosophical underpinning to the constructivist / interpretivist paradigm is informed by hermeneutics and phenomenology, unlike positivism which is informed by realism, idealism, and critical realism (Wagner, Kawulich, Garner, 2012). Hermeneutics is the study of interpretation, which concerns the meaning of human intentions, beliefs, and actions, or the meaning of human experience (George, 2020). Phenomenology is the study that is based on lived experiences and describes the meaning for individuals of a concept or a phenomenon (Creswell, 2009). It is a study of structures of consciousness as experienced from the first-person point of view. It recognises that the centrality of experience is its intentionality, which is directed toward an object by virtue of its content or meaning (which represents the object) together with appropriate enabling conditions (Smith, 2013). Langdridge (2007) argues that phenomenology can be considered as an umbrella term which encompasses both a philosophical movement and a range of research approaches. It is a discipline which "aims to focus on people's perceptions of the world in which they live and what it means to them; a focus on people's lived experience". A phenomenological approach attempts to understand people's perceptions and perspectives (Leedy & Ormorod, 2015).

6.3.   QUALITATIVE RESEARCH

To realise the objectives of the study, which is to explore how boards perform their functions over governance of AI technologies ethics, the researcher used a qualitative research approach. The qualitative interpretation was executed within the discussion of corporate governance and AI ethics. As such, corporate governance and AI ethics research requires more qualitative investigation as "qualitative research can assist policymakers and practitioners to develop more efficient governance mechanisms, by shedding light on the efficacy of policy prescription. Qualitative research provides a basis for rethinking and challenging some of the dominant assumptions and meanings about how governance actors and institutions actually function" (McNulty, Zattoni, & Douglas, 2013).

Qualitative research is the systematic inquiry into social phenomena in natural settings. These phenomena can include, but are not limited to, how people experience aspects of their lives, how individuals and/or groups behave, how organizations function, and how interactions shape relationships. In qualitative research, the researcher is the main data collection instrument (Bogdan & Biklen, 1997; Corbin & Strauss, 2014). Qualitative research involves collecting, analysing, and interpreting non-numerical data. Qualitative research is mainly to understand how an individual subjectively perceives and gives meaning to their social reality.

The motivation for following a qualitative approach is that the domain of AI ethics and governance, particularly reflection on the role of the boards in AI technologies ethics, is still in its early stages. Board governance of AI ethics is a relatively novel phenomenon that requires careful consideration of the ethical implications of AI technologies. As suggested by Creswell & Poth (2018), Kohler (2016), and Leedy & Ormrod (2015), when there is an inadequate theory to guide the research, a qualitative methodology is often the preferred approach for exploring complex ethical issues related to AI. This is because qualitative research allows for in-depth exploration of the nuances and complexities of ethical dilemmas, and can provide rich insights into the social and cultural contexts in which AI technologies are developed and deployed. Most of the knowledge base related to the role of the board in governance of AI ethics is within the minds of experts in the field of corporate governance attempting to exercise governance in this domain, and to some extent, within the minds of AI ethics policy makers. A qualitative methodology allows for in-depth data to be gathered from such key experts, along with related contextual information and associated nuances (Creswell & Poth, 2018; Kohler, 2016).

In combination with the qualitative research undertaken, a critical review of the literature in AI ethics, AI governance, and corporate governance was undertaken to deep-dive into the role the boards play in the ethics governance of AI technologies. Furthermore, general norms of traditional philosophical research methodology were adhered to. This includes the definition and explanation of concepts, philosophical argumentation, the critical analysis and assessment of theoretical frameworks, the identification and critique of presuppositions, and the critical-creative reinterpretation of contemporary literature on the research problem. This work, combined with the interpretation of the qualitative research results, will guide recommendations in chapter 8.

6.4. PHENOMENOLOGICAL METHODOLOGY

This research is designed within the scope of qualitative design. The design uses a phenomenological methodology (Creswell, 2003). The phenomenological research approach is grounded in people's experiences of social reality (Gray, 2004) and makes meaning of their experiences (Cridland et al., 2014; Meadan, Stoner & Angel, 2010). "Phenomenology serves as the rationale behind efforts to understand individuals by entering into their field of perception to see life as these individuals see it" (Bruyn, 1966). Phenomenology is the study of lived experience (Van Manen, 1997), focusing on lived experience, not a reality, separate from the person (Valle, King, & Halling, 1989).

Giorgi and Giorgi (2008) argue that phenomenological studies aim to explore and describe situations as experienced by people in their daily lives. Consequently, the focus is on attaining first-hand descriptions of the individuals' experiences to explore a specific phenomenon. Karlsson (1995) adds that phenomenology provides a personal view of the individual's experiences. It allows for an opportunity to talk about an individual world view rather than describing events that occur outside of that world. It is important to understand personal meanings, perceptions, and experiences. Lindegger (2006) states that phenomenology is especially interested in the individual's perceived relationship to the world in which they live.

This study focuses on the understanding and experience of board members as it attempts to unfold meanings as they live as board members. This approach seems appropriate for exploring directors' and specialists' lived experience of corporate governance, AI technology, and ethics. The study is fundamental and provides a window to all that can be directly known because all knowledge is captured from the board members' and specialists' experiences and perspectives. Polkinghorne (1989) identifies two types of phenomenologically based inquiries. The first type focuses on how the different methods of conscious experience or memory of conscious experience understand an object or subject. Inquires of the second type search for information on how the meanings present themselves in the experience. This study attempted to gain direct knowledge of the experience through a phenomenological investigation of conscious and memory perceptions of board members and specialists. It also investigated how board members understand the meanings of the experience.

In this context, the phenomenological approach matched the assumptions and aspirations of the study. As stated, a phenomenological approach seeks to understand and describe the lived experiences of individuals concerning a particular phenomenon. The lived experience of the board of directors (including specialists sitting on boards) is a blend of their individuality, experience, background, and expertise. Thus, a phenomenological methodology in this study focuses on exploring the subjective experiences and perspectives of individuals involved in or affected by the governance of AI ethics, particularly within the context of board decision-making. The researcher and participant worked together to bring life to the experience being explored (Laverty, 2008).

Phenomenology involves an in-depth exploration of participants' experiences to uncover the essence of their perceptions, beliefs, and interactions related to the research topic. Before conducting interviews, the researcher prepared all the concepts and questions, which focused on experiences and relevant factors. The researcher collected data through in-depth and semi-structured interviews, allowing participants to freely express their thoughts and reflections on the board's role in shaping ethical considerations in AI governance. These interviews aimed to capture the participants' experiences on the underlying meanings and motivations that influence their views.

6.5.   RATIONAL FOR CHOOSING A PHENOMENOLOGICAL METHODOLOGY

There are three main reasons for selecting a phenomenological approach for the study. First is the assumption of the philosophical perspective that every researcher has their own view of what constitutes truth and knowledge (Wagner, Kawulich, Garner, 2012). These views guide our thinking, beliefs, and assumptions and frame how we view the world around us. The researcher has a core belief that social phenomena do not have an independent existence outside the minds of individuals who experience the phenomena in question. Hence, any information that an individual board member and specialist provides about their lived experience of corporate governance, AI ethics and governance informs ascription of meaning to these concepts.

Secondly, the phenomenological approach is more relevant to the study regarding board members. Pettigrew (1992), highly cited in the discourse "On Studying Managerial Elites",

made clear 20 years ago that the study of the processes within the boardroom is required to inform reflection on corporate governance practices. By its nature, phenomenological approach is ideally suited to identifying and understanding corporate governance issues, in particular, the role of the boards. This is so as the phenomenological approach research design can play a significant role in identifying and understanding the inner workings of the board, the so-called 'black box' of corporations (Massie, 2013), i.e. the view that only the board members who are seated at the meeting table understand how the "decision-making machine" works (Guerra, 2021). Over the past years, there have been many calls to open this black box of boards (board of directors) to identify what they do in practice (e.g. Zahra & Pearce, 1989; Pettigrew, 1992; Leblanc & Schwartz, 2007).

Thirdly, a qualitative study is produced through an interaction between the study participant/interviewee and the researcher/interviewer. In this transactional and subjectivist epistemology, the investigator and the object of investigation are assumed to be linked so that the findings are literally created as the investigation proceeds (Guba & Lincoln, 1994). Thus, corporate governance, AI ethics, and governance topics are best studied through the interpretive lens of the observer. In the study, the interpretation took place at two levels: i) the first interpretation is at the participant level, as research participants are interpreting governance structures, processes, and mechanisms within their organisations or organisations in general, and ii) the second interpretation is at the researcher-level, as the researcher tries to understand the data received about the phenomenon under investigation mainly in the context of the literature review.

In conducting this study, the researcher took the stance of a knowledgeable person (along with the study participants) with their domain knowledge in corporate governance, in general, and board governance, in particular. This knowledge allowed the researcher to find "patterns in the data, enabling [him] to surface concepts and relationships that might escape the awareness of informants" (Gioia, Corley & Hamilton, 2013). The researcher is a certified director with the Institute of Directors South Africa and a corporate governance scholar with two master's degrees in the area of governance, which is a build-up to the current study. Further, the researcher is an experienced board member who sits on various boards and board subcommittees such as audit committees. Furthermore, he has authored and co-authored professional articles on the subject of artificial intelligence and board governance.

Lastly, the researcher is a member of 'PhD Students in AI Ethics'[35] an international student association, an affiliate of the Centre for Artificial Intelligence Research[36], and a former contributor to Good AI[37].

Based on previous knowledge and experience in corporate governance issues, particularly board governance of information technology, the researcher has a good understanding of information about technology, AI ethics, and governance processes currently utilised within JSE-listed companies, enabling informed interpretation of information shared by study participants. This knowledge also helped the researcher co-construct how current processes need to be enhanced to fill gaps identified by the participants and improve the overall effectiveness of governance processes related to AI technologies within JSE-listed companies. The study has many built-in validation checks on the findings produced, including the constant comparison of data from multiple sources, the explicit search for disconfirming information, constant reflexivity about the research process and findings, the use of detailed descriptions, and most importantly, member checks to confirm the reasonability of the findings produced with a subsample of research participants.

6.6.   RESEARCH SETTING

The research setting is a "physical, social, and cultural site in which the researcher conducts the study. In qualitative research, the focus is mainly on meaning-making, and the researcher studies the participants in their natural setting" (Given, 2008). Simply put, the research setting is the location where the research takes place. At the start of the research project, the researcher opted to write directly to twenty-plus JSE-listed companies requesting consent to conduct research and access to board members as participants. However, most companies declined to participate, while others ignored the request.

The researcher thus changed the research site to the Institute of Directors South Africa (IODSA). Established in 1960 as a branch of the Institute of Directors in London, the Institute of Directors South Africa (IoDSA) is a non-profit organisation that represents directors of public and private companies, aspirant board members, business leaders and those charged with governance duties in their individual capacities. The IoDSA's main role is to

---

[35] https://www.phdaiethics.com
[36] https://www.cair.org.za/user/217
[37] https://thegoodai.co

develop those charged with governance duties, directors, and board learning and improving board effectiveness. The IoDSA's objectives are expressed as follows (IODSA, nd):

- To promote good corporate governance by empowering boards to effectively and legally discharge those duties
- To safeguard the integrity and status of directorship as a profession by serving as the professional body for boards of directors and through the provision of a professional designation such as Chartered Directors SA and Certified Director
- To uphold and maintain the concept of a responsible corporate citizen and its role in a free enterprise economy
- To provide an effective voice for directors and executives to maintain and strengthen the bonds and reciprocity between the IoDSA and organisations with similar objectives within South Africa
- To undertake, provide and support or assist in research, thought leadership, guidance, advice, opportunities for peer interaction, education and training which may assist in the attainment or advancement of any of the objects of the IoDSA.

Members of the IoDSA benefit from ongoing educational opportunities, leadership publications and updates featuring the latest developments in corporate governance, as well as unique networking opportunities. The IODSA is the only organisation in South Africa that represents directors in their capacities, which made it easier for the researcher to access the JSE-listed companies' directors directly for the study. The IODSA consent letter and conditions are attached in the study Annexure B. The only condition imposed by the IODSA was that the researcher should produce high-level research findings in an executive summary to IoDSA members and/or an article in Directorship magazine within six months of submission of their final thesis. The conditions were accepted by the researcher, who is also a member of IODSA, and this will also assist the researcher in sharing the findings with those in charge of governance in the country.

6.7.   RESEARCH TOOLS AND APPLICATIONS

Data collection enables the researcher to systematically gather theoretical and empirical data relevant to the research question and research problem (Burns & Grove, 2006). Mouton (2001) argues that data collection processes conform to the study's research methodology.

The researcher utilised both primary and secondary data to achieve the objective of the study. The primary data was collected through in-depth and semi-structured interviews; secondary data was collected from literature reviews.

6.7.1. PRIMARY DATA (INTERVIEWS)

One of the most critical qualitative research techniques is in-depth interviews, which has been used by the researcher to explore more on the research question and help unearth the participants' understanding and perspective on the role of the board in governing AI ethics. Semi-structured interviews ensure the uniformity of the research questionnaire to the participants. This essentially means that all participants were asked the same questions in order for the responses to be comparable by means of the sample (Jupp & Sapsford, 2006).

These perspectives are particularly useful in exploring social phenomena, or as a tool to "gather descriptive data in the subject's own words so that the researcher can develop insights into how subjects interpret some piece of the world" (Polit & Beck, 2008). This is also argued by Taylor and Bogdan (1984), who regard encounters between the researcher and participants as directed towards understanding of participants perspectives on their social and organisational lives, experiential elements, and their situatedness in society.

The collection of qualitative data by means of interviews is regarded as purposeful conversation between two or more people (Crookes & Davis, 1998; Bogdan & Biklen, 2007). The researcher selected a semi-structured interview approach for data collection mainly because it allows the participants an opportunity to express their experiences, views, and opinions (Lauer, 2006). Moreover, this allowed the researcher to have an open conversation with board members and AI specialists in order to obtain their understanding of the subject matter. The researcher recognised different types of data collection for a qualitative research study, such as interviews, observations, and documents (Merriam, 2002; Yin, 2003; Creswell, 1998; Crookes & Davis, 1998). In order to extract as much information as possible to a point of saturation during these interviews, the researcher employed a semi-structured interview schedule in order to allow a free flow of information based on specific themes (e.g., Board fiduciary duties, board role in AI alignment to organisational culture, adequate AI governance structures). Given the role played by the board as a focal point of governance, interviews were conducted with board members of different JSE listed companies. The participants represent the custodianship of corporate governance in these corporations.

The researcher aimed to establish and explore the social reality of participants based on their responsibilities as board members. The interviews were advantageous in that social and organisational explanations and arguments could be contextually constructed. Furthermore, interviews establish depth, nuances, complexities, and complementarity of data (Mason, 2002). Information obtained from the participants was audio-recorded and analysed following a specific transcription process (see Section 6.10). The analysis was intended to facilitate data/information congruence, variations, and insights into the research questions.

As a result of the Covid-19 pandemic and travel restrictions, the researcher conducted interviews through virtual meetings platforms. The study developed a model interview guide with a list of questions relevant to the research topic. The interview guide allowed respondents to demonstrate their understanding of AI technologies concepts and their views and opinions on how the board performs its function in governing AI ethics.

6.7.1.1.   Research Questionnaire (Annexure E)

The study intended to answer the complex question regarding the board's role in governing AI ethics. Thus, semi-structured interview questions were used to collect data on four pre-determined themes. The open-ended questions allowed the researcher to unearth the understanding and perspective of board members and specialists in governing AI ethics in the organisation. As Burns & Groves (2011) suggested, the questionnaire was used as a "primary empirical data collection tool and self-reporting mechanism" for information received from participants. In addition, the development and administration of the questionnaire enhanced the qualitative-oriented aspect of the research design. During the interviews, the following factors were laid out by the researcher:

- Participants were informed to respond to questions without "fair or favour" and that there were no "correct or incorrect" responses
- Ethical issues were taken into consideration to ensure that the privacy, anonymity, and confidentiality of participants were and remain guaranteed
- The participants were informed that the interviews were recorded
- Open-ended questions were asked to facilitate unrestricted perspectives by participants.

All questions were thematically linked to the role of the board in governing AI ethics in the context of board fiduciary duties, the board's role in AI alignment to organisational culture, and adequate AI governance structures. The questions were structured as follows:

Theme 1: How does the board discharge its fiduciary responsibility in overseeing AI technologies and its potential impact on the organisation?
- What is your understanding of artificial intelligence?
- What is your understanding of ethical concerns around artificial intelligence technologies in terms of possible societal harm?
- What is your understanding of the general role of the board as far as fudiciary duties are concerned?
- What is your understanding of the role of the board as far as technology. i.e.. artificial intelligence, is concerned?
- How does the board oversee artificial intelligence technologies in the organisation?

Theme 2: What is the board's role in ensuring that AI technologies are aligned with the organisation's ethical culture?
- What is your understanding of how the board play a key role in setting the ethical tone in the organisation?
- What is your understanding of how artificial intelligence technologies may impact on the ethical culture of the organisation?
- What is your understanding of how artificial intelligence technologies strategies are aligned to the organisation's ethical culture?

Theme 3: What is the board's role in ensuring that the research or design or development or deployment or use of AI technologies (whichever is applicable in your organisation's context) does not negatively impact on society?
- What is your understanding of the responsibility of the board in mitigating societal harm from AI technologies in play in your organisation?

Theme 4: What is the board's role in ensuring that adequate AI governance structures are in place?
- Does your board have a governance subcommittee for technology or AI?

- o If yes, what is your understanding of the role of the AI or technology governance subcommittee?
- o If no, do you think such a subcommittee is needed? Why or why not?
- Does your board have a responsible AI governance framework?
  - o If yes, what is your understanding of the framework?
  - o If no, do you think such a framework is needed? Why or why not?
- How does the board monitor the performance and value delivery of artificial intelligence technologies?

6.7.2. SECONDARY DATA

The researcher utilised secondary data to compare the study with readily available information. Secondary data refers to data collected already such as surveys, observations, experiments, questionnaires, personal interviews, websites, books, journal articles, internal records etc. This work is mainly captured in the literature review. The researcher selected and critically appraised research by searching over multiple databases in order to answer a clearly formulated question (Armitage & Keeble-Allen, 2008; Gough et al., 2012). The researcher searched for key information through academic research databases such as Scopus, google scholar, University of Pretoria Library, IEEE Xplore, Directory of Open Access Journals, EBSCOhost, Emerald Insight, SAGE Journals, ScienceDirect, ProQuest, etc., which made it easy to locate relevant literature. The researcher searched data bases using key words such as, Corporate Governance, Artificial Intelligence, AI ethics, AI governance, role of the board, ethics, etc.

The researcher also used a new technology for searching academic articles and journals called Connected Papers[38], an AI powered visual tool to help researchers find and explore papers relevant to their field of work. Connected Papers builds a spider web graph of similar papers in the field with references. The researcher was thus able to have a visual understanding of the key references, popular work, and dynamics of the research topic. The site also has a direct link to the reference articles. Example: consider a search on the role of the board which produced a spider web graph as depicted in figure 11 below. The spider graph shows Lam (2014) as the main recent scholar on the topic, and then shows all the other connected references on the topic itself.

---

[38] Connected Papers https://www.connectedpapers.com
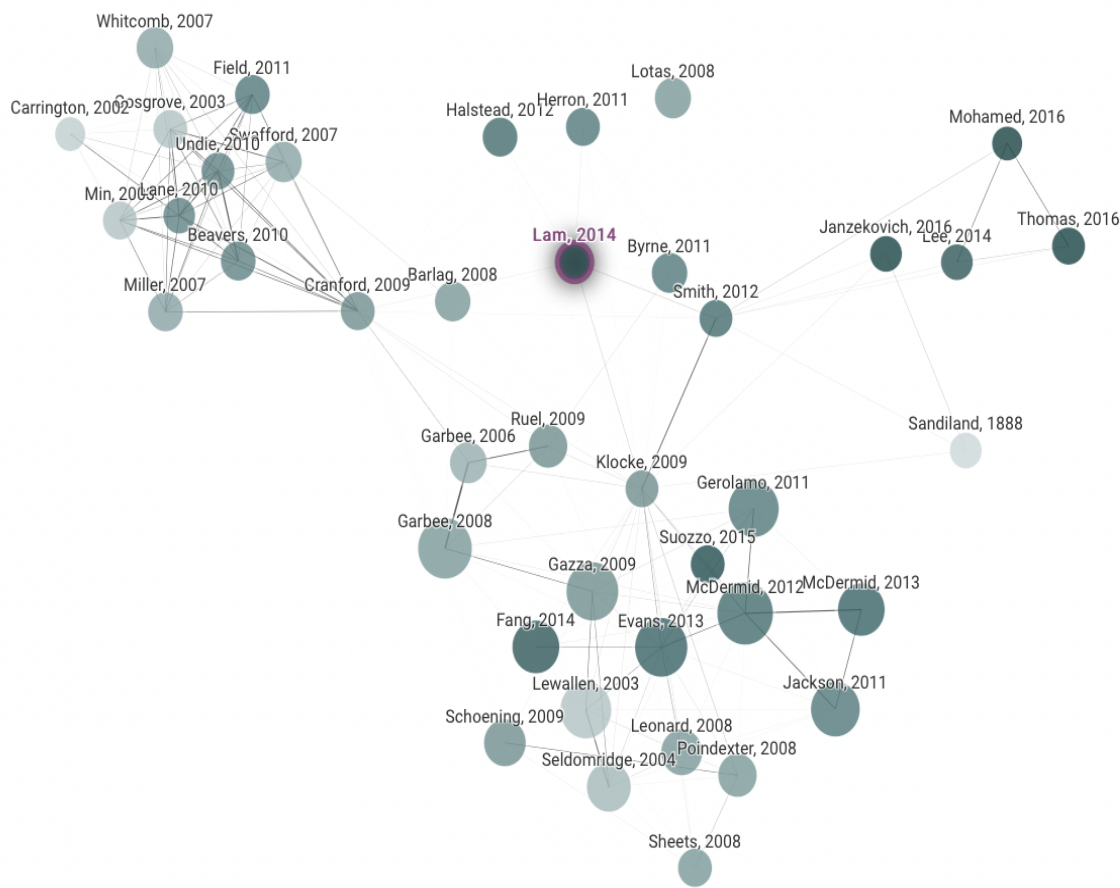
Figure 11: Connected Papers Spiderweb Graph on the Role of Board (Own illustration based on articles searched)

Secondary data collection and analysis provided the researcher with the additional benefit of interrogating more direct data. Secondary data allowed the researcher to substantiate the interviewee's responses during data analysis (Marion and Morrison, 2007). Thus, the analysed secondary data will be used as additional evidence and validity for qualitative research findings (Maxwell, 2005).

6.8. POPULATION

A research population is also known as a well-defined collection of individuals or objects known to have similar characteristics. In simple terms, the population is the entire group that the researcher wants to draw conclusions about. The Johannesburg Stock Exchange has 800 listed securities and approximately 400 listed companies, together with 60 Equity market member firms (JSE, nd). The IoDSA represents over 5500 members. The membership is widely representative of public and private sector commerce, industry,

government, and the business professions. Most members are from JSE listed companies, however, the researcher could not establish the exact number of members that are from JSE listed companies, as the information was not publicly available.

6.9. SAMPLING

Sampling is very important to any study and choosing it appropriately has an impact on the results of the study. Neuman (2006) states that "qualitative researchers focus less on a sample's representativeness than on how the sample or small collection of cases, units, or activities illuminates social life". Accordingly, the main aim of sampling is to collect specific cases, events, or actions that can provide clarity and enhance understanding to social life. Thus, to achieve the objectives of the study, the researcher selected a non-probability sampling method. Even though non-probability sampling tends to be more cost effective and convenient (Wagner, Kawulich, Garner, 2012), the researcher has no way of predicting or guaranteeing that each element of the population will be represented in the sample (Leedy & Ormrod, 2015).

A purposive sampling was used for the purposes of the study. Purposive sampling is one of the non-probability sampling methods. In this sampling method, the researcher does not intend to select research participants using a random sampling but to deliberately choice participants due to their qualities, experience, and expertise (Etikan, Musa, & Alkassim, 2016). The primary goal of purposive sampling is to sample the participants in the research in a very programmatic manner suitable and relevant to the research question. Neuman (2006) argues that purposive sampling is suitable for sampling unique cases or participants that are informative to the study. The researcher chose a purposive sampling method to select participants who are board members of JSE listed companies in South Africa.

The researcher selected twenty JSE listed companies board members who are members of IoDSA and five specialists in the area of artificial intelligence who interact or report at the board level i.e., Chief Information Officers, for semi-structured interviews. However, one participant pulled out at the last minute. Table 2 below shows the anonymised profile of participants, experience, industry, education, and number of boards.

Table 2: Participants Profiles

Participant	Education	Industry(s)	Board Experience in years	No of JSE boards – current and previous	Tech background
Participant 4 (P4)	Masters	Finance and Auditing	22	2	No
Participant 5 (P5)	Masters	Finance and Banking	11	2	No
Participant 6 (P6)	Masters	Telecoms and Financial Services	4	4	Yes
Participant 7 (P7)	PhD	Mining, Manufacturing and Telecoms	4	3	Yes
Participant 8 (P8)	CA(SA)	Financial Services	14	4	Yes
Participant 10 (P10)	PhD	Financial Services	21	10	Yes
Participant 11 (P11)	Masters	Consulting engineering and construction	20	1	No
Participant 12 (P12)	Masters	Energy, Telecoms and IT	30	1	Yes
Participant 13 (P13)	Masters	Health Care and Business Consulting	20	2	Yes
Participant 14 (P14)	CA(SA)	Finance and Retail	20	4	No
Participant 16 (P16)	Masters	Law and Corporate Governance	30	2	No
Participant 17 (P17)	Masters	Governance and Legal	15	2	Yes
Participant 19 (P19)	Honours	Energy, Telecoms and IT	14	8	Yes
Participant 21 (P21)	Masters	Financial Services	10	1	No
Participant 22 (P22)	CA(SA)	Finance	7	4	Yes
Participant 23 (P23)	Honours	Mining, IT and Financial Services	20	4	Yes
Participant 24 (P24)	CA(SA)	Finance, Properties and Marine	13	4	Yes
Participant 3s (P3s)	Honours	Financial and Insurance	9	Advisory at board level	No
Participant 1(s) P1(s)	Honours	Consulting in Banking and Financial Services	8	Consulting at Board Level	Yes
Participant 2(s) P2(s)	Honours	Data Analytics	4	Reporting at Board Level	Yes
Participant 9s (P9s)	Masters	Chief Information Officer	21	4	Yes
Participant 15s (P15s)	Degree	Chief Information Officer	10	Reporting at Board Level	Yes
Participant 18s (P18s)	Masters	Technology Industry (CIO)	12	Reporting at Board Level	Yes
Participant 20s (P20s)	Professor	Academic, AI technologies	15	Academic	Yes

* (s) = Industry Specialists

6.10. DATA TRANSCRIPTION

As already mentioned in Section 6.7.1 above, the researcher conducted interviews through virtual meetings platforms, thus allowing for an opportunity for audio and virtual (some cases) recording of the interviews. Research participants were informed and gave consent to be recorded, in the consent letter (Annexure C&D) and before the interviews.

Bailey (2008) argues that in qualitative research, data transcription is conducted of individual or group interviews and generally written verbatim (exactly word-for-word). The process of transcription may differ depending on the individual researcher. Transcripts that are used mainly to select quotes and sound bites may not be needed at the same level of detail as transcripts which will be systematically reviewed, grouped into themes (often through a process of coding), and analysed for content (Bailey, 2008). Transcription of interviews is often presented uncritically as a direct conversion of recorded audio to text. Transcription is a time-consuming and often tedious task which can take hours and days to complete, depending on typing speed (McMullin, 2021). Transcription is not, however, a mechanical process where the written document becomes an objective record of the event—indeed, "written text varies from the spoken word in terms of syntax, word choice and accepted grammar" (Davidson, 2009).

As technology to facilitate transcription improves, many researchers have shifted to using voice-to-text software that employs AI technologies rather than human transcription (McMullin, 2021). In this case, the researcher employed Otter.ai, an application technology which transcribes speech to text using artificial intelligence and machine learning. Otter.ai, partnered with Zoom and MS Teams, allows the researcher to transcribe audio recordings and has the ability to turn audio conversation to smart notes live as the interview happens (Otter.ai, 2019). It should be noted that the system displayed some level of bias, as it was difficult for it to recognise voices of some participants. Thus, the researcher had to listen to all audio recordings to validate and edit where the text was misidentified or not properly recognised by the system.

The researcher had to make subjective decisions throughout about what to include (or not), and whether to correct mistakes and edit grammar and repetitions (McMullin, 2021). There is thus a spectrum between "naturalized" transcription (or "intelligent verbatim") which adapts the oral to written norms, and "denaturalized" transcription ("full verbatim"), where everything is left in, including utterances, mistakes, repetitions and all grammatical errors (Bucholtz, 2000). The researcher has included the original text transcript from the research participants which will be posted in institutional repository of the University of Pretoria, an open access electronic archive collecting, preserving and distributing digital materials created, owned and hosted by the University of Pretoria. In chapter 7: Data collection and

presentation, the researcher extracts direct quotes and contextualises the text from the participants.

6.11. DATA ANALYSIS TECHNIQUE AND INSTRUMENTS

Data analysis is when data is transmuted into meaningful information that contributes to knowledge production. The process is achieved through "different techniques and methods used to analyse data previously collected from various settings" (Bogdan & Creswell, 2003). Data analysis aims to organise the layout structure and bring out the meaning from data collected (Neuman, 2000).

The interview data was made into a transcript in text data format using Otter AI software. The researcher tried to interpret the ideas, combined with the inductive data, in order to come up with an interpretive paradigm, which is the philosophy of discussion in which he tried to explore the essence in order to conclude the research objectives. As the researcher who used field-based information (Colaizzi, 1978), he had to classify significant statements, themes and lastly essence analysis which contain both statements and quotations. For the structured interview, the researcher needed to use the technique of statement significance and theme which could suit the text information (Creswell, 2003).

As suggested by Cho (2014), data was collected through interviews and review of material. As such, data from interviews was systematically collected and analysed in an attempt to understand the structure, processes, and rational mechanisms inherent within corporate governance and the role of the board in governing AI ethics. The researcher conducted thematic analysis through transcription of verbal data retrieved in interviews with the participants. After familiarizing with data from participants, an initial list of ideas was generated. The researcher then developed significant statements, which were then grouped into a theme. The theme was developed based on significant statements duplicated and repeated by many participants. Another criterion for the development of the theme is that

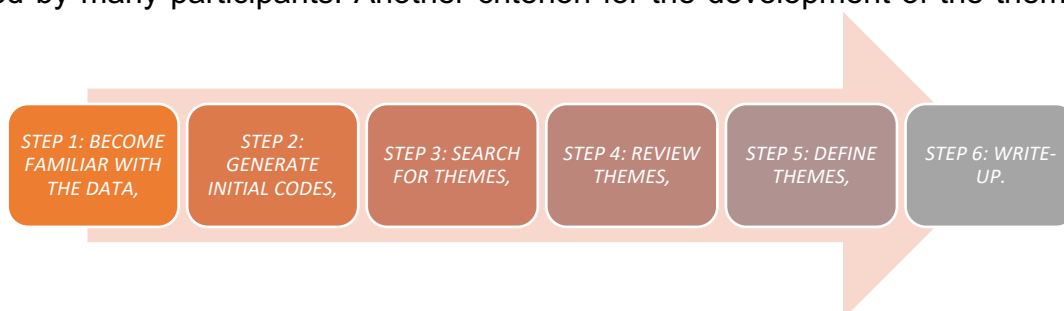| STEP 1: BECOME FAMILIAR WITH THE DATA, | STEP 2: GENERATE INITIAL CODES, | STEP 3: SEARCH FOR THEMES, | STEP 4: REVIEW THEMES, | STEP 5: DEFINE THEMES, | STEP 6: WRITE-UP. |

Figure 12: Braun & Clarke's six-phase framework for doing a thematic analysis

the theme must be an issue or experience related to the research objectives and research problems. The research followed a six-phase guide by (Braun & Clarke, 2006) below.

Step 1: Become familiar with the data

The first step was for the researcher to familiarise himself with the entire body of data (i.e., all participant interviews). This was done through listening, reading, and re-reading the transcripts. At this stage, the researcher made notes and jotted down early impressions. This step helped the researcher to gain a comprehensive understanding of the interview content from A-Z.

Step 2: Generate initial codes

In this step, the researcher meaningfully and systematically organised the data from interviews and secondary sources. Since the researcher was concerned with addressing specific research questions, the analysis was done through a theoretical thematic analysis rather than an inductive one, meaning the researcher did not code every piece of text in the interviews. The researcher coded each data segment relevant to the research question. At first, the researcher tried ATLAS qualitative data analysis, given the large data sets, but it was hard to understand, given that the researcher was not trained in the system. The researcher then used both MS Word and Excel to highlight specific themes.

Step 3: Search for themes

As the process progressed, the researcher explored potential themes, i.e. AI framework, AI risk management, social and ethics committee role, AI experts, etc. This step required the researcher to identify patterns, connections, or recurring elements in the data that signify broader and more abstract themes, contributing to a deeper understanding of the subject matter. As Braun & Clarke (2006) argued, there are no hard and fast rules about what makes a theme. A theme is characterised by its significance; some themes were merged as they overlapped. At the end of this step, the codes were organised into broader themes linked to the research question. The themes were predominately descriptive, i.e., they described patterns in the data relevant to the research question.

Step 4: Review themes

Following theme identification, the researcher reviews and refines the themes to ensure coherence and relevance to the entire dataset. During this step, the researcher reviewed, modified and developed the preliminary themes identified in Step 3.

Step 5: Define themes

This is the final refinement of the themes, and the aim is to "identify the 'essence' of what each theme is about" (Braun & Clarke, 2006). This stage involved developing a clear description of what each theme represents and how it contributes to the overall research question. The final refined themes were focused on understanding of artificial intelligence, ethical concerns around AI technologies, role of the board, role of the board in AI, setting the ethical tone, AI technologies impact on the ethical culture, AI technologies strategies alignment, governance subcommittee for AI, AI governance framework and performance and value delivery of AI technologies.

Step 6: Writing-up

The analysis findings are then incorporated into a narrative or report. This involves presenting the identified themes, supporting them with relevant quotes or examples from the data, and constructing a coherent story that encapsulates the essence of the research. The data presentation was based on four sub-themes drawn from the primary research questions that formed the core of the semi-structured interviews. The findings are conceptualised extracts and quotes from participants' semi-structured interviews.

6.12. TRUSTWORTHINESS AND CREDIBILITY

In quantitative research, the terms 'reliability' and 'validity' are used to determine accuracy; however, qualitative research is not concerned with accurate measurement, and as such, at issue here are the terms 'trustworthiness' and 'credibility' (Wagner, Kawulich, & Garner, 2012; Leedy & Ormorod, 2015. The overall objective of establishing trustworthiness and credibility is to provide considerable evidence that the findings of the study are reliable and accurate, and that the research process was conducted in an ethical manner. To ensure trustworthiness, the researcher ensured that the data collected is safe and non-threatening, and that the relationships with the participants were transparent. During the interviews, the participants were reassured that all the data collected would be treated with confidentiality and with necessary sensitivity (see Annexure C). The qualitative research study was done in such a way as to ensure credibility, dependability, and confirmability (Lincoln and Cuba, 1985).

6.12.1. CREDIBILITY

Credibility refers to the extent to which the research findings are accurate and authentic. Polit & Beck (2012) argue that credibility refers to the accuracy of the data or participant views, and representation and analysis of data by the researcher. Credibility is enhanced by the researcher describing their experiences as a researcher and verifying the research findings with the participants (Cope, 1969). Lincoln and Guba (1988) suggest that a researcher should implement a number of strategies to ensure the likelihood that the findings produced are credible.

In this case, the researcher conducted personal interviews with JSE listed companies board members as well as AI specialists who interact or report at the board level. The researcher also selected participants or board members representing different sectors of JSE i.e., financial, manufacturing, mining, banking, construction, etc. Moreover, the participants represented diverse groups of individuals in terms of demographics, gender, age, and educational levels with different views. These were all strategies employed to ensure that the data emerging was credible. The researcher also used multiple sources of data (interviews and literature reviews) and triangulated their findings (through AI technologies specialists).

Triangulation was used as a key strategy to establish credibility. In qualitative research, triangulation means exploring a research question from different angles in order to enhance the accuracy of the data collected (Nueman, 2006) and to strengthen validity (Davidson & Layder, 1994). It can involve using multiple data collection methods, multiple data sources, and multiple theoretical perspectives to interpret a single set of data (Denzin 1978). This implies that if something is observed from one viewpoint, the chances are that the results will be one-sided; thus, the view or perspective may not be entirely accurate. In qualitative research, triangulation is used to ensure the validity and reliability of the findings, based on data collected from multiple sources.

Morse (1998) has suggested three principles that may be used to improve the legitimacy of qualitative research. Firstly, there should be adequate data to achieve 'saturation'. Secondly, data collected for the purposes of the research should be appropriate to the theoretical basis. Thirdly, it should be possible to verify the researcher's account or interpretation by using secondary informants. As mentioned, the researcher used methodological

triangulation for the latter, based on the use of participants and AI specialists' in-depth interviews. This meant that information could be checked at different levels to enable reasonably accurate interpretation. Thus, data triangulation helped the researcher mitigate biased information as some participants tended to offer what the researcher wanted to hear. Furthermore, it assisted the researcher to increase the validity and reliability of the findings, as it helped to confirm or refute their initial assumptions and interpretations. It is an important part of the phenomenological approach process, as it helps to ensure that the findings are based on a thorough and comprehensive analysis of the data.

The data collected from participants were also reviewed and validated against secondary data. Being reflexive throughout the research process provided an added dimension in terms of striving for validity.

6.12.2.   DEPENDABILITY

Dependability refers to the consistency of the research findings over time. In order to establish dependability, the researcher may use rigorous research methods, document their research process carefully, and seek feedback from research participants (Linclon and Guba, 1988). Dependability and confirmability can be determined through a properly managed audit.

In order to establish dependability, the researcher continuously examined the whole research process, which included the various stages of the research project, as well as the techniques used to analyse the data. In all stages of the research from research proposal, sampling, data collection, data analysis, and interacting with the researcher's supervisor, the aim was to establish that the process followed was applicable in terms of the research problem, that consistent process management occurred, and that all the university processes were followed.

6.12.3.   CONFIRMABILITY

Korstjens & Moser (2018) argue that confirmability refers to the degree to which the findings of the research study could be confirmed by other researchers. Confirmability is concerned with establishing that data collected and interpretations of the findings are not figments of the researcher's imagination, but clearly derived from the data collected (Korstjens & Moser,

2018). This means that the interpretation of data should not be based the researcher's preferences and viewpoints but needs to be grounded in the data. An audit trail is a strategy used to ensure dependability and confirmability in the interpretation process of analysis (Guba, 1979).

In order to establish confirmability, the researcher used transparent research methods which included documentation of the research process carefully. Part of this process included a complete set of notes on the research process (Accessed in institutional repository of the University of Pretoria), decisions made based on supervisor's inputs and guidance during the research, team meetings recordings, sampling, research materials used, emergence of findings, and information about data management. This enables an auditor to evaluate the transparency of the research path.

6.13. ETHICS CONSIDERATION

Ethical considerations are important in research in order to ensure that the rights and welfare of research participants are protected and that the research is conducted in a responsible and transparent manner. The researcher took into account the following key ethical considerations when conducting the research:

- *Informed consent (See Annexure C & D for participants informed consent):* Research participants were fully informed about the nature and purpose of the study, as well as any potential risks or benefits, and were given the opportunity to ask questions and make an informed decision about whether to participate. All participants gave consent in the introductory email which contained a research interview calendar invitation link with it: The participants were informed that by selecting the **"Agree/ Yes"** option, they granted permission for participation in the research interviews and confirmed that they understood the purpose of this research. The consent included the interview responses being used anonymously in the research report and in any subsequent academic publications that might stem from this research report and findings. The participants were also given an option to decline by selecting **"decline/ no"**, which meant they are excising their right not to participate in the research.

- *Privacy and Confidentiality:* The researcher ensured that all ethical considerations were followed, such as protecting the confidentiality of information and participants

and obtaining informed consent from participants. The researcher ensured that the personal information of research participants was kept confidential and not disclosed without the participant's consent. The researcher respected the privacy, anonymity, and confidentiality of the participants. These concerns were addressed in the Letter of Introduction and Informed Consent attached as Annexure C & D. No identifying data of participants was recorded in the study.

- *Risk of harm:* The study was classified as low risk by the University Ethical committee as it did not pose any harm to participants. In addition, the researchers took steps to further minimise the risk of harm to research participants, both physical and psychological. In the informed consent letter participants were informed that there are no penalties should participants choose not to participate or withdraw from the study. They were also informed that they have an option to withdraw at any time or not answer any question if they did not wish to. This was also emphasised before the interviews with participants. The participants were also given email contacts to the University ethical administrators if they had any concerns or complains regarding the ethical procedure of the study.

- *Compliance to POPI Act:* As already discussed in Chapter 1 of this study, South Africa (SA) has introduced the Protection of Personal Information Act, 2013 also known as POPIA which come into effect on 1 July 2020. The Act is designed to protect personal information of SA citizens. If a researcher collects personal data of an SA citizen i.e., through interviews or focus groups, or surveys that collect identifiers such as names, contact information, etc., the researcher should ensure full compliance with the POPI Act. Thus, the researcher ensured that any personal identifying information of participants was stored in a secured, access controlled location to avoid data harvesting or mining or use by third parties. The de-identified datasets of participants will be stored in a safe and password protected cloud storage in the Department of Philosophy and will be kept for 15 years. Furthermore, on-site storage (personal hard drives) is strictly access-controlled to the researcher only. In addition, no personal identifying information was mentioned in the study, the participants are referred to using pseudonyms (alias names) i.e., participants A, B or C.

- Ethical *clearance:* The researcher obtained ethical clearance from the University of Pretoria before conducting the research and followed all the university's ethical requirements (See Annexure A UP ethical clarence).

By considering these ethical considerations, the researcher ensured that the research was conducted in a responsible and transparent manner and that the rights and welfare of research participants were protected. The researcher committed to the highest standard of ethics with the research participants, meaning that all the data collected from participants was treated as raw data, and was not tampered with or misrepresented with a view of producing a quality research report, rather than simply representing the honest views of participants.

6.14. CONCLUSION

The chapter presented the methodology and design adopted for the research to help the researcher achieve its envisioned objectives. The methodology is one of the essential crucial chapters of the research process, as it guides the reader on the coherent steps adopted by the researcher while embarking on the study. The study adopted a qualitative research methodology. To gain direct knowledge of the experience, the phenomenological approach was followed on the conscious and memory perceptions of board members and specialists. The chapter further provides a detailed description of the research approach, tools and applications, primary and secondary data collection and data analysis process methods selected for this study. The sampling procedures, validity and reliability, limitations, feasibility and positionality were also deliberated in detail. Ethical considerations were discussed to ensure that the rights and welfare of research participants were protected and that the research was conducted responsibly and transparently.

CHAPTER 7: DATA COLLECTION AND ANALYSIS

7.1. INTRODUCTION

This chapter provides a synopsis of the data collection conducted and the presentation and analysis of semi-structured interviews and secondary data. The interview responses and analysis have been divided into sub-themes 1 to 4 to answer the main research question, "*The Role of the Board in Governance of Artificial Intelligence Ethics – A Case for JSE Listed Companies?". The division into sub-themes was done to allow for* a smoother transition and to enhance clarity, making it easier for the reader of this study to follow the structure. These four research questions are:

- *How does the board discharge its fiduciary responsibility in overseeing AI technologies and its potential impact on the organisation?*
- *What is the board's role in ensuring that AI technologies are aligned with the organisation's ethical culture?*
- *What is the board's role in ensuring that the research, design, development, or use of AI technologies does not negatively impact society?*
- *What is the board's role in ensuring that adequate AI governance structures are in place?*

The chapter further details data interpretation and analysis. The researcher conducted thematic analysis through transcription of verbal data retrieved in interviews with the participants. The researcher then developed significant statements, which were grouped into themes. The theme was developed based on significant statements duplicated and repeated by many participants. The research followed the six-phase guide by (Braun & Clarke, 2006). The analysis included breaking down the data from "large bodies of text into smaller units and individual words" (Creswell, 2013) and evaluating the relationship between these.

7.2. BACKGROUND AND PROFILE OF RESEARCH PARTICIPANTS/ RESPONDENTS

As highlighted in methodology chapter, the researcher selected twenty JSE listed companies board members who are members of IoDSA and five specialists in the area of

artificial intelligence who interact or report at the board level i.e., Chief Information Officers, for semi-structured interviews. The participants represented 58 companies' both current and previous board memberships. It is a common phenomenon that board members of JSE-listed companies serve in boards of multiple listed companies simultaneous. For example, participant 7 sits on the boards of telecom industry, while also serving on the board of JSE-listed manufacturing and mining.

The researcher interviewed twenty-four participants and one participant exercised the right to decline to participate in the study. Semi-structured interviews were conducted using a purposive sampling methodology. The primary goal of purposive sampling was to sample the participants in the research in a very programmatic manner suitable and relevant to the research question. The researcher chose a purposive sampling method to select participants who are board members of JSE listed companies in South Africa. All participants were interviewed using video-conferencing software, i.e., Microsoft Teams and Zoom conferencing software and recorded through Ottis AI software for data transcription.

7.3.   PRESENTATION OF FINDINGS AND ANALYSIS

The data presentation is based on four sub-themes, drawn from the primary research questions, that formed the core of the semi-structured interviews. The findings are conceptualised extracts and quotes from participants semi-structured interviews[39]. The four sub-themes (each with their own sub-questions) that were posed to research participants were expanded from the primary research question, namely, "what is the role of the board in governance of artificial intelligence ethics".

The interview questions were designed to allow research participants to elaborate, thereby allowing them to demonstrate their understanding of the role the board plays in terms of governance, corporate governance, AI governance processes, and identifying and dealing with the ethical implication of AI technologies. As such, data from interviews was systematically collected and analysed in an attempt to understand the structure, processes, and rational mechanisms inherent within corporate governance and the role of the board in governing AI ethics. After familiarising himself with data from participants, an initial list of ideas was generated by the researcher. The researcher then developed significant

---

[39] The research full transcript can be accessed from UP Repositories Research Data https://repository.up.ac.za

statements, which were then grouped into a theme. The findings and analysis from the semi-structured interviews are presented in the following sub-theme section.

7.4. Research sub-theme 1:

The first research sub-theme: How does the board discharge its fiduciary responsibility in overseeing AI technologies and its potential impact on the organisation? This research sub-theme included 4 questions posed to the participants: (1) What is your understanding of artificial intelligence?; (2) What is your understanding of ethical concerns around artificial intelligence technologies in terms of possible societal harm?; (3) What is your understanding of the general role of the board as far as fiduciary duties are concerned?; and (4) What is your understanding of the role of the board as far as technology, i.e. artificial intelligence, is concerned?

The sub-theme aimed to establish a comprehensive understanding of corporate governance, particularly the board's fiduciary role in governing AI technologies. Additionally, the question aimed to explore whether board members have a basic understanding of corporate governance in the context of AI technologies, including the ethical implications of AI technologies. This helped establish the overall essence of the corporate governance concept, such as the fiduciary understanding of the role of the board.

7.4.1. The term artificial intelligence

Under the sub-question (*How does the board discharge its fiduciary responsibility in overseeing AI technologies and its potential impact on the organisation?),* the participants were asked, "What is your understanding of artificial intelligence". Asking the participants about their understanding of artificial intelligence was an essential step in the research process to establish whether the board members had a basic understanding of AI technologies. Understanding AI technologies is crucial for the board members to fulfil their fiduciary responsibility in overseeing their governance effectively. The question aimed to explore the level of awareness and knowledge that board members had regarding AI technologies, including their potential benefits and risks. The participants' responses helped gauge the extent to which the board members understood AI technologies and identified any knowledge gaps that needed to be addressed. Ultimately, this information was critical

in determining the board's potential preparedness to govern AI technologies effectively and ensure their adoption aligns with the corporation's strategic goals and ethical standards.

The participant's responses were as follows:

Responses from board members, herein referred to as 'Participants (P)'.

AI is a step beyond traditional logic-based computers that make predefined decisions, AI gives machines broad parameters to make decisions and learn from past experiences (Participant 4).

Participant 5 (P5) distinguishes between AI and machine learning. AI is a repetitive decision-making that does not require human judgement, whereas machine learning involves teaching machines to understand certain patterns and behaviours. AI is used when human judgement is not the primary decision-making process and AI involves intelligent automation (P5).

Participant 6 (P6) views machine learning as inputting data into a platform and teaching the machine to understand patterns and behaviours. As the machine continues to observe behaviour, it gets better at predicting outcomes and anticipating behaviour. AI has potential to create digital twins of individuals based on data sources and AI can mimic human behaviour but get better over time (P6). Participant 7 (P7) agrees with the sentiment raised by P6. According to P7, AI is the process of getting machines to do work that humans previously did in terms of analysis and interpretation. Machines now have a level of virtual intelligence previously attributed to humans.

AI is when computers mimic human intelligence, where computers now carry out processes that a human being would do. It is essentially a simulation of human intelligence by computer (Participant 8). However, participant 23 (P23) views AI as an area of computing that can approximate low-level human thinking. They suggest that AI is capable of replacing routine and predictable human tasks and that the outcomes of AI systems are highly dependable and predictable. AI technologies are becoming sentient and able to learn, but this is not yet mainstream (P23).

AI involves developing programs that enable technology to be utilised to effect transactions that would normally be affected by humans, such as opening and transacting of a bank account. This involves using technology and AI to convert mundane tasks that would be accurate, repetitive, and free of human error (Participant 8).

According to participant 11, AI uses computational tools to make our lives easier. It involves using computers to simplify processes, procedures, and tasks (P11).

Participant 12 (P12) believes that AI is an attempt to merge computers and machines with humans. The dream of AI is to give consciousness to machines so that they can replicate human behaviour. However, P12 believes we are still far from achieving "true" AI.

AI is a broad term that refers to how machines, through coding, can do the processing that humans would ordinarily do. It is problem-solving on a binary basis using algorithms, which almost simulate thinking, and that is why we talk about machine learning (Participant 13).

Participant 16 (P16) argues that concerns about AI depends on how one looks at it. It has both risks and opportunities for businesses and society. On the one hand, machines could take over the world like in the "Terminator movies"; on the other hand, machines can help us make decisions by assessing and evaluating data (P16).

AI is a way of obtaining information or processing data into information for decision-making. It involves putting data into algorithms that direct people to ask for information or input, which then provides information as an output (Participant 17).

Participant 19 (P19) indicated that they only have an experience of using bots for debtors and finance and mining big data using algorithms to use in strategy, particularly in retail.

Participant 21 (P21) describes AI as using technology and applying machine learning approaches to solve defined problems. "At the basic end, AI can solve specific problems, but as it moves along the spectrum, it can solve undefined problems and grapple with issues it hasn't been specifically designed to resolve" (P21).

Participant 22 (P22) has an understanding of AI as Robotic Process Automation. Participant 24 (P24) shares a similar sentiment that AI is a system that can improve processes and

analyse data, implying that it can do so in a more efficient and intelligent way than humans could do manually. Both P23 and P24 recognise the power and potential of AI in augmenting or replacing human labour in certain areas.

Responses from specialists herein referred to as 'Participants (Ps)'.

AI is a technology that can mimic human intelligence, such as computer vision, machine learning, and robotic process automation (Participant 1s). Participant 2s (P2s) and participant 20s (P20s) agree with participant 1s (P1s) that AI is an attempt to simulate human intelligence into a machine and program it to think and act like humans. It can execute actions, whereas machine learning only suggests actions. Participant 20s (P20s) added that AI could mimic human intelligence and incorporates automation, "it is a broad term that encompasses anything that can replace a human task" (Participant 20).

AI refers to machine-driven intelligence that mimics cognitive functions associated with the human mind, such as learning and problem-solving (Participant 3s). According to participant 9s (P9s), AI is about creating, interpreting, and generating assumptions based on information. It artificially creates intelligence, and machine learning builds on those assumptions. AI uses technology and processes to improve and automate manual processes, reduce errors, and make services faster with less human intervention (Participant 15s).

AI goes beyond input and output and interacts more deeply and intelligently with users. It tries to anticipate user needs and make recommendations (Participant 18s).

Analysis

The participants' definition of artificial intelligence focused on three key aspects, (i) AI's mimicking human intelligence and behaviour; (ii) AI's evolution from rule-based to machine learning; and (iii) AI ability to automate decision-making and tasks.

AI's mimicking Human Intelligence and Behaviour: Most participants defined AI technologies as systems that mimic or simulate human behaviour (P6, P8, P23, P2s, P3s and P20s). These definitions of AI technologies mimicking or simulating human behaviour and intelligence align with AI's aspiration to replicate cognitive processes. Syam and

Sharma (2018) have argued that AI technologies have the ability to emulate some of the cognitive functions that we attribute to the human mind. By employing techniques like machine learning, AI technologies simulate human-like understanding and problem-solving. This resonates with academic definitions that describe AI technologies as a study of computation activities that can simulate or even recreate the cognitive capabilities of human beings (Hu, 2008; Vassilopoulos & Georgopoulos, 2010). Undoubtedly, AI technologies are improving their learning and adaptation capabilities, furthering its resemblance to human learning and growth. Thus, an interdisciplinary approach is crucial in the study of AI technologies and their behaviour, which includes ethical consideration in the design, deployment, and development of AI.

AI's evolution from Rule-based to Machine Learning: AI technologies have moved from rule-based programming towards systems capable of learning from data and experiences, underscoring AI's transformational nature in reshaping how computers interact with and process information. "AI is a step beyond traditional logic-based computers that make predefined decisions, AI gives machines broad parameters to make decisions and learn from past experiences" (Participant 4). This definition by Participant 4 stems from AI technologies' ability to make decisions beyond predetermined instructions through the machine and deep learning models. On the other hand, rule-based systems are programmed to follow rules, while data-driven AI technologies have the ability to adapt and respond to different dynamic scenarios. Data-driven AI technologies can handle complex and intensive issues in relatively variable environments, whereas rule-based AI systems are easier to interpret.

Automation of Decision-Making and Tasks: The definition by P11 emphasises AI's capacity for repetitive decision-making outside of human judgment. In Chapter 2 of this study, it was concluded that AI could be defined as the science that attempts to have computers perform and execute tasks that humans would perform, in part or full, and in some instances, AI systems perform even better than humans. This definition highlights the concept of automating tasks through AI technologies, leading to efficiency gains and reduced human intervention in routine processes.

Machine learning and pattern recognition notions tie into this idea, enabling AI systems to learn from data and make informed decisions autonomously. Academic definitions often

elaborate on the role of algorithms and data in training AI systems to make decisions, providing a foundation for this interpretation.

In conclusion, the participants collectively reflected on various facets of AI, ranging from its departure from rule-based operations to its role in automating tasks, simulating human-like behaviour, enhancing human capabilities, and representing both a technological and cognitive endeavour. The interplay of these perspectives from participants highlights the diverse and evolving nature of AI technologies as a field that continues to shape our technological landscape. The participants were able to simplify the definition of AI and capture the core themes that surround the concept of artificial intelligence based on their experience as board members.

7.4.2. Ethical concerns around artificial intelligence technologies in terms of possible societal harm.

Under the sub-question (*How does the board discharge its fiduciary responsibility in overseeing AI technologies and its potential impact on the organisation?) The* second sub-question explored board members' understanding of ethical concerns related to AI technologies and their potential risk of causing societal harm. The question (*What is your understanding of ethical concerns around artificial intelligence technologies in terms of possible societal harm?)* was asked because research by academics, non-governmental organisations, and professionals reveals significant implications of AI technologies for society (see Chapter 4). The board members' understanding of these implications is essential for fulfilling their fiduciary responsibility toward the corporation and its stakeholders.

AI technologies have the potential to transform various industries and create new opportunities, but also raise significant ethical concerns. These concerns include the potential for biased decision making, privacy violations, and the displacement of jobs. AI technologies can perpetuate societal inequalities, amplify existing biases, and exacerbate discrimination.

Therefore, board members need to understand the ethical implications of AI technologies, particularly in terms of possible societal harm. This understanding is crucial in ensuring that corporations operate ethically and transparently, and that AI technologies align with

corporations' strategic goals and ethical standards. Understanding these ethical concerns can help board members develop appropriate policies and procedures to mitigate potential risks and ensure that AI technologies are developed and deployed responsibly and ethically.

The participants' responses were as follows:

Responses of board members herein referred to as 'Participants (P)'.

Most participants expressed ethical concerns with regard to job losses. P4's concern was mainly about machines replacing people in decision-making and targeting advertising. Such concern was also raised by P19 who also indicated concern that AI technologies is "using people's information without permission and companies selling information to AI companies". P22 added that there should be a balance between efficiency and job elimination with AI technologies. However, P22's view is that AI technologies can enhance decision-making and allow workers to focus on value-adding tasks, but worry about job loss, such as in the case of Amazon factories.

As the chair of the social and ethics committee, P5 indicated that they always ask the committee about the ethical implications of using AI technologies, such as what it is used for and what are the check and balances in place to ensure customers are treated fairly and that there is no harm to society. P5 emphasises the importance of treating customers fairly in the financial sector, whether it is through artificial intelligence or human intervention. They state that using biased datasets or processes could result in harm to society (P5).

The potential impact of AI is on jobs and productivity, for example, AI technologies have been used to predict customer calls in the telecom industry and mentioned how it reduces call centre staff. AI technologies can perform tasks such as auditing and preparing legal cases, which could make some jobs redundant (P6). P6 further expresses concern about the nature of human interaction being redefined and the moral and ethical implications of too much free time. However, P8 argues that AI technologies are used for the simulation of processes, and the deployment of AI technologies is an advantage to a profession such as accounting and finance. For example, P8 indicated that in their organisation, they deploy AI technologies on voluminous processes to achieve efficiency and speed. Secondly, they deploy AI technologies "so that people can manage rather than do processing but focus on more strategic work or thinking work and then use their time efficiently". The ethical concerns

could be if AI technologies are used as a tool to retrench people; thus, organisations "need to think of re-skilling the people so that they don't lose jobs because of AI technologies deployment" (P8).

P7 indicated that their concern is for "machines to actually do things that are unfair and unjust .... so you have to be ethical in terms of not allowing the use of machines to do things that would be unethical for other people" (P7).

For AI to be ethical, it should not be involved in illegal or dishonest activities. AI processes must be both legal and permissible from common sense, and decency, from an acceptable point of view, from a societal norm point of view (P10). P11 added to the sentiments raised by p10 above that the use of people's data becomes an ethical issue if the data is used incorrectly and without consent.

P12 equates the ethical AI concerns to that of Arthur C Clarke's novel, where the computer suddenly finds a conscience and wants to protect the mission, killing everybody on the ship except one. P12 fears that it may happen that we may believe that we create conscious AI, but we'll create a monster that is actually extremely conflicted. "AI probably doesn't have any beliefs, because it doesn't have any consciousness … so I really don't see artificial intelligence, being able to respond to its code of ethics" (P12).

"I think the first part is the black box of artificial intelligence, the notion that we don't actually understand how the algorithms work and on what basis algorithms make the decisions because algorithms have to use the information on which basis it learns. So they are the issues of it being a black box… Why did it choose to give a home loan to x and not to y sometimes the bank cannot actually explain the choice how the algorithm came to its decision". The second problem is the bias that is introduced by virtue of the information that it has, "whether it is biased on the basis of income, whether it is biased on the basis of age, whether it's biased on the basis of how you look, because of the images that it is being fed on which basis it is learning, it is only replicating what it sees in society" (P13).

P14 indicated that their concern is in the banking sector, where AI technologies are used for credit scoring. "You know, if you were banking and granting credit or something, it's a massive issue because there are ethical questions that need to be considered there. Secondly, it is the risk of job losses as a result of automation which would be a potential

ethical issue given South African high unemployment" (P21). P21 raised the concern about hardcoding pre-existing biases into AI technologies, resulting in perpetuating inequalities and systems, such as decision-making around job applicants, sentencing in courts, and financing decisions. P23 also mentions the replication of biases from the past into new systems, such as credit application processing in banks. P23 emphasises the importance of being aware of biases in historical data and ensuring the desired end result is achieved.

"Let's just take something like the whole cyber-crime, acts of unauthorised access to information that in itself poses a huge risk for organisations, in other words, how do we protect data that we collect" (P16). For example, shops like Pick n Pay and Dischem have data on smart shopper cards, how millions of people shop, and customers' personal data. There is a risk of an unauthorised person gaining access to such personal data and the unethical use of that data (P16).

P17's main concern was "who is checking the data that's been inputted into the algorithms… an example, if we're going to go out there and recruit for a suitable person, using AI technologies to do the recruitment, who checks that the questions in terms of what that search tool is going to do are ethical?". There have been cases where it's been shown that AI systems used for recruitment have been biased against women or in favour of the individual who went to particular schools or has certain education backgrounds.

P24 highlighted the dark side of technology where AI, intended for good, can be utilised for malicious purposes, such as deepfakes and cyber-attacks.

P24 called for ethical standards and regulations to prevent the misuse of AI.

Responses from specialists herein referred to as 'Participants (Ps)'.

Participants 1s, 7s, and 9s express concerns about ethical issues related to AI. Participant 1s highlights that bias in algorithms used in predictive analytics and autonomous decision-making can prevent individuals from accessing credit or healthcare. Moreover, job displacement caused by automation raises questions about how to upskill and retrain individuals who are no longer needed. The participants suggested that organisations prioritise their staff's welfare and consider the real impact of AI on people.

Participant 2s argues that ethical concerns arise when AI actions contradict societal definitions of ethics or are discriminatory, and it is crucial to guard against such potential biases. Participant 7s explains that the ethical concerns of AI are broad and encompass various areas such as machine learning, facial recognition, job displacement, the accuracy of outcomes, and AI applications. Participant (7s) gives examples of the Chinese government's deployment of facial recognition technology in public spaces to identify criminals and the potential biases resulting from using historical data to predict future outcomes.

Participant 9s draws parallels between the "trolleyology" thought experiment and AI, suggesting that remote AI may lead to a lack of moral and ethical accountability for its actions.

P15s raised a concern with the presence of unconscious bias in designing solutions using technology or artificial intelligence, which can lead to biased outcomes. For example, sifting through job applicant CVs using AI technologies can also lead to unconscious bias. P15s believes that this is a problem that will always exist as long as humans are the ones interpreting and writing the algorisms for AI technologies. P15s further expresses uncertainty about how the issue of bias can be resolved in the future (P15s).

P18s raises concerns about the sensitive information that individuals share with AI technologies and how that information is being used. They question the people behind the development of AI technologies and how they program the parameters of such technologies. For example, algorithms used by a bank to give loans to customers from disadvantaged communities, such as the black community, may not be ethical, leading to unfair treatment of certain customers. There is a need for ethical considerations in developing and using artificial intelligence (P18s).

P20s addresses the importance of ethics and laws surrounding the development and use of AI technologies. AI technologies have access to sensitive information, and it is critical to ensure that this information is being used in an ethical manner (P20s). P20s suggested that laws should be in place to prevent AI technologies from circumventing any laws that humans must follow, such as invasion of privacy or accessing unauthorised data. If AI technologies are engaging in activities that would be considered illegal for a human, it should also be illegal for AI technologies. In addition to legal considerations, P20s highlights the importance

of social norms and values that dictate acceptable behaviour. Social constructs and norms play a crucial role in determining ethical considerations in the use of AI technologies. AI technologies should be designed and developed in such a way that they can function within the social environment that it is intended to operate in. P20s added that ethics surrounding AI technologies should not only consider legal or social constructs but also moral considerations. For example, "AI should not show discrimination or bias even if it is not necessarily illegal. This is because AI does not have the same awareness and accountability as humans, and thus it is crucial to ensure that AI functions in an ethical and responsible manner" (P20s).

Analysis

There is no doubt that AI technologies have the potential to transform various industries and create new opportunities, but they also raise significant ethical concerns. The key concerns emanating from participants' interviews include the potential for biased decision-making, displacement of jobs, and privacy and security violations.

a) Bias

Almost all participants (board members) pointed to bias as the biggest ethical concern around AI technologies. P21 raised the concern about hardcoding pre-existing bias into AI technologies, resulting in perpetuating inequalities and systems, such as decision-making around job applicants, sentencing in courts, and financing decisions. The specialist participants raised the same AI technologies bias concern. There have been cases where it has been shown that AI systems used for recruitment have been biased against women or in favour of the individual who went to particular schools or has certain educational backgrounds (P17).

"The presence of unconscious bias in designing solutions using AI technologies can lead to biased outcomes" (P15s). Participant 1s highlights that bias in predictive analytics and autonomous decision-making algorithms can prevent individuals from accessing credit or healthcare (P1s). P15s believes that this is a problem that will always exist as long as humans are the ones interpreting and writing the algorisms for AI technologies. P15s further expresses uncertainty about how the issue of bias can be resolved in the future (P15s). The statement raised by P15s reflects a valid concern on the problem of AI bias in that it may persist as long as humans are the ones interpreting and writing algorithms.

135

AI technologies can be biased in various ways, such as reflecting the bias of their developers or being trained on biased data, which can result in discrimination against certain groups. There has been consensus that AI technologies may amplify and contribute to existing structural bias and lack of accountability and transparency, exacerbated by the fact that there is no one set of human values for it to adopt (Marwala, 2020). A human trains AI technologies on datasets containing human data; therefore, existing biases may be learned in the process and be displayed and amplified in real applications.

Researchers have pointed out that bias in AI technologies can stem from multiple sources, including biased training data (Barocas & Selbst, 2016; Boyd & Crawford, 2012), biased design choices (Selbst & Barocas, 2018) and biased interpretations of algorithms (Caliskan, Bryson, Narayanan, 2017). There has been well-documented evidence of bias, such as facial recognition (Buolamwini & Gebru, 2018; Khalil, Ahmed, Khattak & Al-Qirim, 2020), criminal justice (Wu & Zhang, 2016; Hamann & Smith, 2019), hiring processes (Datta, Tschantz, & Datta, 2014; Barocas & Hardt, 2017), and loan approvals (Crawford & Calo, 2016), among others. One additional key challenge lies in the fact that humans, as the creators of AI technologies, bring their own inherent biases, prejudices, and assumptions into the development process.

P23 emphasises the importance of the board being aware of bias in data and ensuring the desired end result is achieved (p23). The board should not only be aware of the bias in AI technologies, they should play an active role in addressing them. The board should put in place processes and structures to identify and understand AI bias, ensure the diversification and representation of data sources, evaluate and audit AI, and ensure transparency and human oversight.

b) Employment

Participants (including specialist participants) also raised job displacement as a major concern for AI technologies possible societal harm. They indicated that AI technologies may disrupt traditional forms of employment, potentially leading to job loss. Participant 6 mentioned that AI technologies have been used for customers calls in the telecom industry and it has reduced the number of call centre staff, which had an impact on jobs and productivity.

Rapid technological innovation holds a potential threat to employment, but this is not new in terms of technological evolution and as such dates back to the first Industrial Revolution. Technological advancements can impact employment in different ways. Acemoglu & Restrepo (2018) have provided a theoretical framework to understand the impact of new technologies on employment. They categorised the impact of AI technologies on employment into three broad effects: (i) displacement effect, by directly displacing employees from tasks they were previously performing; (ii) productivity effect, increasing the demand for labour in industries or jobs that arise or develop due to technological progress; and (iii) reinstatement effect where AI technologies can be a platform to create new jobs especially where labour has a comparative advantage relative to machines, boosting labour demand.

The combination of a rising need for employment and parallel displacement of jobs because of the automation and adoption of AI raises job fears in employees. As indicated by participant 6, the impact on displacement and productivity has already been felt in the call center industry due to AI adoption. What is clear from the participants is the acknowledgment that the impact on employment would be both positive (increase) and negative (job displacement). P8 indicated that AI technology can be deployed on "voluminous processes to achieve efficiency and speed". Deploying AI technologies on voluminous and repetitive tasks, can increase efficiency and productivity, and free up employees time to focus on more creative and higher-level task. Participant 8 noted that AI should be deployed "so that people can manage rather than do processing but focus on more strategic work or thinking work and then use their time efficiently". The same view was raised by P22s who indicated that "AI technologies can enhance decision-making and allow workers to focus on value-adding tasks".

The negative side which was raised by participants was the ethical concerns that AI technologies could be used as a tool to retrench people; AI technologies can lead to job displacement as machines and algorithms take over tasks that were previously performed by humans. This could lead to a decline in wages and employment opportunities. As suggested by participant 8, organisations "need to think of re-skilling the people so that they don't lose jobs because of AI technologies deployment" to reduce the impact of job losses and take advantage of the reinstatement effect by placing them in newly create new jobs. Assessment of the impact of AI technology on the future of work is crucial for the board to

mitigate job losses while increasing efficiency. AI technologies should work hand-in-hand with people to maximize their efficiency and performances overall.

c) Security & Privacy

The third ethical concern raised by both sets of participants (board members and specialists) was security and privacy in terms of issues such as cyber-security, data protection, mass surveillance, and acts of authorised access to information as a huge risk to corporations. The concerns reflected by participants are valid. AI technologies are processing vast amounts of data to learn and make predictions, which raises concerns about the collection, processing, and storage of such data. This data can include personal information such as personal identification numbers, names, addresses, financial information, and sensitive information such as medical records. Thus, these processes introduce new vulnerabilities and cybersecurity risks of exploitation by malicious actors who can gain unauthorised access to sensitive data or disrupt critical systems.

Some participants addressed the importance of ethics and laws surrounding the development and use of AI technologies. They suggested that laws should be in place to prevent AI technologies from circumventing any laws that humans must follow, such as invasion of privacy or accessing unauthorised data.

There has been calls for greater oversight of the AI technologies and to ensure appropriate safeguards are in place to maintain privacy. It is the responsibility of the boards to ensure establishment of AI governance structures and frameworks to ensure that the development and deployment of AI technologies are done responsibly, with a focus on cybersecurity and data protection. The board should ensure transparency and accountability, through governance structures that identify who is responsible and accountable for data processes during the different stages of the AI lifecycle and ensuring human oversight and controls (Kompella, 2022). Human control is also linked to accountability, and AI technologies should be built in such a way that humans can intervene in their actions or decision-making (Fjeld et al. 2020). The capability for human intervention during both the design and implementation, including the possibility to override a decision made by the AI technologies when it violates the law (Kumar et al.,2020), would ensure greater accountability of AI technologies.

7.4.3. The general role of the board

The third sub-question focused on the role of the board. The fiduciary duties of the board refer to the board of directors' legal and ethical responsibility to act in the best interest of the company and its stakeholders. The sub-question aimed to explore the board member's understanding of their general role regarding fiduciary duties. This question was asked because board members have a fiduciary responsibility to act in the corporation's and its stakeholders' best interests (including in terms of the adoption of AI technologies). To fulfil this responsibility effectively, the board members must clearly understand their roles and responsibilities.

Therefore, understanding the general role of the board as far as fiduciary duties are concerned is essential for effective corporate governance. It is essential for the board members to have a clear understanding of their responsibilities and obligations to fulfil their fiduciary duties effectively. This understanding can help the board members to make informed decisions and provide appropriate oversight, ensuring that the corporation operates efficiently, effectively, in trustworthy ways, and ethically.

The participant's responses to the question were as follows:

Responses from board members herein referred to as 'Participants (P)'.

Most of the participants P4, P5, P7, P8 and P14 indicated that the board's role is to provide strategic direction, guidance, resource allocation, setting the ethical tone, holding management accountable, evaluating performance, and monitoring and oversight.

The role of the board is "to deliver the shareholder's expectations in the business. They need to be held to account for delivery of shareholder's expectation because, in a lot of cases, the shareholders cannot do that themselves directly" (P6). The board has a strategic role in the long-term sustainability of the corporation (P6). The sentiments were also echoed by P11 who also added that the role of the board is "making sure the company survives and grows into the future". The board must establish governance structures such as an ethics committee and protect people and the environment (P11).

"The role of the board is to understand the processes and the return on investment used to measure executives' performance so that shareholder value is not eroded by unnecessary

spending" (P8). As a far as performing fiduciary duties the role of the board is to monitor management performance (P8).

"The board does have four key functions, the policy, the strategy, the overseeing of the executive, and the accountability reporting" (P12).

"The most important board's fiduciary duty is to set the tone at the top because that's where it permeates through to the rest of the business. The organisational culture and ethics start at the board. The role of the board is to ensure that all stakeholders are treated equally, not only your shareholders but all stakeholders including employees, and society at large" (P5)

P16 indicated that "in the complex world we live in, it is becoming more and more challenging to be a board member; let me start there because I think … the external environment in which we operate are becoming so complex, from a regulatory perspective from a risk perspective, with global events, global pandemics, et cetera, et cetera. So for me, one of the key roles of a board is to have, as part of their fiduciary duty and duty of care and skill and diligence, … a good understanding of the external environment within which this organisation operates and the risks and opportunities, we should not just focus on risks only, but also take advantage of the opportunities" (P16). P16 further indicated that one does not need to be an AI specialist to sit in the board but "you do need people on the board that have a basic understanding of data, technology and development in technology and how that potentially can impact your organisation".

The board is expected to act in the best interest of all stakeholders, not necessarily in the shareholders' best interest but in the company's overall interest (P21). P21 argued that "the board acts [as] an interface between management and the shareholders of the company, by which I mean, they sort of deal with the agency problem where management may have misaligned incentives with the owners of the company and so the board to a certain extent is sort of expected to act as a counterweight to that potential agency problem. But ultimately, it is in the shareholders, the director's obligation to act in the company's best interest over the long term" P21. P24 echoed this by stating that role of the board is to ensure that the relevant stakeholders are catered for, not only limited to shareholders, but including other stakeholders, such as the communities within which the organisation operates (P24).

"I mean, the general role of the board sets the strategy, is to protect the stakeholders, or is to make sure that all the stakeholders are treated equally, stakeholders such as employees, shareholders, communities where we operate; hence the boards are looking at the unemployment rate, you ask yourself in a company to say, what are we doing about that, it's no longer something that's for government to deal with but every company these days, is asking themselves because they now have an understanding that when the riots happen, what the impact of that is on their business, it's an issue that no government can deal with. And also like looking at the risks, looking at the ways in which services can be delivered to people in a cost-effective manner" (P22).

According to P23, the board is predominantly required to focus on the strategy of the business, incorporating the business purpose, the vision, and what the outcomes need to be. The board delegates these tasks to the CEO and management to develop, but the board is responsible for that strategy and measuring the performance of management, in terms of the achievement of that strategy, and incorporated in that is not only business results, but also how management executes. The board is responsible for promoting ethical leadership. The board is responsible for creating an enabling culture that attracts, develops and retains the right employees. The board is responsible for having a responsible mindset about the macro environment within which the company operates that stakeholder interests are considered and catered for, not just the interests of shareholders and profits, but considering the short, medium, and long-term implications of the decisions and the role that the company plays in broader society (P23).

Responses from specialists herein referred to as 'Participants (Ps)'.

"It is the top leadership of the company, and it's where the strategic planning and decision-making is done. It's where the interests of the shareholders are protected. But also, the well-being of the organisation should be protected at that level" (P1s).

The role of the board is to ensure that the corporation's affairs are aligned with meeting the interest of shareholders and other relevant stakeholders within that corporation's ecosystem (P2s). The role of the board is independent oversight of the management team. To maintain a role of independent oversights, the board cannot be too operationally involved unless there is some sort of governance issue that requires a board to actually step in at that level (P7s).

"Prof King always says that the organisation or company is headless, it's mindless, it's, and he calls it an incapacitated person because it's a juristic person as incapacitated, the Board provides it with the capacity, the capacity to capacitate, that organisation so the board determines the heart, mind and soul of the organisation" (P9s). This means that the board is accountable to the organisation to be its heart, mind, and soul. The King IV principle 6 recommends the role of the board as strategy, policy, oversight and disclosure. Core to the role of the board is stakeholder engagement, ethical, effective leadership, risk governance, performance, sustainability, and social responsibility (P9s).

The board also designs strategy and provides oversight on various components, such as governance structures that need to be in place (P20s). The board is expected to behave with integrity, competence, responsibility, accountability, fairness, and transparency (P18s).

Analysis

The findings on the third sub-question focused on the role of the board. The sub-question aimed to explore the board member's understanding of their general role regarding fiduciary duties. This question was asked because board members have a fiduciary responsibility to act in the corporation's and its stakeholders' best interests (including in terms of the adoption of AI technologies). To fulfil this responsibility effectively, the board members must clearly understand their roles and responsibilities.

The participants responses were based on six key roles of the board which were combined by the researcher into three key areas i.e., (i) strategic, effective control and oversight role; (ii) shareholder vs stakeholder view; and (iii) ethical leadership role.

i. Strategic, effective control, risk management and oversight role

Most of the participants indicated that the board's role is to provide strategic direction, guidance, resource allocation, setting the ethical tone, holding management accountable, evaluating performance, monitoring and oversight. Based on the findings, it appears that the board is perceived to have several crucial roles and responsibilities within the corporation.

- **Strategic direction** – The role of the board involves setting strategic direction for a corporation by ensuring realisation of the corporation's vision, mission, and

objectives, and that all decisions are aligned to these strategic goals. This also involves the board approving and monitoring short, medium, and long term strategy, thus taking care of risks and opportunities (King IV, 2016). The board delegates the task to develop the strategy to the executive management, but the board is responsible for the oversight on strategy execution and monitoring management performance. The board is ultimately accountable for the corporation strategy.

- **Guidance and advisory** – In today's fast-paced world and complex business challenges, the board is required to provide guidance to the executive management. This may involve offering expertise, insights, and advice to support effective decision-making. Therefore, it is important that the board is composed of members with the right mix of skills, background, knowledge, and expertise to ensure that the executive management is guided through a diversity (multidisciplinary) of thoughts in decision-making.

- **Holding management accountable** – The board has fiduciary duties to act in the best interest of the corporation and its stakeholders. This means that the board is required to oversee and hold executive management accountable for their actions and decisions and to ensure that they are aligned with the corporation's strategic objectives and long-term sustainability. The board is also involved in evaluating the performance of the executive management and, key company personnel such in the C-suites.

- **Monitoring and Oversight:** Participants acknowledge that the board's role includes actively monitoring the company's performance and overseeing its operations. As argued by P2s, the board needs to maintain independent oversight of the management teams. This is to ensure that the corporation remains on track and compliant with relevant laws and regulations. The board is responsible for strategic direction and management, and it typically delegates authority for the day-to-day operations of the corporation to the executive management. Once the board has delegated broad authority, its primary responsibility is to oversee management on strategic initiatives, financial performance, and the integrity of financial statements and accounting and financial reporting processes, risk management, and compliance with laws and regulations.

These findings highlight the board's critical role in providing direction, guidance, and accountability to the organisation. Their involvement in resource allocation, ethical

leadership, and performance evaluation further strengthens their position as key stewards of the company's success.

ii. Ethical leadership role

Participants recognised the board's responsibility in setting the ethical tone for the organisation. Participant 5 noted that "the most important board's fiduciary duty is to set the tone at the top because that's where it permeates through to the rest of the business. The organisational culture and ethics start at the board". The King IV Report (2016) which JSE listed companies are mandated to apply in terms of all its principles and code practices, states that the board must set the ethical tone from the top and provide ethical leadership. Ethical leadership highlights the ethical dimension of the board. It refers to the board values, ethical traits and ethical behaviour in the corporation, and the way such vales relate to stakeholders and the boarder society.

P18s indicated that "the board is expected to behave with integrity, competence, responsibility, accountability, fairness and transparency". King IV, Principle 1 sets the basis of ethical leadership as exemplified by integrity, competence, responsibility, accountability, fairness, and transparency also known as ICRAFT. While leadership starts with each individual board member, it finds its expression through the board as a collective, setting the appropriate example and tone, and ensures it permeates through to the rest of the organisation, as alluded to by P5. Accordingly, the board provides effective leadership based on an ethical foundation, which entails:

- Evaluating the effects of the company's strategy on the economy, society, and the environment, both in the short and long term.
- Conducting business with integrity and adherence to ethical principles.
- Preserving and safeguarding the natural environment.
- Acknowledging and considering business risks on both internal and external stakeholders.

Furthermore, the board builds and sustains an ethical corporate culture by determining clearly articulated ethical standards and values, which include a code of ethics and business conduct, and aligning internal and external ethics performance around the same ethical standards.

iii. Shareholder vs stakeholder view

Some participants argued that the role of the board it to look after shareholder interest. P6 indicated that the role of the board is to "deliver the shareholder expectation in the business". The argument was also supported by P1s who indicated that "… its where the interest of the shareholder is protected". However, most of the participants argued against this view. P5 argued that "the role of the board is to ensure that all stakeholders are treated equally, not only your shareholders but all stakeholders including employees, and society at large" (P5). The board is expected to act in the best interest of all stakeholders, not necessarily in the shareholders' best interest but in the company's overall interest (P21).

From the agency theory perspective (Berle & Means, 1932; Jensen & Meckling, 1976; Hillman & Dalziel, 2003), managers are expected to act and make decisions in the shareholder's interest. The shareholder approach (Donaldson and Preston, 1995; Fontaine et al., 2006; Wagner et al., 2012) means that the board does not consider other stakeholders as an instrument to serve the shareholder's interest, but rather that they have intrinsic value for decision-making in the best interest of the corporation. Thus, the board exercises oversight, supervision, and monitoring to ensure that management performs in the shareholder's best interest (Hillman & Dalziel, 2003). However, JSE listed companies have adopted the stakeholder-inclusive approach in line with the King IV report. According to the stakeholder approach the board should identify material stakeholders and their interests and expectations relevant to the corporation's strategic objectives and long-term sustainability. This implies that the board should consider the interests of all legitimate stakeholders, like employees and creditors, and not just those of the shareholders. The various interests of different stakeholders should be determined on a case-by-case basis, and the decision to act in the best interests of the company.

In chapter 3 of this study it was argued that incorporating a stakeholder-inclusive approach in the AI environment, would not only result in collaboration amongst various stakeholders during the planning, design, development, and implementation of AI technologies, but may also help address global challenges related to AI ethics. The study by Mhlanga and Molio (2020) found that corporations that embrace a stakeholder approach better serve society at large.

7.4.4. Role of the board in technology and artificial intelligence

The fourth sub-question focused on the board's role in technology and artificial intelligence. The question was asked to explore the board members' understanding of their role regarding technology, specifically AI. AI technologies have significant implications for any corporation, and, as mentioned already, the board members' understanding of these implications is essential in fulfilling their fiduciary responsibility to the corporation and its stakeholders.

The role of the board in overseeing AI technologies includes setting the strategic direction for AI adoption, establishing governance structures, ensuring that AI technologies align with the corporations' ethical standards and strategic goals, and providing appropriate oversight and risk management. Additionally, the board is responsible for ensuring that the corporation has the necessary resources and expertise to develop, deploy and maintain AI technologies effectively.

The understanding of the board's role as far as technology, specifically AI, is critical for effective corporate governance. By understanding their roles and responsibilities regarding AI technologies, the board members can make informed decisions, provide appropriate oversight, and ensure that the corporation operates efficiently, effectively, and ethically.

The participant's responses to the question were as follows:

Responses from board members herein referred to as 'Participants (P)'.

"In the old days, we had information technology (IT) in the back office, and these days, IT is now in the middle of business operations. It's certainly helping the business in many more respects than just pure support, such as running HR, payroll and accounting systems" (P4). The directors (board) need to become more aware of AI technology risks such as cybersecurity, privacy, the use of data, and big data, especially if that data falls into the wrong hands or is abused (P4).

P13 indicated that; "in terms of the strategic part of the board role when it comes to artificial intelligence. I think that all companies must recognise that if we're going to disrupt, we are going to be disrupted. So, from a strategic perspective, if you haven't invested in the digital transformation of the company, then you're not being relevant, which means you do not

have an understanding of the impact of machine learning, 3d printing, 5G, which could disrupt the business, or it could be a major innovation drive" (P13).

The board needs to have responsible use of technology framework in place, ensuring the appropriate use of technology and AI technologies, data protection, and cybersecurity. The role of the board is to ensure that the protection and usage of technology are done responsibly and adequately (P5).

The board has a governance, strategic, and performance oversight role in AI technologies (P6). However, P6 also argues that "many board members are clueless on AI globally", and they are responsible for approving huge capital expenditure (Capex) for AI technologies. In their (P6's) experience, executive management is more conversant on AI technologies, and in the board meeting, management invites specialists or experts to present proposals for the adoption of AI technologies. P10 argued that AI is an emerging subject, and there aren't sufficiently qualified people on the board to interrogate this sufficiently. P6 indicated that "… it's more like the board listens, told what is happening out there. And then, too afraid to embarrass themselves and ask questions they don't understand, they basically sign off the stuff that they don't understand, because of the threat of being left behind, because everyone is doing it… because they (experts) come from Google or Facebook or Amazon, then you like, this is a feature I must sign-off".

The boards need to be better trained on AI technologies, and there is a "huge gap with technology; many board members can barely use an iPad" (P6).

P10 further indicated that "the board must have one or more professionals on the board as well as a specialised committee of the board, dealing with information technology that is sufficiently aware of the dynamics, complexities, implications and outcomes of using the technology. So, it is imperative that if the board is pursuing a technology-driven agenda, one or more board members must be familiar with the technology so that they could ask and participate in robust question making" (P10). AI emergence has only taken precedence over the last 10 to 15 years, and the people that have insights are considered to be a bit young to be appointed to the board. Therefore, the board needs to have a structure going forward for alternate directors so that people with skill sets and experience can be appointed as alternative directors to the primary director in the board meeting to make inputs on artificial intelligence, ethics, and governance (P10).

In providing that strategic oversight, the board must embrace technology and be aware of what technology can do for the organisation while appreciating and understanding the risks and opportunities involved with using technology (P17). The role of the board in technology and information governance is to oversee investments in technologies, protection of information, and ensuring that technology is used in a manner that creates value and not in a manner that destroys value. The board must set policies and integrate technology or digitisation into a business strategy. The policies and strategy will guide the organisation's use and implementation of artificial intelligence technologies (P7). The board needs to have a framework for the responsible use (and adoption) of technology in place, ensuring the appropriate use of technology and AI technologies, data protection, and cybersecurity. The role of the board is to ensure that the protection and usage of technology are done responsibly and adequately (P5).

According to P12, at the policy level, the board need to be clear on the direction it wants to take regarding AI. The board is responsible for implementing these policies and the strategy, which is easier if the policies are well defined. However, the biggest challenge is that there are so many unknowns with AI technologies. The overseeing role of the executive is probably the easiest because it won't be different from what it is today, except that the power of AI may be even more discriminant in appointing our executives. Openness, accountability and the reporting role of the board would be challenging: "how can you be accountable for something that's unpredictable?" (P12).

The fiduciary responsibility is setting strategic direction, governance, and risk management, including financial and ethical risk management. In terms of technologies, the board is responsible for oversight and risk management. Part of that oversight and risk management is to ensure transparency and that AI is not biased, which can cause reputational risk. We cannot explain why a loan was given to x and not to y because there were no right checks and balances from the board on algorithm decision-making on behalf of the company (P13).

Technology is a major cause or source of potential risk for a board, and a board has responsibility or oversight and risk function. Most boards have risk committees, sometimes combined with an audit committee, but every company should have some form of a risk committee. Technology is obviously a major potential cause of risk as a result of increasing cybersecurity concerns and reliance on technology. And with respect to AI, it depends on

the industry, so the role of the board is to be alert to issues, always to be alert to potential ethical issues (P14). P14 gave an example of an asset management company in which AI technology is a massive issue because of the amount of data in play, and the fact that the data has to be accurate. So, in this company, the board had an IT committee which was composed of executives and board members. The committee reported to the risk committee and the board. The committee required more skills, at least two more IT expects, but the company did not want to increase the size of the board. So, they formed a specific committee that is answerable to the board (P14).

The board has a responsibility to remain aware and educated about emerging technologies and to understand their implication, both potential positive and negative implications of AI technologies to their business (P21). As a starting point, "the board must be aware of emerging technologies, understand their application, understand their usage, understand the potential risks associated with their deployment, which may include bringing in outside experts to provide that kind of insight" (P21). P21's view was also supported by P22 who stated that "the role of the board is to keep themselves abreast of the new technologies and how they affect their businesses because if as a business, you don't do that, you will be disrupted" (P22).

P23 argued that the board is wholly responsible for whatever it uses to enable the strategy; therefore, as it is wholly responsible for its employees, it is wholly responsible for technology. Technology is an enabler of business objectives and outcomes, and the board is responsible for the technical decisions. In terms of the broader decision-making of the board, the board has to ensure that it understands the digital revolution and innovation. The board must understand the role of information and technology in terms of the best way to deploy the best outcomes for the business, how to best outsource and insource and stay abreast of new technologies like AI, and in terms of risks like cyber security. The JSE-listed companies have a mechanism for managing, at the board level, the detail associated with IT decision-making. P23 indicated some boards delegate responsibility to "the audit and risk committee. Some do it through a separate risk committee, and some do it with a separate IT Committee. It doesn't really matter how each company needs to assess its level of risk, but it needs to know that at a board level, it cannot shy away from its responsibilities in focusing on these very important assets and capability for the company" (p23).

AI technologies' potential risk fall under ICT risk mitigation processes, for the most part, depending on the nature of the company. Generally, from a governance structure perspective, it will fall under the audit committee as part of the audit committee mandate or be separated as its own committee of the board. Depending on the level of risks, AI can be discussed under an audit committee, but where risks are extreme or catastrophic or potential business continuity disruptors, the board needs to have a separate committee to deal with it (P24).

P16 indicated that for thirteen (13) years, they had been a member of the King committee and very much involved in the writing of the King IV Corporate Governance Report. One of the key positive highlights of King IV was specifically the section on technology and information. The section unpacks the board's oversight role in terms specifically of a technology perspective, responsible use of technology, and disposal of technology. King IV also gives guidelines for oversight of how to manage information, cybercrime, and social media risks, etc. (P16).

P19 also referred to the King IV reports, which requires the board to be responsible for technology. Most boards delegate to the Audit Committee, which obviously doesn't mean the rest of the board isn't responsible, but the board is responsible for IT Governance. The board must be aware of what management is doing to ensure the controls are in place. On AI technologies, the board needs to understand what the company is doing in that realm, which includes issues such as risks, including ethical and reputational, cybercrime, disaster recovery, and continuity planning, that fall within and outside of IT (p16).

P8 did not fully understand how the board play a role in AI.

P10 indicated that they had been on the board of the company for a longer time, but stated they never dealt with AI. They nonetheless believe that the social and ethics sub-committee of the board would be the right platform to govern AI technologies (P10).

Responses from specialists herein referred to as 'Participants (Ps)'.

"The ownership of that technology should sit at a board level, not at some technical level, somewhere down the hierarchy, the fiduciary responsibility, the legal responsibility, the potential reputational responsibilities, the ownership of this tech must be at board level"

(P1s). Technology will impact the future of corporations; it will greatly impact the staff with potential job displacements and how they interact with customers through digital channels. Therefore, the board should have a good enough understanding, "even if it's not a technical expert understanding", but everyone on the board, irrespective of their role, should have a good enough understanding about the impact of this technology so that they can make the right decisions (P1s).

According to P3s, the role of the board is to ensure appropriate governance of technology, including its use and application. The board needs to ensure the ethical use of technology, particularly AI and ensure that it is free from bias. The board is responsible for overseeing the methodologies and processes of AI to ensure it is used ethically. The board needs to keep regular board practices and structures of good governance in place, such as considerations around materiality, governance, ethics, and professional development. The board needs to understand the technology and its capabilities to approve requests for budget and resources for leveraging these technologies in the organisation (P3s). P3s further indicated that the board needs to have a sound understanding of concepts like data management, data governance, data protection, and data privacy, particularly when the corporation is working with large amounts of data. The board needs to have a general grasp of how technologies are used and their risks and benefits. The board's role is to mitigate risks and recognise opportunities to leverage these technologies in the corporation. The board can call on relevant experts to inform them and ensure they have a sufficient understanding of the technology (P3s).

P6s argues that the board is accountable for everything that the organisation does, including strategy oversight and policy development, while the responsibilities to implement strategy lie with the executives. However, the board's accountability to an organisation is often overlooked and not effectively implemented. According to ISO 37000[40], the board's role includes purpose, values, and value generation, as well as enabling activities such as stakeholder engagement, ethical and effective leadership, risk governance, sustainability, social responsibility, and data-driven decision-making. The board's responsibility is to govern the organisation, and it must ensure that the organisation operates in accordance with its purpose and values while generating value for stakeholders (P6s).

---

[40] ISO 37000:2021is a Governance of Organizations - Guidance issued by International Standard on Governance of Organizations

P18s argues that the "board is there to consider the negative and the positive impacts and potential harm and potential benefit that the organisation can benefit from this from technology and artificial intelligence, but also the potential harm that it can bring to the organisation and its stakeholders and the public". The board's role is to decide if the technological tools can help reach the strategic objectives, the ethical use, potential risks, and a responsible way to use technology (P20s).

According to P15s, the board should evaluate AI technologies based on whether they will benefit the company, customers, and other stakeholders, rather than simply buying it for the sake of having new technology. P15s cautions against ego-driven decision-making and suggests engaging with executives and independent consultants to ensure alignment and to evaluate options and their potential business impact. The focus should be on how AI technologies can help drive the company's strategy faster and deliver tangible benefits to all stakeholders.

P2s argues that technology is no longer just a supporting department in a corporation but closely linked to the entire business model. Technology is now a key factor that affects how everyone in the corporation carries out their duties, and it should be thought of more strategically beyond just being allocated to a specific department that supports the organisation. The board should carefully consider the impact of technology on the corporation in a wider context.

Analysis

The findings reveal the role of the board in technology and artificial intelligence as mainly setting strategic direction and ensuring governance structures and processes are in place in terms of raising AI awareness and training, and risk management and oversight.

(i)   Strategic role

Most participants among both the board members and specialist groups believe that the board is responsible to oversee tools and processes which enable the company strategy; therefore, as it is wholly responsible for its employees, it is wholly responsible for technology. Therefore, the board is accountable for everything that the organisation does, including strategy oversight and policy development, while the responsibilities to implement strategy

lie with the executives. The strategic board involvement should cover mission development, strategy formulation and strategy implementation (with the executives). The strategic board involvement is viewed to be important from a practical viewpoint as it offers board reform activists a basis to delineate the sphere of the directors' potential role in strategy.

The King Code IV (2016) states that the board should govern technology to support organisations in setting and achieving strategic goals. Technology governance offers valuable guides to operationalising AI governance. However, because of their learning capabilities and dynamic evolution, AI technologies should initially be viewed as special cases to IT systems that require distinct governance tools compared with general IT governance. As AI governance matures, it may eventually be subsumed under IT governance.

The board needs to ensure that technology governance aligns with the corporate strategy (Weill and Woodham, 2002). Van Grembergen, De Haes & Guldentops (2004) argued alignment between technology strategy and organisational strategy is crucial for technology to deliver business value. Technology strategic alignment seems to be a critical element of ensuring that technology investment adds value to the organisation. Thus, it is imperative that technology supporting business objectives is not only aligned but integrated into the strategy. One of the main reasons for the lack of return on investment is the lack of alignment between technology strategy and business strategy. Furthermore, lack of alignment may also lead to information silos, bad business management, and failure to gain business value from technological investments.

(ii) Board structures and processes

The board must set policies and integrate technology or digitisation into a business strategy. The policies and strategy will guide the organisation's use and implementation of artificial intelligence technologies (P7). The board needs to have a framework for the responsible use of technology in place, ensuring the appropriate use of technology and AI technologies, transparency, data protection, oversight and responsibility, and cybersecurity. The role of the board is to ensure that the protection and usage of technology are done responsibly and adequately (P5). P10 further indicated that "the board must have one or more professionals on the board as well as a specialised committee of the board, dealing with information technology that is sufficiently aware of the dynamics, complexities, implications and

outcomes of using the technology. So, it's imperative that if the board is pursuing a technology-driven agenda, one or more board members must be familiar with the technology so that they could ask and participate in robust question making" (P10).

The role of the board is to ensure appropriate governance of technology, including its use and application. The board needs to ensure the ethical use of technology, particularly AI and ensure that it is free from bias. The board is responsible for overseeing the methodologies and processes of AI to ensure it is used ethically. As outlined in the AI board toolkit on empowering leadership (WEF, 2020), boards need to keep regular board practices and ensure structures of good governance are in place, such as considerations around materiality, governance, ethics, and professional development. For example, the board should have the following structures and processes in place:

- **Board oversight** – Through the establishment of board committee(s) such as AI ethics board or social and ethics committee, audit and risk committees, AI committee, etc.
- **Risk assessment and opportunities** – The board should assess the risks associated with AI implementation and ensure that proper risk management processes are in place. This should involve identifying potential risks, evaluating their potential impact, and developing strategies to mitigate these risks and take advantage of opportunities.
- **Monitoring, audit and response** –The board should ensure that the AI initiatives adhere to ethical standards and responsible practices. The board should establish guidelines for the design, development and deployment of AI technologies that consider factors like fairness, transparency, accountability, and human-centered evaluating processes to ensure compliance and ethics assurance programmes and the response to allegations of violations.
- **Reporting to shareholders** – as required by law, ensuring reports to shareholders and filings to regulators include information about the risks of AI. They should also report on the use of AI in financial reporting and auditing to verify the accuracy of financial statements.
- **Stakeholder engagement** – The board should ensure that the interests of stakeholders, including shareholders, customers, employees, and the broader community are respected. Therefore, the board should always engage with these

stakeholders to ensure that AI governance aligns with their expectations and concerns.

- **Training** – The board should ensure training programmes are in place for all employees, outside contractors, and other third parties. Moreover, training programmes must also be tailor-made for board members themselves to fully full comprehend AI technologies.

Such a proactive involvement of the board in strategic planning, policy formulation, risk management, oversight of ethical practices, resource allocation, monitoring, stakeholder engagement, and maintaining alignment with company values would ensue that the board navigates the complexities of AI technologies while promoting responsible and beneficial outcomes. Therefore, the boards have a responsibility to ensure that the corporations have effective and efficient governance structures in place for setting objectives and monitoring performance. That structure must provide for how the board, management, shareholders, and other stakeholders participate in decision-making, and ensure the board has the information required to provide oversight, and that shareholders have information about the strategy, risks, and performance of the corporation.

(iii) Risk management

Some participants indicated that technology, and AI in particular, is a major cause or source of potential risk for a board, and a board has responsibility or oversight and risk function. The directors (board) need to become more aware of AI technology risks such as cybersecurity, privacy, the use of data, and big data, especially if that data falls into the wrong hands or is abused (P4). The board needs to have a general grasp of how technologies are used and their risks and benefits. The board's role is to mitigate risks and recognise opportunities to leverage these technologies in the corporation. The board can call on relevant experts to inform them and ensure they have a sufficient understanding of the technology (P3s).

Part of that oversight and risk management is to ensure transparency and that AI is not biased, which can cause reputational risk. As already argued above, the board should assess the risks associated with AI implementation and ensure that proper risk management processes are in place. This should involve identifying potential risks, evaluating their potential impact, and developing strategies to mitigate these risks and take advantage of

opportunities. This implies that the board should be knowledgeable enough about AI technologies and have a perspective on critical issues such as AI ethics, risk, and governance. The board should keep abreast of the impacts (and potential) of AI technologies to avoid shocks to the corporation and to be open to the impacts that may reasonably be expected to arise during the course of executing their duties as directors.

(iv) Training and awareness

In providing strategic oversight, the board must embrace technology and be aware of what technology can do for the organisation while appreciating and understanding the risks and opportunities involved with using technology (P17). The board has a responsibility to remain aware and educated about emerging technologies and to understand their implication, both potential positive and negative implications of AI technologies to their business (P21). As a starting point, "the board must be aware of emerging technologies, understand their application, understand their usage, understand the potential risks associated with their deployment, which may include bringing in outside experts to provide that kind of insight" (p21).

The second fiduciary duty is the duty of care, and it requires that the directors be attentive to details when making decisions on behalf of the corporation (Veasey, 1997). The duty requires directors to act with care and diligence when performing their responsibilities. In other words, directors are required to take reasonable steps to inform themselves before making business decisions. This implies that the boards have a responsibility to take reasonable steps to gain deeper insights into the impact of AI technology in the corporation. Exercising care, skill, and diligence means that the board, individually and collectively, need to constantly obtain and uphold adequate information and understanding to execute their oversight and control mandate effectively. Even if AI technology is being used in boardrooms themselves, it does not replace the board's responsibility to exercise due care and loyalty, to be fully informed and make its own decisions in the best interests of shareholders and all other stakeholders. Therefore, the boards need to be better trained on AI technologies, to close the knowledge gap as indicated by P6.

7.5. Research sub-theme 2:

The second research sub-theme: What is the board's role in ensuring that AI technologies are aligned with the organisation's ethical culture? The research sub-theme included three questions posed to the participants: (1) What is your understanding of how the board plays a key role in setting the ethical tone in the organisation?; (2) What is your understanding of how artificial intelligence technologies may impact on the ethical culture of the organisation?; and (3) What is your understanding of how artificial intelligence technologies strategies are aligned to the organisation's ethical culture? The purpose of the sub-theme was to explore the board's role in ensuring that AI technologies are aligned with the organisation's ethical culture.

The questions were asked because the ethical implications of AI technologies are becoming increasingly important and a significant concern for organisations, and the board has a critical role in ensuring that AI technologies align with the organisation's ethical standards and culture. Overall, this sub-themes of three questions aimed to explore different aspects of the board's role in aligning AI technologies with the organisation's ethical culture, including the board's role in setting the ethical tone in the organisation and understanding the potential impact of AI technologies on the organisation's ethical culture.

7.5.1. The role of the board on setting the ethical tone

The question "What is your understanding of how the board plays a key role in setting the ethical tone in the organisation?" was asked as part of the second research theme sub-question "What is the board's role in ensuring that AI technologies are aligned with the organisation's ethical culture?".

In the context of AI technologies, the board's role in setting the ethical tone is particularly important. AI technologies have the potential to impact society in significant ways, and their development and deployment must be guided by ethical considerations. The board must ensure that the organisation's ethical standards are integrated into the development and deployment of AI technologies. By understanding the board's role in setting the ethical tone of the organisation, board members can provide appropriate oversight, guidance to management to ensure that ethical considerations are taken into account in all decision-

making processes and mitigate potential ethical risks. The participant's responses to the question were as follows:

Responses from board members herein referred to as 'Participants (P)'.

"In setting an ethical culture, mainly comes from the tone from the top, starts with the chairman, the CEO, being able to abide by the ethics of the company and the code of conduct of the company, and observing the conflict of interest because that's where most of the ethical dilemmas come where conflict is either known or unknown" (P22).

P5 indicated that the ethical tone comes out when there are grey areas, where the board then needs to clearly state the ethical terrain on what is acceptable and not acceptable (P5). "People watch how the bosses behave and will mimic what happens at the top…so the board has always been responsible for setting the tone and ensuring that it cascaded through the organisation for everyone to have the same mindset and view" (P4). The board is responsible for setting the tone at the top as required by the JSE listing requirement and the King IV (P8).

The board plays a crucial role in setting the tone for ethics in an organisation. They do this by implementing policies such as a code of conduct, monitoring ethical conduct, and establishing a reward system. It is important for the board to ensure that their policies are clear on addressing potential unethical issues and deliberately taking ethical steps, such as reaching out to vulnerable populations and making technology accessible. The board's responsibilities involve preventing unethical conduct and setting a clear direction for the organisation to implement deliberate steps towards ethical behaviour, such as sharing technology and making it user-friendly for those who may not have had access before (P7).

"I think it is very critical to set the tone at the top, is the whole aspect of saying the fish rots from the head, so if governance structures are not as effective as they are intended to be, then it permeates through the organisation because the oversight role is lacking. To the extent that the potential gap is not dealt with head-on, it may lead to unhealthy corporate, which then pulls even bigger risks" (P23)

The role of a board is to give the right message across and walk the talk regarding the values and the organisational culture they want to see. The board needs to have proper

consequence management to deal with unethical behaviour, not a "mindless tick box approach" (P16). P16 indicated that in their thirty years' experience, some organisations who had "made all the right noises with a beautiful, framed picture of the value are just window dressing with tick box mentality" (P16).

Defining individual board members' beliefs, values, and ethics is crucial in establishing a clear set of values for the company. Without this, there may be value differences, leading to failure. It is important to go beyond slogans and posters on the walls and establish fundamental beliefs and values. The boards are professionals chosen for their values; part of their values lies in their ethics and beliefs. Therefore, defining these values is among the significant differences between boards in the 20th century and today (P12).

P13 argued that "the board does not set the ethical tone; the ethical tone gets set by the management team". The board should not be seen as operational; the tone must be driven by the CEO and his or her leadership team. However, the board must be very clear about its expectations. The board's mandate and its charter must be very clear on where it stands on matters of ethics inside the organisation and what it will tolerate and not tolerate amongst itself as board members, but also in terms of how it sets the tone for the organisation. So, it is a supportive role to the executives and its leadership team (P13). P19 agrees with the sentiment above that the board should delegate the function of ethics, but it should be very clear in how the company sends the ethical message. They added that the board charter should cover ethics, while the audit or social and ethics committee should deal with incidences of unethical behaviour and be responsible for how ethics is rolled out throughout the organisation. It is not only setting the tone from the top but also the oversight role (P19).

The board will not be able to set the tone for the entire organisation, but the board certainly sets the expectation in relation to the CEO, executives, and board members. For example, "setting risk tolerances will give an indication of what is acceptable behaviour. The board is responsible for selecting the CEO and the extent to which ethical considerations feed into that selection process. The board oversees these core values and value statements concerning the business (P21).

The board culture is impacted by the ethical tone of the board, and it is critical that the board is aware of and is alert to its own ethical culture in the board. For example, does the company

board have codes of conduct, conflict of interest policies, ethics, and whistle-blower policies in place (P17).

The role of the board is to ensure that the company is operating ethically and to understand the impacts on society and the environment. The board needs to ensure that there are no conflicts of interest, full disclosure, and no attempt to try undue influence. Additionally, the board must ensure that there are structures in place for whistleblowing (P14).

Responses from specialists herein referred to as 'Participants (Ps)'.

"The role of the board as the highest decision-making body in an organisation can never be over-emphasised in setting the culture right. So that it flows down and to consistently reinforce the commitment to these values in line with organisational mission efficiency" (P2s)

"The board is vital to setting the ethical culture, and if you look at King IV, principles 1,2 and 3, it is talking about leadership, ethics, and corporate citizenship. So, Principles 1, 2, and 3 actually talk about how the board is responsible for setting an ethical culture and also being responsible for being a good corporate citizen in the organisation. To governance ethics, it starts on the board and cascades down into the organisation" (P18s). P3s also emphasised the importance of the King Report in governing and promoting ethical, sustainable, and transparent corporate culture. P3s argues that the board of directors should strongly understand King IV and apply its principles through a social and ethics committee. This committee plays a crucial role in discussing and addressing ethical, cultural, and compliance matters. Culture can pose risks to the organisation, and the board needs to track and assess these risks, with relevant metrics reported to the risk committee and ultimately to the entire board (P3s).

The board holds a key responsibility in shaping the values and culture of the organisation. The board's role is crucial in aligning business strategy with the desired values to be instilled in the organisation. The organisation's ethical approach must align with the general values shared by all humans and the specific values of the areas in which they operate. P1s highlighted the importance of ensuring that the culture goes beyond superficial appearances and is driven by meaningful and impactful values rather than simply aiming to look good in the public eye (P1s).

Establishing and maintaining a culture within an organisation is challenging, and the board holds accountability for that culture. For example, the Volkswagen scandal, where unethical practices were allowed, indicates that such culture starts from the top. Creating a culture requires significant effort, and there is no guarantee of getting it right, as people have their own vested interests and individual values and morals. Merely displaying values on the wall is insufficient; they must be actively lived and ingrained in the organisation (P9s).

P20s states that the board should set the ethical tone for the company and align it with the strategic direction of the organisation. The board is responsible for establishing ethical guidelines or creating platforms for discussions and debates about ethical practices. However, P20s also acknowledges that boards sometimes neglect their ethical responsibilities, as it may not be a legal requirement for them to do so (P20s).

"I think it goes back to what your vision and your mission is and the principles within a company are based on, that will drill down into the capabilities that you are going to pursue. So does it align with vision, mission and values" (P15s).

Analysis

In agreement with the findings in the previous section, the findings in this section reveal that almost all participants agree that the board has a role to play in setting the ethical tone of the organisation. The board sets the tone (an ethical or unethical) from the top and has a trickle-down effect on lower-level employees. This means that, if the board upholds ethical standards and integrity, such will permeate down to employees. But if the board appears unconcerned with ethics and focuses solely on the bottom line, it may lead to unhealthy organisational culture (Doug Cornelius, 2009), which then poses bigger risks, and employees will be more likely to commit fraud and feel that ethical conduct isn't a priority.

Central to the overall ethical environment of the organisation is that the tone from the top should be set by example and action by the board. For example, the tone from the top starts with the board members, the CEO, being able to abide by the ethical standards of the corporation and the code of conduct and noting any conflict of interest because that's where most of the ethical dilemmas occur. As outlined in the King IV™ report, the board of directors sets the tone for any ethics and compliance program in a corporation. King IV™ explains that the board must govern ethics to ensure that the ethical culture within a corporation

aligns with the tone set by the board. As noted already in the previous section, according to King IV, the way to achieve this outcome is for board members to cultivate integrity, competence, responsibility, accountability, fairness, and transparency (ICRAFT). Underlying this analysis is the critical assumption that ethical behaviour, especially on the part of board, leads to ethical culture and the best long-term interests (reputation and its relationship with all stakeholders) of the corporation.

Establishing and maintaining an ethical culture within an organisation is challenging, and the board holds accountability for that culture. The board needs to have policies in place such as a code of conduct, disclosure of interest, monitoring ethical conduct, and establishing a reward system. It is important for the board to ensure that their policies are clear on addressing potential unethical issues and deliberately taking ethical steps. The board needs to ensure that there are no conflicts of interest, full disclosure, and no attempt to try undue influence. As suggested by P20, the board should establish ethical guidelines or creating platforms for discussions and debates about ethical practices.

P13 argued that "the board does not set the ethical tone; the ethical tone gets set by the management team" as the board should not be seen as operational. However, contrary to the view expressed by P13, the board should set the tone by exemplary leadership, establishing governance structure and putting in place policies, and hiring a CEO. The board will thus delegate the oversight responsibility of ethics and compliance to the risk and audit committee, and social and ethics committee. The implementation of ethical culture will be delegated to the CEO as the face of the corporation, the figurehead to whom employees ultimately look for vision, guidance, and leadership. The tone at the top should demand that the board and especially the CEO should find ways to connect all stakeholders and openly communicate their values on an ongoing and transparent basis. Lastly, the tone at the top should ensure that the board should embody and not merely pay "lip service" to compliance and upholding ethics but should be honest, show integrity, and uphold an ethically-correct corporate culture.

The overall comprehension of the board's role in setting the ethical tone of the corporation by participants means that board members can provide appropriate oversight and guidance to management to ensure that ethical considerations are taken into account in all decision-making processes. Such understanding may assist to mitigate potential ethical risks

associated with the development and deployment of AI technologies and promote long-term organisational success.

7.5.2. AI technologies' impact on the ethical culture of the organisation.

The second sub-question "What is your understanding of how artificial intelligence technologies may impact on the ethical culture of the organisation?" was asked as part of the second research theme sub-question "What is the board's role in ensuring that AI technologies are aligned with the organisation's ethical culture?". The question was asked to explore the board members' perception of the potential ethical implications (challenges and opportunities) that AI technologies may have on the organisation's overall ethical culture.

Understanding the potential impact of AI technologies on the organisation's ethical culture is crucial because it enables the board to proactively address ethical concerns and ensure that AI technologies are developed, deployed, and utilised in a manner that aligns with the organisation's ethical values and standards. It also highlights the need for the board to provide guidance and oversight to management in navigating the ethical complexities associated with AI technologies.

Responses from board members herein referred to as 'Participants (P)'.

"I think there is this tendency to believe that if a computer does decision-making, it is somehow objective, and the extent to which we devolve decision-making to AI technology, we need to be careful that we are not evolving our ethics to the AI as well, or that we are saying we are comfortable that our ethics and our ethical standpoint is embedded in the AI and its application. Because I think the danger is or the concern for me would be we run the risk of abdicating our ethical responsibility by not thinking through the implications of certain decisions being left to artificial intelligent technology" (P21).

According to P6, there are two main perspectives regarding the impact of AI technologies on organisational culture. From a positive standpoint, if AI technologies are correctly deployed, they can contribute to developing a high-performance learning culture. The rapid advancements and changes in AI technologies necessitate a combination of performance improvement and a willingness to learn from everyone involved, which can lead to

organisational success. However, there is also a negative impact on the culture to consider. The proliferation of sensitivities and rules, such as data privacy regulations like the POPI Act, can significantly impact how the organisation operates. Concerns about potential data leaks can make people overly cautious and prevent data sharing between departments, limiting the organisation's ability to effectively leverage business intelligence (P6).

In terms of the working of AI technologies, rules are set and applied without fail, which is good, but on the other hand, the human element is required to assess the situation and make their own decisions. AI technologies may not take into account the need of the other party and its current circumstances (P4).

Allowing the systems, processes, and algorithms to make decisions without testing them regularly for bias is problematic. Companies should adopt an annual board IT testing framework which can be used to check cases of biases continuously, and the board can play an oversight role quarterly or annually (P5).

There is a level of leadership required to ensure that the organisation is prepared for the change "because when you implement artificial intelligence, there is a change element involved as you are bringing new ways of doing things there will be fears, resulting in resistance to the change; fears, resulting in people being anxious". Thus, the board needs to play a very important role in ensuring that the organisation is prepared for the change regarding AI (P7).

"Since organisational culture is how people behave and do things, I do not see how AI impacts culture because an AI is a tool is an enabler for efficiency, achieving or setting targets such as customer satisfaction or achieving certain efficiencies. So, I don't see how it can affect culture. The only way I can see AI affecting the culture is when AI technology deployment becomes part of KPIs" (P8). P12 agrees with the argument above, and they state that "I don't think you'll have an impact because in the sense that artificial intelligence is not what is going to give the direction on the culture of a company. It's only a tool, it's a thing, it's going to become an amazingly powerful tool, but it's an only tool, not the brain" (P12).

"Artificial intelligence and machine learning, and all sorts of digital technologies are seen purely as a technology, and actually not seen as something that changes the entire fabric of

our business. But it does; it changes how people operate, the type of work we do, and the ethical dilemmas we face. It changes the responsibility that every individual has to keep the organisation safe from a cybersecurity perspective. It changes what training we give people, and it is likely to change where our reputational gaps might be. So, it's not a technology solution; so, if the board sees artificial intelligence or digital technologies purely as technology solutions and not changes in business models, and changes in the way that an organisation works, then they are underestimating the impact of artificial intelligence" (P13). The board still sees AI as a pure technological investment; it is still discussed under the IT or finance committee, but it is not discussed under the social and ethics committee, which would be concerned with the impact of ethics or ethical dilemmas. "The social and ethics committee deals with issues of reputation, they deal with issues of ethics, but that's not where digital transformation or artificial intelligence gets discussed. Because people don't, at this point in time, think about the technology implications or artificial intelligence implications on ethics or the ethics of technology. They only think about technology as an investment" (P13). Moreover, different subcommittees need to discuss the impact of AI technologies. The board members and the executives should have a strategic discussion of the impact of technology beyond just a cost conversation; it should be part of the board's agenda, like a risk register (P13).

P16 emphasised the need to avoid a solely risk-averse mindset when it comes to utilising AI technologies. They further argued the importance of recognizing that AI technologies will shape the future workforce, and that organisations should proactively focus on reskilling employees. The participant highlights the existing challenge of unemployment, which is exacerbated by AI's potential to replace jobs, and suggests that businesses should take responsibility and consider how they can address this issue, as this has a huge potential to impact organisational culture (P16).

P17 argued that the organisation must establish the purpose of using AI technologies and have checkpoints to check the data used to develop such a technology. "For example, if the organisation wants to have a culture of promoting people from within the organisation, and you create the AI tool to capture the data of all your employees to then work out who's been here the longest to who's doing what particular roles, and then you let the AI tool say, these are the following people". However, building such a tool, the company should guard against bias such as qualification and years of experience; thus, organisations need to be very

careful about using AI alone in that context unless it is adaptable to account for those elements (P17).

The organisation's culture may be impacted if someone is selling data inappropriately, which means that the deployment of those AI may also impact the culture because employees may be using the same system to collect data and still sell them to the next unauthorised person (P19).

P22 argued that if AI processes are not set up properly, the decisions that will be taken could be dire to the organisation. If the system can be hacked, the organisation runs the risk of outsiders controlling what decisions can be taken by these processes that the organisation has implemented. Thus, the board should understand how the organisation can be secured from a cyber risk point of view and that decision-making processes are not left to AI technologies (P22).

From an organisational cultural point of view is about understanding the role of securing people's data and information. It is critical when it comes to a culture of protection of information and even within teams or units within the organisation and sharing data or leveraging data that may impact an organisation's culture, especially now with the POPI act in place (P11).

Responses from specialists herein referred to as 'Participants (Ps)'.

Specialist Participant 3 (P3s) argues that the impact should be considered from two angles, risk, and opportunities. P3s highlighted that the application of artificial intelligence, in some instances, could influence the external perception of an organisation's culture. For example, an organisation using artificial intelligence in an application perceived to be highly controversial or very biased has a reputational impact on the organisation. From a human resource perspective, "it's going to potentially have a massive impact with people who suddenly become very fearful of it because they think they might become redundant, that the technology might replace them, so then it introduces considerations and cultural debates" (P3s). Boards will have to consider the impact on people as artificial intelligence can significantly impact the people working within an organisation in terms of skill sets. From an opportunity point of view, artificial intelligence can motivate and stimulate a culture of

innovation as the use of AI can bring more efficiencies and new product offerings for clients (P3s).

AI can impact organisational culture when it starts making decisions because "it has that automation layer to the intelligence, in these instances you run the risk of AI, running sort of independent law not so much augmented into day to day decision-making that people would have made as the machine takes over making the decision. By default, you don't get to apply what I would call the domain knowledge to some of the decisions that may be necessary because AI follows algorithms which follow data and data sometimes, there may be inconsistency with what is considered to be right" (P2s). P2s gave an example that AI technologies are used to evaluate credit rating in the banking sector. AI technologies use an algorithm which can decline the right people in terms of them being of higher risk if there is credit default risk because their incomes are lower, without taking into consideration that we are living in a society where people tending to earn lower incomes are people of a specific race, and thus the algorithm is inevitably discriminating. Therefore, such a decision will need a human touch, who understands transformation imperatives is an organisation that is aligned to a bigger societal goal; they will put measures in place that will foster and promote better financial inclusion, but in a responsible way so there is a role for balance without necessarily compromising the integrity of the commerciality business model (P2s).

Using AI technologies to predict customer behaviour is now starting to get into the public domain. The challenge with these technologies is that they are becoming more intrusive, i.e., monitoring customer purchases, tracking which customers are going to which stores, and sharing such data with third parties. The board must be cognizant of potentially violating the constitutional rights of our customers and check whether we are not violating privacy and data privacy. The board must always ensure that the corporation is an ethical and a law-abiding corporate citizen (p18s).

Utilising AI technology in healthcare to make a certain decision, there are legal requirements and issues around accountability, and you can't hold an AI accountable for making a mistake; the Doctor needs to be held accountable. In such instances, how do you hold the doctor accountable when AI makes a decision? There are still debates around legalities and accountability structures, which South Africa hasn't ironed out yet. Thus, at the moment, most AI technologies are functional support; they don't function as replacement technologies; they don't replace humans but support humans (P20s).

"I think, maybe it's not ethical, but from a culture point of view, the fear that AI technology instils fear of losing my job, privacy concerns" (P1s).

Analysis

AI technologies have the capacity to transform various aspects of organisational operations, decision-making processes, and interactions with stakeholders. As such, it is essential for the board to consider the potential impact of these technologies on the organisation's ethical values, norms, and practices. Thematic analysis of the first sub-question also indicated some risks or opportunities and/or positive or negative impacts of AI technologies on the organisational culture.

From a positive standpoint, if correctly deployed, AI technologies can contribute to a high-performance and learning culture. The rapid advancements and changes in AI technologies necessitate a combination of performance improvement and a willingness to learn from everyone involved, which can lead to organisational success (P6). Furthermore, AI can enhance performance and stimulate a culture of innovation as the use of AI can bring more efficiency and new product offerings for clients (P3s).

AI improves operational efficiency and helps achieve effectiveness in organisational operations (Tursunbayeva and Renkema, 2022). The study conducted by Malik et al. (2021) found that AI technologies positively impact work-related flexibility and autonomy, creativity and innovation, as well as overall enhancement in job performance. The study conducted by Noy & Zhang (2023) also found improved efficiency and decision-making and improved team morale, collaboration, and collective learning.

AI improves business intelligence and performance (Selene and Gong, 2014). Existing research has well documented the positive impact of adopting AI technologies to improve productivity at the organisational level (financial, marketing, and administrative) and streamlining organisational processes for effective decision-making and tasks (Arslan et al., 2021). In the right setting, AI technologies can enhance the performance of individuals and teams, streamline workflows, and even complete some tasks to enable employees to focus on more meaningful work and work-life balance. An optimised AI technology system can produce the same outcomes as a trained individual in certain business areas. Doing so frees

up time and human resources to develop and innovate in other business areas that require new human learning. In addition, these new, innovative learnings can be passed down and taught to machines over time.

The study by Malik et al. (2021) found that creativity and innovation positively impact adopting AI technologies and enhances employees' talent by increasing their job learnings, allowing them to indulge in creativity and innovation in management processes and functions by freeing their time from mundane tasks. Further, it was found that an increase in the inputs of employees' decision-making is due to increased access to information and real-time data provided by AI, which can facilitate the prediction of the future and faster judgment by employees (Malik et al.,2021). AI technologies have introduced a new level of efficiency, impacting the organisational culture.

However, the findings also reveal AI technologies' negative impact on organisational culture. The participants highlighted issues such as unemployment, exacerbated by AI's potential to replace jobs, the risk of abdicating ethical responsibilities, privacy, transparency, accountability, and ethical dilemmas. Indeed, if not properly managed, AI can have a negative impact on the organisation. For example, AI and automation could lead to job displacement, which could create anxiety and insecurities among workers, thus leading to a decline in morale and trust, which will impact performance. Another example could be collecting and analysing vast amounts of data by AI technologies, which can raise privacy concerns among employees. If data is mishandled or used inappropriately by AI technologies, employees may be held responsible, especially with the introduction of the POPI Act, which may lead to jail time.

Lastly, a lack of transparency and accountability in decision-making by AI technologies could lead to ethical dilemmas wherein employees may be placed in a situation where they must make ethical judgments based on AI-driven actions but lack understanding of how the AI arrive at certain decisions. An example was given by P2s that "AI technologies use an algorithm which can decline the right people in terms of them being of higher risk if there is credit default risk because their incomes are lower, without taking into consideration that we are living in a society where people tending to earn lower incomes are people of a specific race, and thus the algorithm is inevitably discriminating."

Understanding the potential impact of AI technologies on the organisation's ethical culture is crucial because it enables the board to proactively address ethical concerns and ensure that AI technologies are developed, deployed, and utilised in a manner that aligns with the organisation's ethical values and standards. It also highlights the need for the board to provide guidance and oversight to management in navigating the ethical complexities associated with AI technologies.

In organisational ethics, in particular, the rise in AI technologies has brought a spotlight on how to ensure that AI technologies companies adopt function in line with organisational core values. Establishing an organisational culture for AI technologies' research, design, development, and deployment is of central importance for effective AI ethics and governance. AI ethical principles, i.e., accountability and responsibility, fairness, transparency, explainability, and privacy, should be closely linked to the overall organisational ethics. It has been well-researched that without AI ethics, corporations can break customer trust and perpetuate bias and safety concerns – all of which generate risk to the organisations. The bottom line of these findings provides valuable developmental information for boards to deal with opportunities and risks presented by the development and adoption of AI technologies.

7.5.3. How to align AI technologies strategies with ethical culture.

The third sub-question "What is your understanding of how artificial intelligence technologies strategies are aligned to the organisation's ethical culture?" was asked as part of the second research theme sub-question "What is the board's role in ensuring that AI technologies are aligned with the organisation's ethical culture?". The question was asked to assess the board members' perception of the alignment between AI technology strategies and the organisation's ethical culture. Incorporating AI technologies into an organisation's operations and decision-making processes requires careful consideration of ethical implications.

The researcher aims to understand the board members' awareness and understanding of the importance of aligning AI technology strategies with the organisation's ethical culture. He seeks to explore whether they recognise the need for ethical considerations to be integrated into developing and implementing AI strategies. The question provides an opportunity for board members to reflect on the extent to which they consider ethical implications in the formulation of AI technology strategies. It aims to explore their

understanding of how AI strategies can reinforce or shape the organisation's ethical culture and whether they recognise the potential risks and challenges that may arise if these strategies are not aligned with ethical values.

Responses from board members herein referred to as 'Participants (P)'.

"My view is that you need digitisation in the business strategy, not a separate strategy for digitisation and a separate strategy for the organisation. So, you first say, where is this organisation going? What is our strategy? And then how is digitization going to help us? And part of that is how artificial intelligence that is going to help us in the achievement of this particular strategy?" (P7).

P8 argued that AI is an enabling tool to achieve the company's strategic objectives; these technologies have to be aligned with the organisational strategy. They cannot be off-tangent to the organisation's goals (P8). P24 agreed with the sentiments above that AI is an enabling tool that would give the company leverage to achieve its strategic objectives. It is important to ensure alignment with the strategy, and AI is there to support the employee in achieving the objectives (p17).

"You can't have one set of values for the organisation and a different set of values for, you know, artificial intelligence in isolation, so consistency in our values is critical. Those values should speak to every aspect of the business" (P16).

"My starting point is to the extent that people deploy technology, they do it in pursuit of a strategic aim" (P21). There is a need for a strategic alignment as the deployment of AI technologies should be in pursuit of strategic objectives (P21). The adoption of AI technologies should consider outcomes which are the strategic objectives of the corporation; thus, IT strategy and the business strategies should be aligned (P23).

Responses from specialists herein referred to as 'Participants (Ps)'.

P2s argued that the best way to align is actually to not have too many strategies; organisations need only one business strategy. There should not be a separate AI or data strategy because organisations spend a lot of time trying to work on that alignment. The organisation should be deemed as working towards one organisational strategy. AI

technology should not have a strategy that exists in isolation, it should actually be a function of the organisational strategy, and AI technology should only exist to support what is the organisational strategy that would have been approved at the board level, with the goals and objectives and targets that the executive leadership would have been in agreement about with the board (p2s).

It should be absolutely aligned, but the focus should be on business strategy. AI is not the solution to the business strategy, but as part of that business strategy and the organisational planning, we have to consider how do we interact using technology (P1s).

"Whenever you have a strategy that's not the organisational strategy, the strategy should start with the word 'therefore'. So, this is the organisational strategy; therefore, the AI strategy is this. So, the AI strategy always has to be subservient to the organisational strategy. So, there's an overall value statement, code of ethics, organisational strategy, and it's on that basis that these sub-components should be built" (P9s).

"So basically, what happens is that the board will set its culture to say that look, we are becoming a more digitally driven, increasingly digitised, and digitally-driven organisation. So, digitization and being digitally driven are two different things. But what we're going to be doing as a board is we're going to be setting a culture such that the organisation needs to embrace our use of technology. But we also need to do this in a responsible way such that we do not harm our stakeholders, we have zero harm towards our stakeholders and employees, and we have the greatest value to them as we actually do our business going into the future. So that's [how] our board would approach it. So, when you set the culture, you're actually setting a tone for how you expect the board or the organisation to be able to say, so from the board down to the executive, from your executive down to your senior management, and your senior management into the organisation and … the organisation out for your customers and other stakeholders, suppliers, etc." (P18s).

P20s indicated that the organisation could not introduce AI technologies or a data-driven strategy if the culture is completely the antithesis of that, there needs to be a mindset shift in staff for a cultural change (p20s).

P3s argued that it is about ensuring that the AI tools used are aligned with your business strategy. A clear business strategy should be in place, so everyone understands and knows business objectives.

Analysis

The finding reveals that most participants did not fully comprehend the question; thus, they did not respond to it. However, few participants understood the question of aligning AI technologies to the organisation's ethical culture. As already stated, the question was asked to assess the board members' perception of the alignment between AI technology strategies and the organisation's ethical culture. This seeks to explore whether they recognise the need for ethical considerations to be integrated into developing and implementing AI strategies. The question provided an opportunity for board members to reflect on the extent to which they consider ethical implications in formulating AI technology strategies. It aimed to explore their understanding of how AI strategies can reinforce or shape the organisation's ethical culture and whether they recognise the potential risks and challenges that may arise if these strategies are not aligned with ethical values.

P16 argued that an organisation "can't have one set of values for the organisation and a different set of values for artificial intelligence technologies in isolation, so consistency in our values is critical. Those values should speak to every aspect of the business". AI technology should not have a strategy that exists in isolation; it should be a function of the organisational strategy, and AI technology should only exist to support the organisational strategy that would have been approved at the board level, with the goals and objectives and targets that the executive leadership would have been in agreement about with the board (P2s).

Incorporating AI technologies into an organisation's operations and decision-making processes requires careful consideration of ethical considerations. The strategies employed to develop, deploy, and utilise AI technologies should be aligned with the organisation's ethical values, principles, and goals. Thomaz et al. (2021) argued that AI ethics should be a logical extension of a corporate strategy and objectives. An AI ethics statement should serve as a blanket declaration that the development, deployment, and adoption of AI technologies are consistent and ethically caried out. For the purpose of enhancing transparency, the statement should include the hierarchy of ethical values of the corporation.

The alignment of organisational cultural norms with AI technology development and deployment is key to the uptake and efficacy of explicit rules and regulations. Individuals are more likely to adhere to the organisational culture and norms they believe in (Seger, 2022). Explicit rules and requirements often generate extra work, and time-consuming requirements like checklists, self-evaluations, and impact statements may lead to resentment by AI developers if seen as unnecessarily restricting (Baum, 2017). Such resentment may, in turn, lead to rejection of the very principles being promoted (Pettit, 2002).

P8 argued that AI is an enabling tool to achieve the company's strategic objectives; these technologies have to be aligned with the organisational strategy. P24 agreed with the sentiments above that AI is an enabling tool that would give the company leverage to achieve its strategic objectives. It is important to ensure alignment with the strategy, and AI is there to support the employee in achieving the objectives (P17). AI technology alignment aims to ensure that powerful AI is properly aligned with human values (Russell, 2019). It 'is necessary to align the processes of AI technologies with corporate strategy, ethical values and overall corporate objectives in order to explicitly translate relevant ideas into specific norms that influence concrete design and governance to shape the development and use of AI technologies. Organisations should create AI ethics frameworks that are actionable, with clear accountability and concrete methods of monitoring. To align AI technologies with corporate values and business ethics implies that explicability, transparency, accountability, fairness, responsible use of data, guarding against criminal misuse, aligning AI with values and being mindful of the societal implications of AI are all essential pillars of ensuring responsible use of AI technologies.

The organisational AI technologies strategy provides a general direction and manages expectations regarding the overall set of AI technologies that the organisation intends to use and what they are meant to achieve. The strategically aligned use of AI requires organisational resources, capabilities, and processes that can be ensured (Shneiderman, 2020). Ethical culture (values) requires boards to state the value base and AI ethics the organisation adheres to, and adherence to these ethical values and principles across the organisation's AI systems must be ensured. Including strategic and value alignments as separate themes highlights alignments at different levels (Whittlestone et al., 2019).

7.6. Research sub-theme 3:

The third sub-theme, "The board's role in ensuring that the research or design or development or deployment or use of AI technologies does not negatively impact on society", focused on the board's responsibility in mitigating potential societal harm arising from the various stages of AI technologies within the organisation.

Only one broader question was posed to participants: "What is your understanding of the responsibility of the board in mitigating societal harm from AI technologies in play in your organisation?" under this theme. The context of this question lies in addressing the ethical and social implications associated with AI technologies. It aims to assess the perceptions of the board's awareness of the potential risks, challenges, and unintended consequences that may arise from the organisation's involvement with AI technologies. It also emphasises the board's role in ensuring that the organisation operates in a way that aligns with ethical standards and avoids or minimises harm to society. It helps establish a context for evaluating the board's commitment to promoting responsible AI practices and their efforts in considering the potential social impact in the decision-making processes related to the research, design, development, deployment, and use of AI technologies.

7.6.1. The responsibility of the board in mitigating societal harm from AI technologies.

The question "What is your understanding of the responsibility of the board in mitigating societal harm from AI technologies in play in your organisation?" was asked as part of the broader research objective to understand the board's role in ensuring that the research, design, development, deployment, or use of AI technologies (whichever is applicable in the organisation's context) do not have negative impacts on society. By asking this question, the research aims to understand the board members' comprehension of their responsibility to safeguard against societal harm arising from the organisation's use of AI technologies.

Understanding the board's responsibility in mitigating societal harm is vital because it enables board members to provide effective oversight and guidance to management in developing policies and strategies prioritising the ethical and responsible use of AI technologies. It also highlights the need for the board to establish mechanisms for

monitoring and evaluating the impact of AI technologies on society and taking corrective measures when necessary.

Responses from board members herein referred to as 'Participants (P)'.

The board needs to broadly understand the ethical risks of artificial intelligence in terms of the stakeholders that the use of artificial intelligence technology can impact. The board also needs to understand the opportunities that emanate from using artificial intelligence. Where there are risks or opportunities, the board delegates to management to draw up risk mitigation plans or a plan to maximise opportunities that will flow out of the implementation of artificial intelligence. The board is to ensure consultation at the level of society where they are most likely to be harmed by these AI technologies (P7). P7 indicated that they found that it is not enough to just deal with harm from one organisation's point of view because there are multiple stakeholders who are involved in the creation of these AI technologies. Thus, the board needs to ensure partnerships with suppliers of artificial intelligence if outsourcing.

Furthermore, the board must ensure that consumers are educated in using these technologies. Lastly, the need to adopt a cradle-to-grave approach; the board should ensure the safe disposal of the technologies after use as some harm to the society or environment may emanate from the waste. Basically, to ensure that it is safely disposed of, it does not cause harm to the environment, does not endanger children, and ends up affecting people's livelihoods (P7).

"Well, if we're going to be a responsible corporate citizen, then we have a responsibility to ensure that we are mitigating against the harm to society. Part of that responsibility would also be making sure that when the board approves the use of AI for a particular business, the board is made aware of any risks. The board then ensures that potential risks are monitored and reviewed regularly" (P17). P19 also indicated that any potential strategic risk to the company is the board's responsibility. P21 argues that the board is clearly responsible for assessing the risks associated with AI technologies. The board must carefully consider how deploying a particular AI technology will impact their business. The broader issue of societal harm raises a larger question about the board's responsibility in mitigating such harm. This goes beyond just AI and encompasses the board's obligation to consider the organisation's impact on communities and societies as outlined in the King IV report. If the

deployment of AI poses a risk of harm to society, the board should take it into account and assess the implications for the company. They should implement measures to mitigate these risks, either through the development and deployment of AI itself or through subsequent actions to address any negative consequences. Ultimately, it is part of the board's responsibility to evaluate the overall impact that the company has on society rather than solely focusing on AI (P21).

The board is responsible for holding executives accountable for the strategy. The strategy includes the responsible use of artificial intelligence such as customer complaints, societal compliance, and environmental harm (P5). The board needs to ensure policy development and seek the best practice from other companies or industries. "I don't imagine technicians sitting somewhere in the lab and developing these AI dolls may necessarily have gotten board approval or gone to the board, but the policies that govern the organisation should direct them on what to do" (P6). The board must do everything to prevent societal harm by ensuring the appropriate organisational culture percolates down to the lowest level of the organisation (P12).

The board should establish a social and ethics committee to ensure that the deployment of AI technologies does not impact people and the environment. For example, if the technologies include using chemicals, the board, through the social and ethics committee, needs to ensure that such chemicals are safe for the environment and reduce pollution levels (P8). The social and ethics committee ensures that the decisions that are made also talk to the communities that the organisation is servicing and how those potential risks are beyond just the exposure but also the reputational risks. The social and ethics committee also need to ensure stakeholder engagement to the extent of open communication around interventions being affected to ensure remedial action is executed (P24). The board has the same responsibility related to sustainability, ethics, and artificial intelligence, and it cannot walk away from accountability for the environment and society (P11).

P13 argued any competent board member would recognise the need to consider mitigating societal harm from AI technologies. However, there is no specific guidance available. For instance, in King IV, there is no mention of mitigating societal harm from AI technologies. The focus is primarily on data and data governance, with no reference to artificial intelligence. There is certainly no mention of artificial intelligence in relation to ethics or the role of algorithms in designing and developing products. Nonetheless, it is evident that just

as climate change responsibilities have become a concern for boards, the same will likely happen with regard to artificial intelligence, ethics, and bias. "I think that investors will start asking the question, just as they ask questions about the company's carbon footprint, they're going to ask the question about the artificial intelligence footprint of the company. And they're going to ask for transparency of the artificial intelligence utilisation and the explanations and explainability of bias that could be introduced. So that there's social equity as part of ESG environmental, social and governance" (P13).

The board needs to be aware of all the impacts of any decision, be it social impact, human capital impact and or environmental impact. The board's responsibility is to make sure that when they make a decision, they're making it with the full understanding of all the impacts and that they have considered all of them (P14). The board must fully understand the extent to which AI can be used to empower communities and society and actively address the potential risks for society from AI technologies (P16). The board is fully accountable and responsible for the environment and the society. Thus, the board's decision should not only be about the bottom line but also think about the impact on society in general (P22).

"Let's take something like Waze (GPS), using all the data of all the people travelling on Ben Schoeman Avenue, working out where the traffic jams, and it certainly benefits society. I think it becomes dangerous when you start manipulating people's emotions, like the Cambridge Analytical scandal on elections. So, I can see problems but don't know the solutions" (P4).

Responses from specialists herein referred to as 'Participants (Ps)'.

The boards are required to ensure that their organisation does not cause harm to society. The board should fully understand the stakeholders, especially those who will be materially impacted by the development or deployment of AI technologies, both at value creation and harm. The board should also set a strategy for AI development or deployment, which is guided by the organisation's governance policies. To demonstrate accountability and transparency, the board should also implement effective corporate reporting using an integrated report to disclose information that's not in the financial statements, including the impact of AI technologies in society (P9s).

From a legal perspective, the board is jointly and severally liable for whatever the organisation does when it conducts its business; thus, the board needs to understand their fiduciary responsibilities. They should exercise the utmost care and accept full responsibility regarding how the organisation will deal with technology and artificial intelligence and how it is deployed in the organisation (P18s). The board needs to understand the risks that are associated with AI technologies. They (boards) need to ensure that all the audits and governance processes around that are aligned to identify any risks upfront and manage them accordingly (P15s).

The board should adopt the triple bottom model, which is able to balance environmental, social and governance priorities. Such a module is more sustainable because the business exists within the society, and they cannot ignore the needs of society. Therefore, the board cannot be implementing or leading AI technologies that are potentially working against the social, environmental or governance imperatives because then that will actually work to extinguish society and, by default, the business's sustainability in the long term (P2s; P3s).

P20s argue that boards are not necessarily obligated to care about societal harm as long as their actions are within the boundaries of the law. This applies to any kind of tool or technology similar to how pollution exists, because individuals are not compelled to care unless it is illegal or enforced by external factors like government regulations. The same principle applies to technologies and their impact on society. Unless there is a tangible incentive, many companies may not prioritise societal consequences unless they are engaging in illegal activities, which they typically try to avoid. This is why there is a growing emphasis on Explainable AI and responsible data science, aiming to raise awareness for responsible AI (P20s).

Analysis

The board should adopt the triple bottom model (Slaper & Hall, 2011; Gupta, Mishra & Tandon, 2020; Pedroso et al., 2021) which is balancing financial, environmental and societal priorities. Such a model is more sustainable because the business exists within the society, and the needs of society cannot be ignored (P2s; P3s). The board is fully accountable and responsible for the environment and the society. Thus, the board's decision should not only be about the profit but also about the impact on society in general (P22). The triple bottom line is a sustainably focused model which requires the corporation not only to focus on the

profits (financial bottom line) but also include broader economic, social, and environmental concerns to help measure how a corporation affects its stakeholders, such as the surrounding community and the environment as a whole. In recent years, AI and financial performance-related research have been growing, and so has research that addresses the impact of AI and sustainability (Keil et al., 2020).

Under sub-theme 1, the participants have already highlighted issues such as unemployment, exacerbated by AI's potential to replace jobs, the risk of abdicating ethical responsibilities, privacy, transparency, accountability and ethical dilemmas. Moreover, AI can directly impact society by compromising personal information, biased decisions (based on flawed historical data), and lack of transparency (how decisions are made). Additionally, AI technologies can have an environmental impact, such as training an algorithm to identify reliable patterns, which can require heavy energy to analyse millions of datasets.

P17 has argued that "if we're going to be a responsible corporate citizen, then we have a responsibility to ensure that we are mitigating against the harm to society. Part of that responsibility would also be making sure that when the board approves the use of AI for a particular business, the board is made aware of any risks. The board then ensures that potential risks are monitored and reviewed regularly". The board is responsible for ensuring that the corporation is seen as a corporate citizen, meaning that the corporation is operating responsibly and ethically towards society. To be considered responsible corporate citizens, the boards need to consider the responsible corporate use of AI technologies and the potential impact on key stakeholders such as employees, customers, society, and the environment. The board should ensure that AI technologies comply with the constitution, laws and standards, company policies and procedures, strategy, and codes of conduct and ethics. Through engagement with stakeholders and communities, the board should also ensure that AI technology advancements improve the material well-being of the societies in which the organisation operates (Mulamula, 2022).

The board must carefully consider how deploying a particular AI technology will impact their business. The broader issue of societal harm raises a larger question about the board's responsibility in mitigating such harm. This goes beyond just AI, and it encompasses the board's obligation to consider the organisation's impact on communities and societies as outlined in the King IV report. If the deployment of AI poses a risk of harm to society, the board should take it into account and assess the implications and reputational damage to

the corporation. They (board) should implement measures to mitigate these risks, either through the development and deployment of AI itself or through subsequent actions to address any negative consequences (P21). This will ensure that the potential impacts of AI technology on society are carefully considered and addressed. This may involve conducting risk assessments and developing strategies to mitigate potential negative impacts. As already mentioned in the previous section, the board needs to understand the risks that are associated with AI technologies and need to ensure that all the audits and governance processes around that are aligned to identify any risks upfront and manage them accordingly.

The board should ensure consultation at the level of society where they are most likely to be harmed by these AI technologies (P7). P7 indicated that they found that it is not enough to just deal with harm from one organisation's point of view because there are multiple stakeholders who are involved in the creation of these AI technologies. Thus, the board needs to ensure partnerships with suppliers of artificial intelligence if outsourcing. The board should consider engaging with stakeholders, including employees and customers, to better understand the potential risk and impact of AI technologies and identify ways to minimise any negative impact. This may involve working with multi-disciplinary experts and engaging in dialogue with diverse groups to ensure that the voices of those most affected by AI technology are heard and considered.

The board should provide oversight and guidance to ensure that AI technology initiatives are aligned with its values, mission, and long-term goals. This may involve setting goals and metrics to track progress and holding executive management accountable for ensuring that AI technology is used in a way that is consistent with the values and mission. This could involve establishing policies and guidelines for the development and use of AI and monitoring and reviewing the implementation of these policies to ensure that they are being followed.

The board should establish a social and ethics committee to ensure that deploying AI technologies does not impact people and the environment (P8). The electronic waste (e-waste) produced by AI technologies poses a serious environmental challenge. E-waste contains hazardous chemicals, including lead, mercury, and cadmium, that can contaminate soil and water supplies and endanger human health and the environment. The World Economic Forum (WEF) projects that by 2025, the total amount of e-waste generated

will have surpassed 120 million metric tonnes (Kanungo, 2023). Thus, to avoid environmental harm and minimize the release of dangerous compounds, the board must adopt a cradle-to-grave approach, which includes proper e-waste management and recycling to ensure safe disposal to avoid a risk of harm to society or the environment.

In addition, it should also ensure data wiping of personal information in line with Protection of Personal Information Act 4 of 2008 ("POPIA"), which places an obligation on the corporation. POPIA section 14(5) states that personal information must be destroyed or deleted in a manner that prevents its reconstruction in an intelligible form.

The social and ethics committee ensures that the decisions that are made also talk to the communities that the organisation is servicing and how those potential risks are beyond just the exposure but also the reputational risks. The social and ethics committee also needs to ensure stakeholder engagement to the extent of open communication around interventions being affected to ensure remedial action is executed (P24). The board has the same responsibility related to sustainability, ethics, and artificial intelligence, and it cannot walk away from accountability for the environment and society (P11).

The board should play a critical role in promoting transparency and accountability around the design, deployment and development of AI technologies and the establishment of mechanisms for addressing any concerns or risks related to AI. The board's responsibility in mitigating societal harm from AI technologies is to ensure that the company is acting responsibly and ethically. Understanding the board's responsibility in mitigating societal harm is vital because it enables board members to provide effective oversight and guidance to management in developing policies and strategies prioritising ethical and responsible use of AI technologies. It also highlights the need for the board to establish mechanisms for monitoring and evaluating the impact of AI technologies on society and taking corrective measures when necessary. The board's fiduciary oversight obligations extend to the use of AI technologies; thus, it should understand how AI impacts the corporation and its obligations, opportunities, and risks and apply the same general oversight approach as it applies to other management, compliance, risk, and disclosure topics.

7.7. Research sub-theme 4:

The fourth and last research sub-theme was: "What is the board's role in ensuring that adequate AI governance structures are in place?" The research sub-theme included three questions posed to the participants:

(1) *Does your board have a governance subcommittee for technology or AI? If yes, what is your understanding of the role of the AI or technology governance subcommittee? If no, do you think such a subcommittee is needed? Why or why not?*

(2) *Does your board have responsible use of AI governance framework? If yes, what is your understanding of the role of the framework? If no, do you think such a framework is needed? Why or why not?*

(3) *How does the board monitor the performance and value delivery of artificial intelligence technologies?*

As AI technologies become increasingly integrated into various aspects of business operations, it is crucial for organisations to have robust governance structures in place to ensure their responsible and ethical development, deployment, and use. AI governance encompasses policies, procedures such as impact assessments, and frameworks that guide decision-making, risk management, accountability, and transparency related to AI technologies.

The question aims to explore the extent to which extent board members recognise their role in overseeing and establishing these governance structures. By asking this question, the research seeks to understand if the board members are aware of the importance of having adequate AI governance structures in place, as well as their understanding of the specific steps and responsibilities involved in informing, establishing, and monitoring these structures.

The research sub-theme and its questions provide a context for understanding the board's role in establishing and overseeing AI governance structures. The question addresses the importance of governance in the context of AI technologies and aims to understand the board's role in ensuring that organisations have adequate structures in place to govern AI effectively.

7.7.1. Board governance subcommittee for technology or AI

The sub-question "*Does your board have a governance subcommittee for technology or AI?*" was asked as part of the fourth research theme sub-question "What is the board's role in ensuring that adequate AI governance structures are in place?". The question about whether the board has a governance subcommittee for technology or AI, and further exploring the understanding or need for such a subcommittee, is asked to assess the organisation's current approach to governing technology and AI and to understand the board's perception of the importance of governance oversight.

By asking this question, the research sought to gather information on whether the organisation already has a dedicated governance subcommittee for technology or AI. If such a subcommittee exists, the subsequent question aims to explore the understanding of the board members regarding the role and responsibilities of this subcommittee. Understanding their perception of the subcommittee's role helps in assessing how well the organisation's governance structure aligns with best practices and industry standards.

If the organisation does not currently have a dedicated governance subcommittee for technology or AI, the research aims to explore the board members' views on whether such a subcommittee is needed. This question helps in assessing the board's awareness of the potential risks, complexities, and ethical considerations associated with technology and AI. It also provides an opportunity for board members to express their opinion on whether establishing a dedicated subcommittee would be beneficial for the organisation's governance framework.

Responses from board members herein referred to as 'Participants (P)'.

"I serve in one board that does have IT as part of another committee. The other board also has IT as part of one of the existing committees, so they don't have a separate IT committee primarily because the Tech specialist said they should not create a separate committee yet. We felt that it would be boys and their toys which would explode the IT budget…. So, in both organisations, we've kept this part in other committees such as audit risk and IT committee and other one it's got like a whole host of things but includes IT" (P4).

The board cannot always have specialists in everything that happens in a business, and the board delegates certain authorities to board sub-committees. P5 argues that in their case, AI technologies are a topic of three board sub-committees; operations & IT risk committee, social & ethics committee and risk capital, compliance and conduct sub-committee, which is a cross-membership. The social and ethics committee would look after the ethics and conduct of AI, technology, and machine learning. In contrast, risk conduct and compliance committees have a specific role in measuring risk outcomes. These committees are composed of risk, compliance, and ethics specialists. The committees report back to the board on a quarterly basis. "We have got cross-membership between those three committees of independent directors sitting on the board, and I think your most important thing for governance is to have that cross-membership" (P5).

According to P7, AI is discussed in the board's technology and information governance sub-committee. The committee deals with the strategy as far as the strategic inputs are concerned and deals with the risks as far as the risks of technology are concerned, that will be risks of acquisition, risks of use, as well as disposals, but it also deals with the monitoring of the reports that come out of the technology governance. IT is also discussed in other board committees to the extent that IT impacts the social and ethics committee. But the audit and risk committee also discusses the ethical practice guiding artificial intelligence adoption; the ethical risks will be identified by the risk committee, and the assurance in as far as those ethical risks go, is provided by the audit part of the audit committee (P7).

"I chair the audit and risk committee. King IV requires that the Board should have an information technology agenda. So, most boards have delegated to the audit committee. So, when I chaired the audit committee, we had an IT governance agenda. So, within that IT agenda, we look at IT governance, we have policies in place, we look at cybersecurity, we look at artificial intelligence, and we will look at emerging technologies. And [the] social and ethics committee will look to add the safety of our children and the safety of the environment" (P8).

P6 argued that there is a need for a subcommittee on digital technology. P6 suggested that a social and ethics committee could be the committee to address the responsible use of AI in the organisation. Further, they highlight that many organisations may not be aware of the extent to which they are utilising AI because these technologies are often obtained from the open market or through open-source platforms. They mention that the JSE (Johannesburg

Stock Exchange) also requires this understanding of data utilisation as a listing requirement (P6).

P10 indicated that the IT subcommittee's role is to look at technology and artificial intelligence comprehensively and ensure that information's integrity and protection. Further, the sub-finance committee would look at the budgets and the audit committee would be dealing with risks (P10).

P13 indicated that they do not have a specific subcommittee for AI or technology in general. The board only has an investment committee which focuses on investments in technology, mainly investment decisions, in terms of cost of implementation. The investment committee does not focus on the impact, risk, governance or ethics of technology and data utilisation. "There's no subcommittee like that, and the board has a standing item where the CEO comes in and explains how far down the line they are with the implementation of the digital transformation. And there will be some general questions asked, but there isn't any specific subcommittee that deals with artificial intelligence beyond the financial committee, the investment committee, and sometimes the audit and risk committee" (P13).

"So, one board has a technology committee, and the other board does not. In the mining industry, each mine has its own technology. It's not that there's not that much discussion, to be honest, at a board level. And then the retail companies, the board I sit on the holding company, where each of the retail operations underneath have their own boards that deal with most of the operational stuff. I know that … there are risk committees, and I know that there is an internal audit function; we don't get reports on IT functions specifically. So, I suspect it's not a major focus of the companies (P14).

"Not in the board" (P14). Such oversight is lower down the steering committee at the management level. As part of the mandate, the audit and risk committee also looks at technology and information, risks and opportunities. So, it normally would come to be channelled through that committee into the boardroom (P14).

"So none of the boards I'm sitting on at the moment, AI as not necessarily an agenda item at the moment but technology is" (P17). The subcommittee provides oversight on the technology deployed or explored. The committee is aligned with King IV on IT governance and enterprise risk oversight (P17). P19 added that IT has to be considered within the

enterprise-wide risks due to technology risks where companies are using AI. For a company or organisation which uses significant technology, it makes sense to have an IT committee and at least one independent board member on that committee (P19).

"Yeah, so we don't have a dedicated technology committee; we have a risk committee, which is where most technology discussions will, in fact, take place; broader strategic conversations take place with the full board. So it's not a subcommittee; it's a full board-related technology risk issue, which typically will be that where the discussions take place around, you know, deployment, technology, failures, costs, etc., will take place in the risk committee, though, they will also be picked up, for example, in the audit committee, and we also have what used to be called a social and ethics committee is now just called a sustainability committee, where to the extent that there are AI or technology specific considerations that fall under that rubric, that would be [where] we would pick that up. With our organisation and the way we are structured, we wouldn't have a technology committee specifically. We think about it in the context of risk, strategy and sustainability" (P21). The purpose of subcommittees will be to apprise the board on particular risk issues that were discussed, what the resolutions were that were reached, and any sort of specific action items coming out of them. So, the full committee, the full board, will always have oversight of the full set of issues through a discussion but enables a smaller group of people to deal with the specific issues concerned (P21).

AI technologies are delegated to the audit committee. The audit committee is responsible for any new technology developments or deployment. The committee, in that regard, reviews internal auditors' reports, reviews the progress from the system developers, and reviews the feedback from the business units regarding the functionality of the systems. Basically, this is focused on getting feedback from stakeholders and, most important, feedback from the internal audit around the controls that are embedded in the systems or the controls that are outside of the systems. The committee is also responsible for monitoring the cost versus the budget of the system and return on investment (P22).

The audit committee is mandated to look at audit issues, including statutory reporting methods, as well as ICT issues. A separate risk compliance committee which is not an audit committee provides risk oversight on ICT. "In my other world, they sit as part of the risk compliance and ICT Committee, which has a broader mandate. EXCO will specifically look

at AI and report back to the board through the committee that's responsible for ICT governance" (P24).

P23 argues that whether it is through audit and risk or through an IT committee that discussion of governance issues happens, such committees have the mandate to consider the implications of information, technology, and new technologies such as AI. It has the mandate to consider all of the factors that surround those areas and highlight concerns and issues to the board and deal with the implications.

Responses from specialists herein referred to as 'Participants (Ps)'.

"In the case of our board, yes, but there is no committee specifically only for technology at the board level; technology is reported on the risk and audit committee. But at the executive level, there is an IT steering committee" (P3s). The risk and audit committee is responsible for ensuring that there is an appropriate governance structure for risk management. Risks are appropriately leveraged and mitigated, and the organisation is operating in its risk appetite; that risk is properly embedded and entrenched in an organisation. Further, to ensure that the same principles are applied when technology is introduced, such as compliance with regulations, risk matrices and risk maturity around technology and data protection, and penetration tests, the appropriate security is in place. "The audit committee would also additionally apply the lens of control from an internal control perspective and ensures that the appropriate controls are in place in executing auditor findings and the financial components" (P3s).

"I think the best thing is that there is an advisory board and working with the main board. The advisory board should consist of multidisciplinary people. So, you need some really good technologists who can explain the impact of technology. But you also need your human resources people to understand how it will impact culture and, of course, your legal people when it comes to privacy concerns, market perception and the like. But that is because I think that everyone on the board will not go into all this detail. But when there's a new initiative that's being proposed, or a new question or problem, that they're looking for a technology answer for this advisory board, and that's why it should be multidisciplinary. The problem is that advice is only coming from technology voices, just this tech is great, it's going to save us a million Rand and so forth. But they should agree before they present to the board, taking into account legal risks, impact on people, and what the technology can do;

you can even have marketing people involved regarding perception in the market, and so forth; you could have your environmental people involved in the case of mining. So, I would say the boards need a multidisciplinary advisory board, through which all these recommendations come in that they are taken together. So, it's not just one initiative in legal …, and so forth. And then they need to make decisions that are relevant to your directors to make, and that's around risk, perception risk, future risk, competitor risk, impact on people risk, but then at least I get that multi-lens view from the advisory board and not just the technology" (P1s).

"We do have a subcommittee from an IT perspective, but not specifically an AI one, my views as to the effectiveness thereof with the context of being able to regulate or manage developments within the AI space, I will say that it's far from being effective or adequate. My reason for that is that AI, for me, speaks to business model evolution altogether, and that cannot be the job of IT in the first place. It is the very same way we think about digital. Digital is not technology; technology is just an enabler to digital, actually digital, of which AI is part, is a key player. Digital is all about migration, traditional and analogue business models into platform business models that inherit high levels of customer network effects because of the interconnected nature of platforms, devices and so forth" (P2s). The role of the subcommittee is to regulate and mandate the acquisition and rationalisation of technology tools across the organisation and ensure that they are in line with the IT strategy. Secondly, it is to ensure that all the elements that support effective IT execution in support of business are in place and functioning, such as security or disaster recovery. Lastly, its role is to ensure that this integration with the other committees, such as finance and investment committees.

"The board uses subcommittees to provide additional technical expertise, to help with the balance of power, and to bring in additional independence or diversity of thought. So if the board finds that the organisation is looking to develop AI, implement AI, or use AI, and doesn't have enough sufficient skills, practical working knowledge of AI, it should look to bring in additional skills, and not to use a board meeting, to have oversight of that, when most of your board is it doesn't understand the topic, it would make sense to create a subcommittee … but it provides recommendations to the board, just like the audit committee" (P9s). AI is part of technology; technology also has risk and also has financial implications. The risk committee certainly has expertise in understanding risk assessment, risk treatments, and internal control. So, technology risk needs to be discussed in the technology and AI committee, and the risk committee needs to understand the treatments

to assess risk, and that risk has been assessed. The risk committee responsible for risk discipline, makes sure that the enterprise risk management framework is working effectively and risk is discussed in various committees of the organisation (P9s).

"To be honest with you, in most instances, those boards are useless. It would be very dependent on the quality of the individuals. You need people that are really entrenched in the field to sit on those boards, and are aware of upcoming technologies, the impact they can have on the company, the impact it can have on their clientele, etc. So, really [is] the quality of a board like that is completely dependent on individuals. But ideally, having an AI subcommittee like that should be able to provide the board with guidance on utilising tools appropriately, ethically, responsibly, legally, etc., etc. The different kinds of pitfalls that may exist, the different kinds of considerations that have to take into place. So ideally, a board like that can be extremely useful if it's got the right kind of individuals that are really entrenched in the right fields sitting on it" (P20s).

Analysis

The findings reveal that most boards have subcommittees at the board level that deal with broader technology issues, including AI. The importance of board committees has grown over time due to increased legal requirements and the growing complexity of the corporate environment. In organisational theory, Lawrence and Lorsch (1967) argued that the board establishes specialised structures to handle complex challenges.

"The board uses subcommittees to provide additional technical expertise, to help with the balance of power, and to bring in additional independence or diversity of thought" (P9s). The subcommittees apprises the board risk issues, key resolutions reached, and any specific action agenda item worth elevating for the board attention (P21). Specialised committees, such as the audit committee, investment committee, technology or ICT committee, social and ethics committee, and compensation committee, enable boards to handle this complexity through subgroup-focused responsibility and expertise while limiting the demands placed on individual directors. Therefore, the specialised expertise of board committees is critical to the board's ability to reach effective decisions and fulfil its fiduciary duties.

In other words, the board ensures the establishment of governance structures and sub-committees at the board level to assist in discharging its role and duties. Although the board may delegate authority, it should not result in the abdication of powers by the board. Board committees are an aid to assist the board and its members in discharging their duties and responsibilities, and boards cannot shield behind these committees.

In the case of AI, board committees are responsible for overseeing the development and deployment – and overall adoption – of new technologies are done in a responsible and ethical manner, compliant with the laws, aligned with strategic objectives, and above all, accountable. Thus, the board committees dealing with AI should have in-depth knowledge and experience of AI technologies or at least have an independent expert advising at the committee and board level.

The study observed that some participants were of the view that technology and emerging technologies such as AI should have a standalone board committee. In contrast, others suggest a multi-disciplinary approach as AI technology now affects every facet of the organisation, i.e., finance, ethics, risks, operations, etc. A single board committee could be indicative of how a board has identified the need for and prioritised technology within the corporation. From other respondents' perspective, not all corporations have the capacity or need of a standalone committee.

Some participants indicated that technology and emerging technological issues such as AI are at the purview of the audit committee. "In the case of our board, yes, but there is no committee specifically only for technology at the board level; technology is reported on the risk and audit committee" (P3s). A similar sentiment has also been raised by P22: "AI technologies are delegated to the audit committee". The IODSA research report on the application of technology governance by boards in 2021 found that most corporations (45%)
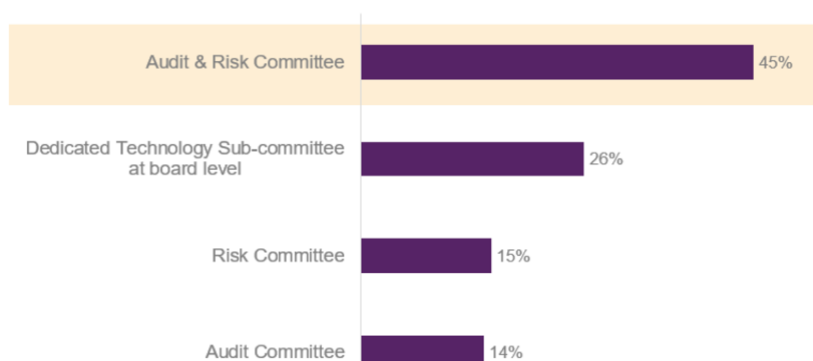

Figure 13: IoDSA Research Report on the Application of Technology Governance by boards 2021

provide oversight of technology through audit committees, and only 26 per cent have a dedicated technology subcommittee at the board level.

A similar study conducted by Cooper, Lamm, and Morrison (2023) found that AI technology oversight is mostly done at the audit committee level and only 5% by a dedicated technology subcommittee. The studies did not provide an analysis or background of why this has been a trend. The research conducted for this study has observed that the current composition of audit committees in South African firms is mainly people with financial backgrounds, i.e., Chartered Accountants, not technologists. It should be noted that the Companies Act requirement for Audit Committee composition is a minimum of three directors, with independent directors forming a majority; there are no minimum qualification requirements for members of an audit committee. However, the reality is that most companies (ABSA Bank[41], BHP Group[42], Richemont[43], Vodacom[44], Shoprite[45], etc.) prefer members with a financial background to serve on the audit committee. Also, the majority of boards still have traditional knowledge backgrounds in finance and law. The study by Gordon (2021) found that there is a lack of skill and knowledge on the depth of AI technology and digital transformation leadership and solid execution experience of AI technologies at the board level.
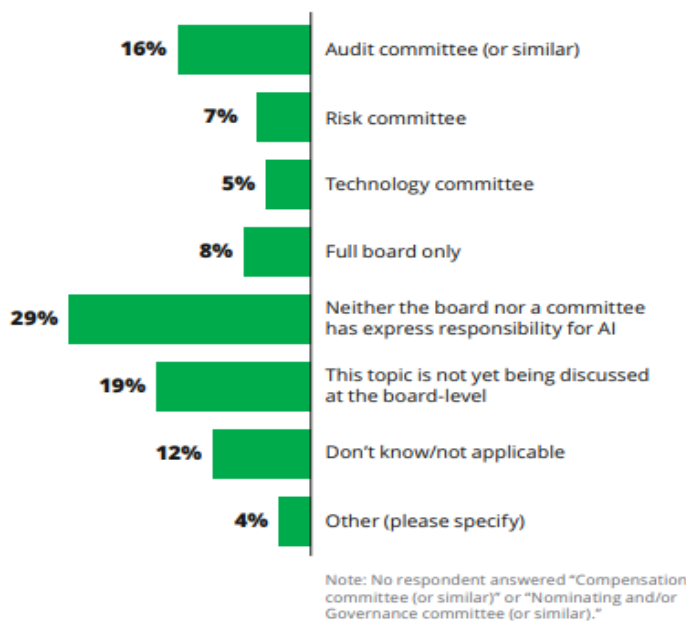

Figure 14: Board Practices: Artificial intelligence (Cooper, Lamm, and Morrison, 2023)

[41] https://www.absa.co.za/about-us/absa-bank/absa-bank-board/
[42] https://www.bhp.com/about/board-and-management/terry-bowen
[43] https://www.richemont.com/en/home/about-us/corporate-governance/josua-malherbe/
[44] https://www.vodacom.com/board-of-directors.php#0
[45] https://www.shopriteholdings.co.za/group/directors.html

Thus, the study suggests that there is a need for the creation of a dedicated AI and technology board committee supplemented by other board committees (multi-disciplinary approach) such as audit, risk and social and ethics sub-committees. Boards should contemplate the formation of a specialised board committee focused on AI and technology governance, considering factors such as the corporation's scale, the intricacy of its technological landscape, and the extent to which its business and operations rely on technology. Creating a dedicated committee will add technical expertise at the board level (Brennan et al., 2019) and ensure that the board is to advise appropriately on technology matters.

As Nolan & McFarlan (2005) suggested, the technology committee should be composed of a majority of non-executive directors who are skilled and experienced in understanding the impact of technology in business and society. However, the creation of this dedicated board committee should be evaluated in the context of potential redundancy and intersection with issues that might resurface within the purview of an audit, risk, social and ethics committee. Redundancy in handling technology-related concerns across different board committees can lead to inefficiencies and place extra responsibilities on both management and the board.

P5 has argued that in their case, AI technologies are a topic of three board sub-committees: operations and IT risk committee, social and ethics committee and risk capital, compliance and conduct sub-committee, which is a cross-membership. The social and ethics committee would look after the ethics and conduct of AI, technology, and machine learning. In contrast, risk conduct and compliance committees have a specific role in measuring risk outcomes. King IV (2018) encourages a holistic and integrated approach when the board is establishing board committees. This means that the board should set the committees to achieve the following:

- Effective collaboration through cross-membership between the committee and their members,
- Ensuring complementary work is produced in instances of overlapping of jurisdiction between committees so as to avoid competing approaches, and
- A balanced distribution of power of membership within the committees to avoid dominating member/s and undue reliance on specific member/s.

This approach allows for more effective committees and encourages the board to be all-encompassing and effective in their committee creation. The effective collaboration through cross-membership can also be cascaded down to management-level sub-committees that support the board committees, i.e., ethics steering committees, AI steering committees, risk committees, audit steering committees, etc. At both board and management level, the cross-membership approach can be a key driver in introducing an AI Centre of Excellence (CoE), wherein AI technologies projects and governance are not only the responsibility of IT units but for everyone in the corporation.

An AI CoE is a dedicated unit within a corporation that focuses on advancing the adoption and implementation of AI technologies. Its primary purpose is to promote AI best practices, streamline AI initiatives, and drive AI innovation within the corporation (Davenport & Dasgupta, 2019). The CoE typically consists of a team of experts who collaborate with different business units and departments to identify AI opportunities, develop AI strategies, and oversee AI projects (Dilmegani, 2023). Key to such a CoE is the centralisation of AI expertise, governance, knowledge sharing, and support for AI initiatives across the corporation. The specific structure and roles within a CoE can vary depending on the size, industry, and AI maturity of the corporation. An AI CoE can foster collaboration between board committees, business units, IT departments, and data scientists and encourage an inclusive culture where multi-disciplinary perspectives to solve complex AI challenges have a mutually enriching effect. Corporations such as Deutsche Bank, J.P. Morgan Chase, Pfizer, Procter & Gamble, Anthem, and Farmers Insurance have adopted a CoE model to oversee AI technologies (Davenport & Dasgupta, 2019).

In the context of cross-membership, the participants also argue for the board social and ethics committee as a key oversight body for responsible use of AI technologies. P6 suggested that a social and ethics committee should be the committee to address the responsible use of AI in the organisation (P6). South African listed companies, as well as unlisted companies that reach 500 on their public interest score, are required to have social and ethics committees at the board level. The social and ethics committee and also the audit committee are the only prescribed board subcommittees by the Companies Act[46]. The functions of the social and ethics committee are mainly social, commercial, and

---

[46] 71 of 2008.

environmental matters. As outlined in the Companies Act regulation 43(5)(a), the role of the committee is very broad and covers elements of social responsibility, sustainability, and corporate citizenship, in addition to transformation issues of particular relevance to South Africa and consumer protection. The functions are divided into five categories, namely: (i) social and economic development; (ii) good corporate citizenship; (iii) environmental, health and safety concerns; (iv) consumer relations; and (v) labour and employment.

To the extent that AI technologies positively and negatively impact the stakeholders, society, and the environment i.e., unemployment, AI bias, the risk of abdicating ethical responsibilities, privacy, transparency, accountability and a number of ethical dilemmas, the social and ethics committee seems a fitting space for oversight over adoption of AI technology. The social and ethics committee deals with matters relating to good corporate citizenship, the promotion of equality, the prevention of unfair discrimination, environment, health, and public safety issues, as well as consumer relationships, all of which can be impacted by the deployment and development of AI technologies.

The study views the social and ethics committee as a viable committee to fulfil the role of an AI Centre of Excellence (CoE) at the board level. Thus, the study proposes as an additional role for the social and ethics committee to oversee issues of ethics, data safety and robustness, of fairness, accountability, and transparency. It would oversee the smooth adoption of AI technology overall. Additionally, the role of the committee would be to provide oversight on responsible use of the AI framework, risk assessment, impact assessment, and evaluation of the transparency of the AI's decision-making process. The committee must recommend to the board whether the AI implementation may have a beneficial or potential negative impact on environment, health, public safety, or society as a whole. The social and ethics committee must examine the risk of bias becoming embodied and amplified in the AI technology at issue and propose measures to prevent and mitigate bias. As per the Companies Act requirement (Section 72(8)(e)), the committee should report annually to the shareholders at the company's annual general meeting.

7.7.2. Responsible use of AI governance framework

The sub-question "*Does your board have responsible use of AI governance framework?* was asked as part of the fourth research theme sub-question "What is the board's role in ensuring that adequate AI governance structures are in place?". The question about whether the

board has a responsible use of AI governance framework, as well as exploring the understanding or need for such a framework, is asked to assess the organisation's approach to governing the responsible use of AI and to understand the board's perception of the importance of having a structured framework in place.

By asking this question, the research aims to gather information on whether the organisation already has a responsible use of AI governance framework in place. If such a framework exists, the subsequent question aims to explore the board members' understanding of its role and significance. Understanding their perception of the framework's role helps in assessing how well the organisation addresses ethical considerations, risk management, and accountability in relation to AI technologies.

If the organisation does not currently have a responsible use of AI governance framework, the research aims to explore the board members' views on whether such a framework is needed. This question helped in assessing the board's awareness of the potential ethical and societal implications associated with AI technologies. It provides an opportunity for board members to express their opinion on whether having a structured framework would be beneficial for ensuring responsible use of AI within the organisation.

Responses from board members herein referred to as 'Participants (P)'.

P5 indicated that they have mechanisms in place to assess the responsible use of the AI framework, which is signed off by the social and ethics committee, which is a delegated board subcommittee. The role of the framework is basically to have a consistent way that the ethics and conduct around the use of artificial intelligence, data and machine learning are monitored. "So, that basically tells you how you will behave in terms of the ethics of AI, how do we apply it, so it doesn't tell you the operations that you must go step ABCD. But it tells you these are the things that guide us in terms of how we ethically apply these processes, so the framework would set out the board's role" (P5). The framework prescribes how the board discharges responsibility and delegates to the responsible board committees. For example, "the board delegates, let's say culture and conduct to the social and ethics committee, it delegates the risk to the risk compliance and conduct committee, it delegates the operational implementation to the operations and IT risk committee". The board is ultimately responsible and accountable for the ethical use of artificial intelligence, data use and machine learning in the organisation. "So, it doesn't matter if you're, whether you're

sitting in insurance in Nigeria, or asset management in Zambia, or home loans in South Africa, you've got the same framework and in that is a non-negotiable, board approved framework" (P5).

"One of the boards that I sit on have that, what we have sort of a digitisation process, we don't call it an AI process, because we still retain a level of decision-making. It is early in the process because one other thing that I've got an understanding of is that we do not have data stored in our system so it's very difficult to have this artificial intelligence set-up; you basically need to have what we are now building, have it running for a year or two. And then with the data, now start setting parameters for decision making, which is where the AI becomes handy" (P22).

"Not yet. As far as I'm aware, they all have responsible use of technology frameworks but not AI. The current framework doesn't cover everything, but the practical use of AI in the businesses that I'm working with is still relatively new. So, I do think it's it is needed. So, I will put it on the in the job jar of important things to do" (P23).

"No. I think it will become more topical, I think it will become necessary, and in some businesses more than others. I can see this need in insurance companies and banks for example, but I think in other environments, probably not" (P4).

"No, I'm not aware of any tool that is being used. I sit on a board, which is an industrial and a private equity board. There you need it because we have diversified and disinvestment. And then I sit on a listed estate real estate investment company… there it's also quite sensitive because we've got 12 billion rands worth of properties under management. So, we have put in thousands of tenants, individuals who are renting flats that are owned by this company, who've got all their data, everything from ITC to their movement data, of payment behaviour, or all of those things and who they are with their parents, and where [they] work and you know, we don't know necessarily how management is using that data. They tell us that they apply the POPI act, and we say, okay, yes, until they get into trouble, and then we have to get out" (P6).

"No. We don't have" (P7). However, P7 believes that companies should have the responsible use of AI framework. "No. I have never come across such. I am going to look for it now that you told me" (P8).

"No. I've looked at a couple of documents, and I think it's pretty good; it's essential for boards to consider that and to work with it. I said, especially in organisations that, you know, AI, is dominant and forms a big part of their work … be aware of where the company is, in terms of AI, the use of AI, you know, because obviously, not every company, you know, will want to put a framework in place or major framework, there'll be parts of it, that they may want to put in place and just be aware. Of course, as I said, our AI is mainly used to do machine calculations to improve our capability in terms of doing design work to design roads, bridges, and buildings. So, from an ethical point of view, there's very little when it comes to using information that goes around things like facial recognition, people, you know, all of those sorts of things" (P11).

"No. I can't really put together governance and artificial intelligence, and artificial intelligence is a tool. Governance is an art … governance is the art of moving the company forward … [what is] keeping it under prudent control" (P12).

"No. Companies that are known as technology companies have that. For example, Microsoft and Google have developed the framework. I don't think that South Africa has one yet. I think South Africa is still far it, if I am not mistaken, it's still very much in the POPI part of the equation, thinking just about data governance, but hasn't really thought about responsible AI… I think for Responsible AI governance, by the time South Africa wakes up, we would have sold all our data. Just like we sold all our gold before we knew how valuable the gold was, at the moment, they still are not enough governance that stops bad agents from misappropriating the data of people inside this country just because we don't have enough protection around our IP, and around our data, and we don't fully understand the value of it, or how it can be misused. So, I think how it can be used for surveillance as an example. So, all of those things, which is really how your data gets used against you or how your data gets sold. So, I think it's absolutely necessary" (P13).

P13 argues that it is necessary to have one framework for all companies. They indicated that most companies do not want the AI framework because if it is in the public domain, the law uses such governance frameworks to create a compliance culture, which is costly. "So, at the moment, you're managing it quietly as a governance structure to keep yourself honest. But the moment it falls within the regulatory space, it becomes a real cost because you now have processes, and you have to assign people to manage that as a formal thing. And

nobody wants to do that until they are forced, that means they can hide things that they do wrong ... We all have companies [that] are going to be fined extensive amounts of money for data breaches and issues of bias that are created by algorithms" (P13).

P14 "No, I don't know anything about it". P14 suggests a framework in the banking sector where AI influences business practices. They also indicated that it does not matter what it should be called as long as there is a guidance document or principles which guide decision-making on AI.

"No. Look, whether it's part of the strategy, technology information strategy, so that, you know, it would not necessarily call for a dedicated document, but it would really form part of the strategy on technology and information generally. You know, I think as a subset of those policies and frameworks; definitely, I can see that being heavy in that, but depending, again, depends on the nature of the business. You know, the more AI will have a potential impact on the business, obviously, the higher on the priority list, you know, the topic comes" (P16).

"No. I think it's an interesting one. Because there is like the cybersecurity framework, I can see how AI could come into that as well" (P17). "No. I'm not aware. I'm not aware of them using the framework. It's very important. Because with you know, there isn't even a really good IT management framework. I mean, COSO[47] and ITIL[48] look at it more from a management level" (P19).

Responses from specialists herein referred to as 'Participants (Ps)'.

P3s indicated that they do not have an AI governance framework but suggests that an AI governance framework is required. They argue that the framework sets out a good mechanism to outline tangible control around AI policies; it sets a starting point and the baseline for people to refer to in terms of how the technology should be adopted, executed, and set considered roles and responsibilities. There are significant ethical considerations around AI, and it is important for the organisation to be aligned. P3 further indicated that

---

[47] COSO (Committee of Sponsoring Organizations of the Treadway Commission) was established in 1985 to improve corporate governance, risk management, and combat fraud. It is best known for developing the widely adopted COSO Internal Control Integrated Framework, which helps organisations enhance internal controls, and the Enterprise Risk Management (ERM) Integrated Framework, which guides risk management practices.

[48] ITIL (Information Technology Infrastructure Library) is a best practices for IT management, offering a framework to enhance the quality, efficiency, and alignment of IT services with business needs. It provides structured guidance across various aspects of IT service management, helping organisations optimise their IT operations.

their company is "on a journey of automating and making many of our currently manual processes much more efficient. And in that instance, if you think you might adopt some form of robotic process automation, which is then a form of artificial intelligence, you know, might not be machine learning and the kinds of things people typically default to. But then yes, it is, you know, and so I think what's important is, this kind of framework must also set the tone for and the alignment around what is artificial intelligence" (P3s).

"So, for me, the board needs some sort of a tool, a multidisciplinary tool, to plan and to measure. It can just be dashboards… let's call it responsible technology use framework is perhaps the best way for the board to manage that at a high level" (P1s).

P2s responded that it is a "yes and no". They indicated that they have frameworks and policies around data process, IT tools, and acquisition, all the components that actually make AI work, have now been distributed to various frameworks, but they have not been necessarily consolidated to converge at an AI framework. Thus, there is no framework for AI specifically, but AI has components from data to cybersecurity to its storage platforms (P2s). The framework is needed, as AI is fast becoming the dominant technology, and we are starting to get to a point where you cannot talk about technology without talking about AI in a conversation. "It does probably necessitate a greater focus in one form or shape, whether framework or not, and it might vary from one use case to another; there might not necessarily be a [one] size fit all. But I think it is more investment of existing framework to be co-opted in the data or IT part or wherever we need to see a more amplified view [of] machine and prominence of AI, organisational frameworks and policies" (P2s).

"No, we don't have it". The AI governance framework is a picture that sets out how the board delegates authority for artificial intelligence and then holds accountable those responsible, i.e., management or board subcommittees. Governance is at a board level, and that entails that strategy, oversight, policy, accountability, power, or authority (P9s).

"Yes". P15s indicated that they recently started working on the framework through the data governance committee. The frameworks and standards are a rule book from a guidance perspective of how you should engage with a solution or process (P15s).

"None of the boards that I've sat on have an AI governance framework" (P18s). King IV's report is an excellent start when it comes to the overall governance of technology and risk

governance. P18s argues that increased transparency must not come at the risk of over-regulation of a certain type of sector because it becomes cumbersome, making it difficult to do business in those types of environments. They also do not favour self-regulation in the AI governance framework for greater transparency (P18s).

"To be honest with you, I think, for most of them, it's entirely irrelevant at present because they are not up to the level of AI at all" (P20s).

Analysis

The findings reveal that most companies the participants represent do not have a responsible AI governance framework. Implementation of responsible use of the AI governance framework creates a common language for corporations to articulate common values, management systems, and measures that will sustain trust and ensure integrity among all internal and external stakeholders (Carpenter, 2021). Most of the participants who responded "no" indicated a need for a responsible use of the AI governance framework in their organisation. Only P5 indicated they already have a responsible use of the AI governance framework in their organisation.

The AI governance framework plays an important role in guiding the responsible development, deployment, and deployment of AI technologies. The framework establishes a set of principles, guidelines, and best practices, providing a roadmap for the development, deployment, and adoption of AI technologies. The framework ensures the alignment of AI technologies with societal values, legal requirements, and corporate objectives. P5 indicated that "the role of the framework is basically to have a consistent way that the ethics and conduct around the use of artificial intelligence, data and machine learning are monitored". The framework sets out a good mechanism to outline tangible control around AI policies; it sets a starting point and the baseline for people to refer to in terms of how the technology should be adopted, executed, and set considered roles and responsibilities.

Additional AI governance framework serves as a mechanism for risk management, promotes transparency and accountability, facilitates compliance with legal and regulatory requirements and is a comprehensive tool to navigate the ethical and societal implications. An effective AI governance framework must encompass key principles such as transparency, explainability, accountability, fairness, bias mitigation, safety, and security.

The corporations must adopt a stakeholder-inclusivity approach, which includes ensuring that the public understands and trusts AI technologies, fostering accountability and mitigating the risks associated with biased algorithms. An AI framework should instil accountability, responsibility, and oversight throughout the AI development and deployment process, which will contribute to building a foundation of trust and accountability in the AI ecosystem.

The framework prescribes how the board discharges responsibility and delegates it to the responsible board committees. The AI governance framework is a picture that sets out how the board delegates authority for artificial intelligence and then holds accountable those responsible, i.e., management or board subcommittees. Governance is at a board level, and that entails strategy, oversight, policy, accountability, power, or authority (P9s). For example, the board delegates culture and conduct to the social and ethics committee, delegates the risk to the risk compliance and conduct committee, and delegates the operational implementation to the operations and IT risk committee.

The AI governance framework provides oversight and supervision at the board level while entrusting executive management with core technical aspects, such as data engineering, building models, and choosing best-fit AI platforms. AI governance frameworks assist corporations in governing, monitoring, and providing oversight over design, development, and deployment of AI technologies. The board should also establish clear policies and guidelines within the framework, ensuring that ethical considerations are embedded in decision-making processes related to AI technologies. Regular monitoring and assessment of AI technologies, along with addressing any ethical concerns that may arise, are crucial aspects of the board's responsibility in implementing and sustaining an effective AI governance framework.

7.7.3. Board monitoring performance and value delivery of artificial intelligence technologies

The sub-question "*How does the board monitor the performance and value delivery of artificial intelligence technologies?*" was asked as part of the fourth research theme sub-question "What is the board's role in ensuring that adequate AI governance structures are in place?". The question was asked to understand the board's oversight mechanisms and

processes for evaluating the performance and value generated by AI technologies within the organisation.

The question helps establish the context by addressing the need for board members to be actively involved in monitoring the performance of AI technologies. It highlights the importance of understanding how the board evaluates the effectiveness, efficiency, and impact of AI technologies on the organisation's strategic objectives. Understanding the board's monitoring processes allows for an assessment of the board's awareness of the potential risks, challenges, and opportunities associated with AI technologies. It also provides insights into the board's commitment to ensuring that AI technologies are aligned with the organisation's goals and objectives. By understanding how the board monitors value delivery, the research seeks to determine the board's understanding of the broader impact of AI technologies on the organisation's success and stakeholder satisfaction.

Responses from board members herein referred to as 'Participants (P)'.

P7 indicated that their organisation uses analytical intelligence systems to track online sales and the success of online strategies and to identify opportunities. The system can take from two different data sets and correlate the sales data and opportunities to explore. The board sets out key performance indicators (KPI) in terms of technology and use of technology. P19 added that the KPIs are included in different parts of the IT (P19). "The board will then look at the sort of the KPI elements, the return on the investment on the system implemented. So, the return of the investment can be measured in … various way[s]" (P22). The board monitors the delivery of the AI against the operational results.

The audit committee's role is to monitor the cost versus the budget of the AI system, such as getting the reports back from management. The board provides oversight and monitors investment returns (P22). The board monitors performance based on the promise of return on investment. The board will look at the return on investment, and monitors capability and achievement since AI technology was introduced and how it has enabled the organisation (P8). "Anything around technologies and performance would typically be part of normal operations. The evaluation will also be on the efficiencies of our system, what they've done to improve efficiencies, and what they do to improve our ability to deliver and grow our business" (P11).

"Ultimately, I think, is it helping the business, … that's going to be how most boards assess performance. So, if the business is extracting value, they'll be happy. I don't know that our boards are mature enough to figure it out. You might be extracting value, but it might not be the best and most efficient... Boards look at performance as a whole, and it's multi-dimensional. It's the money, people, and the technology all together give you an outcome or output that that they can measure" (P4).

"So, at the moment, the way that the board measures is purely on its outcome. So, the board has no measure of what artificial intelligence does or what technologies do, and I'm not sure necessarily that it needs to. I think it needs to have a metric for it, but I'm unsure if its job is to monitor artificial intelligence" (P13). The board currently looks at products and services that are being produced by utilising artificial intelligence, but it doesn't have a performance dashboard (P13).

P24 states that the company has an IT steering committee at an EXCO level which deals with performance. The board only assesses how the company is faring relative to the approved project plan. P24 further indicated that the board doesn't get involved in the detail, but tracks the progress of emerging issues, identifies potential risks that management may have overlooked, and considers risk mitigation (P24).

"The reality is that most boards do not have the expert technology expertise to ask the right kind of questions or interrogate the information that comes from management. One of the key themes coming through in the King IV report is this concept of independent assessments on a regular basis, every three years or five years, to get independent experts. Again, it will depend on the materiality of the organisation, our risk appetite, and risk tolerance and all those concepts that will identify the areas where we need an independent voice to come and look at what's happening here. For example, these days, we have this on more than one occasion; now, I have been exposed to what they call ethical or organised hacking, to see, without warning management, authorised to try and see whether they can hack the system. So that's the kind of things that would have to come from the board, risk committees, sort of radar screen and work plan… I think we are going to need independent assessments of AI in this current environment in which we operate, and the rapid development and pace with which information and technology are developing, boards, or, in my opinion, definitely not, in front of this we lacked very much behind" (P16).

P5 indicated that it would be hard to separate AI performance from normal business performance because they are effectively a technology and data-driven company; the bank part is just regulation. AI is just another technology that would be deployed. The board would look at the extent to which technology is being deployed and would apply traditional measures such as delivering on schedule, according to budget and realising the benefits anticipated from the deployment. Furthermore, the board assesses the potential risks associated with it and monitors the risks, risk tolerance and the associated risks (P21).

Most participants did not answer how the board monitors the performance and value delivery of artificial intelligence technologies.

Responses from specialists herein referred to as 'Participants (Ps)'.

P1s suggested that performance metrics need to be agreed upon upfront; the board is primarily interested in increasing market share, which could be achieved by delivering better service to customers in a more cost-effective way possible. However, P1s argued that the board cannot just measure the bottom line, even though it's very important. The board should have a multidisciplinary committee which should develop a matrix that impacts the environment, people, and society (P1s).

"I think we will monitor this more from how decision-making and machine intelligence [have] improved [and] which paths can be attributed to what. So actually, from an organisation perspective, we actually very much talk data and not so much AI, but to be able to exploit data effectively, you require AI. So, in that way, we monitor how as an organisation, we are becoming more data-driven and effectively using data to understand better revenue opportunities. So, there is a bit of an interchange between usage data and AI, but I think we'll just broaden it up to say there are cases that don't necessarily require a complex application of AI technology. But the simple data is a solution that the organisation can provide" (P2s).

"Probably because we don't have it in place, it will fall under the IT standards currently" (P15s). On a monthly and quarterly basis, a report is sent to the board on breakages, outages, data leakages, or user access that does not adhere to the standards. In the case of the bank, if the incident is critical, they report it immediately to the reserve bank with a mitigating plan. The board's role is oversight and monitoring (p15s).

Most specialists did not answer how the board monitors artificial intelligence technologies' performance and value delivery.

Analysis

Corporations are significantly investing in AI technologies to achieve a competitive edge in the market; thus, it is important to understand how the board monitors the corporation, which means growing revenue, increasing performance, improving customer satisfaction, increasing market share, reducing cost, enabling new product/or service, etc. Therefore, it is important for boards to evaluate their effectiveness, efficiency, potential risks, challenges, and opportunities associated with AI technologies on organisational goals and strategic objectives.

As a subset of IT governance[49], one of AI governance focus areas is ensuring performance and value throughout the delivery cycle, ensuring that AI technologies deliver the promised benefits against the strategy, optimising costs and proving the intrinsic value. AI performance and value delivery should be part of the AI strategy that ensures alignment of AI investment with business objectives and demonstrates the value of AI technologies to stakeholders in the corporation.

Some participants indicated that they use key performance indicators to measure AI technologies' performance and value delivery. As highlighted by (P19), the board sets out key performance indicators (KPI) in terms of technology and use of technology. KPIs are measurable values that indicate how the corporation is achieving its goals and objectives. Existing KPIs, which are used to measure performance and delivery for IT projects, also apply to AI technologies such as mean time to repair (MTTR)[50] and first contact resolution rate (FCRR), which indicates what percentage of problems are resolved by level 1 IT support (basic support) without needing escalation (Murphy, 2023).

KPIs can be used as indicators that reflect on the contribution of AI technologies to the corporation's strategic objectives, such as growing revenue, increasing performance, improving customer satisfaction, increasing market share, reducing cost, enabling new

---

[49] Chapter 4: AI governance is a subset of corporate governance and IT governance (Mäntymäki et al., 2022).
[50] MTTR (mean time to repair) is the average time it takes to repair a system (usually technical or mechanical).

products/or services, etc. This means that KPIs can measure AI technologies' success by demonstrating a concrete return on investment (ROI), expressed as time, money, or labour. The KPIs indicators should be aligned with the strategic objectives of the corporation. As suggested by P1s, the board should not only measure the bottom line (profitability), but they should have a multidisciplinary committee in this case, the social and ethics committee, which should develop a KPI matrix that also measures the impacts on the environment, people, and society.

Another best practice which can be used to measure the performance and value delivery of AI technologies is the balanced scorecard. Kaplan and Norton initially developed the balanced scorecard (BSC) as a performance management system that enables businesses to drive strategies based on measurement. In recent years, the BSC has been applied to information technology (IT) to measure performance and return on IT assets investment (Van Grembergen & De Haes, 2005). BSC can facilitate the link between AI activities and outputs to the business outcomes and value metrics and provide a holistic view of AI performance. The BSC can help monitor and communicate the performance of AI technologies across four perspectives: financial, customer, internal, and learning and growth, as depicted in Figure 15 below.
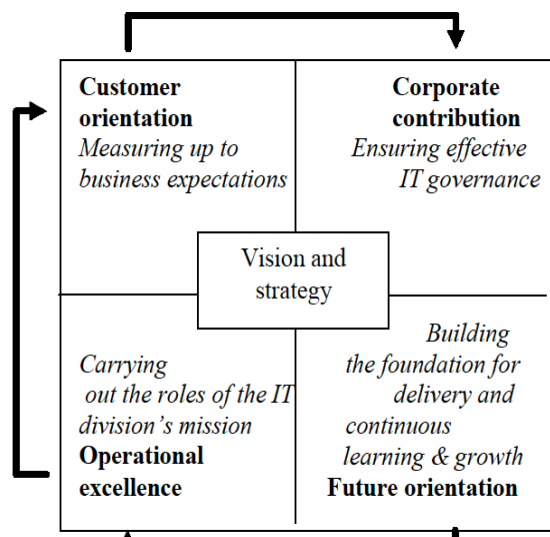

Figure 15: Standard IT Balanced Scorecard (Grembergen, 2000)

- **Corporate Contribution Perspective:** This perspective aims to achieve business contributions to AI investment, which includes measures such as operating income, profitability and return on investment.

- **Customer Orientation:** Focuses on evaluating AI performance from the view of the users (internal and external), i.e., customer satisfaction, AI integration, successful development or deployment and service success rate.
- **Operational Excellence Perspective:** This perspective focuses on how internal business processes are linked to strategic goals. Operational excellence has an important contribution because it results in two things: product quality and AI cost emphasis.
- **Future orientation:** Represents learning and growth. This should include holding training on a regular basis for those responsible for developing and deploying AI technologies to improve their expertise and ensure that they are aligned with corporate values.

Grembergen (2001) argued that each of these four perspectives should be transformed into matching metrics and measures to assess the current environment. The assessments need to be continuously conducted and aligned with the pre-determined objectives and benchmarks with best practices. A well-developed scorecard contains a good mix of outcomes, measures, and performance drivers.

7.8. CONCLUSION

In this chapter, the researcher presented the data collected from the semi-structured interviews. These interviews, which were conducted with twenty-four research participants (board members and specialist participants), have yielded insights that were pivotal to our research. As a result, the data collected was infused with qualitative depth from the participants' lived experiences and personal narratives as board members. The chapter further highlights the analysis of the data collected from participant interviews. After familiarizing with participant data, the researcher developed significant statements, which were then grouped into a theme. The findings and analysis from the semi-structured interviews were presented as sub-theme sections 1 to 4. The insights gleaned from these interviews are instrumental in furthering our understanding of the role of the board in governing artificial intelligence ethics, and this chapter interpreted each facet of research sub-themes, setting the stage for the discussions and recommendations in the next chapter.

CHAPTER 8: SUMMARY AND RECOMMENDATIONS

8.1. INTRODUCTION

The chapter provides an overall summary of the thematic analysis and provides recommendations in terms of each set of findings related to each sub-theme. The research question, which seeks to explore the role of the board in governing AI technologies, is thoroughly addressed through the alignment of the adopted corporate governance definition from the King IV report and the stakeholder-inclusive approach with the research findings. The adopted definition of corporate governance from the King IV report, which emphasises ethical culture, good performance, effective control and legitimacy and the creation of sustainable value for a broad spectrum of stakeholders include the society as whole, aligns seamlessly with the findings of this research. This perspective is rooted in stakeholder theory, which underscores the interconnectedness and influence of various stakeholders in the operations of the corporation. The study highlights the board's crucial role in adopting an inclusive approach to AI governance. This holistic view ensures that all material stakeholders are integral to strategic decision-making processes, fostering a governance framework that is ethically sound in the context of AI technologies.

Furthermore, the stakeholder-inclusive approach is reinforced by integrating principles from agency theory. Traditional agency theory has often been criticized for its narrow focus on shareholder primacy, which can lead to conflicts between principals and agents. Expanding this scope to include all stakeholders helps to mitigate these agency conflicts through improved transparency, accountability, and ethical approach to AI technologies. The findings, which advocate for AI training at the board level, inclusion of tech experts on boards, and the establishment of robust AI risk governance, reflect this comprehensive approach. The recommendations in this chapter such as board stakeholder engagement and the creation of a technology governance committee, social and ethics committee demonstrate a commitment to a stakeholder approach and inclusive governance framework. This approach not only aligns with the King IV definition but also extends it by ensuring that AI governance is integrated into the broader corporate governance landscape, thereby promoting responsible and ethical use of AI technologies within JSE-listed companies.

8.2. SUMMARY AND RECOMMENDATIONS OF ANALYSIS OF THE BOARDER RESEARCH QUESTION THEME ONE – HOW THE BOARD DISCHARGES ITS FIDUCIARY RESPONSIBILITY IN OVERSEEING AI TECHNOLOGY.

This sub-theme aimed to understand corporate governance comprehensively, particularly the board's fiduciary role in governing AI technologies. Additionally, the questions in this sub-theme aimed to explore whether board members have a basic understanding of corporate governance in the context of AI technologies, including the ethical implications of AI technologies. This helped establish the overall essence of the corporate governance concept, such as the fiduciary understanding of the board's role concerning AI technologies governance.

In this case, participants (board members and specialists) could comprehend and simplify the definition of AI and capture the core themes surrounding the concept of artificial intelligence based on their experience as board members. The participants collectively reflected on various facets of AI, ranging from its departure from rule-based AI to its role in automating tasks, simulating human-like behaviour, enhancing human capabilities, and representing both a technological and cognitive endeavour. The interplay of these perspectives in participant responses highlights AI technologies' diverse and evolving nature as a field that continues to shape our technological landscape. Asking the participants about their understanding of artificial intelligence was an essential step in the research process to establish whether the board members had a basic understanding of AI technologies.

Regarding the ethical concerns, the participants' key concerns were mainly biased decision-making, job displacement and privacy & security violations. The general agreement amongst the participants was that the directors need a basic understanding of AI risks, which may emanate from biases, job displacement and privacy & security violations. The board should not only be aware of these risks posed by AI technologies but also actively address them. Thus, understanding AI technologies is crucial for the board members to fulfil their fiduciary responsibility in overseeing their governance effectively and to determine the board's potential preparedness to govern AI technologies effectively and ensure their adoption aligns with the corporation's strategic goals and ethical standards.

The proactive involvement of the board in strategic direction, guidance, resource allocation, setting the ethical tone, holding management accountable, evaluating performance, and

monitoring and oversight, as raised by the participants, ensure that the board and management navigate the complexities of AI technologies while promoting responsible and beneficial outcomes. Boards are responsible for ensuring corporations have an effective and efficient governance structure for setting objectives and monitoring performance. Those structures must provide guidance on how the board, management and other stakeholders participate in decision-making and ensure the board has the information required to provide oversight and that shareholders have information about the corporation's strategy, risks, and performance as these relate to AI governance.

The overall findings reveal that the participants and specialists had a broad understanding of the fiducial duties of the board. The duty requires directors to act with care, skills and diligence when performing their responsibilities as set out in section 73(3)(c) of the Companies Act of 2008, which is applicable to all companies in South Africa. Furthermore, companies should disclose under section 30(1)(b) of the Companies Act of 2008 material information that may affect the company. Thus, if AI technologies are integrated as essential elements of business operations or strategy, the board should disclose information regarding the use, risks, and benefits of AI to shareholders and stakeholders. Additionally, the board should ensure that the design, development, adoption and use of AI technologies comply with data protection, consumer protection, and other relevant laws and that the organisation's AI practices are transparent and accountable.

In this context, the board of directors are required to gather adequate information before making business decisions. This implies that boards bear the responsibility of taking appropriate measures to acquire a deeper understanding of the implications of AI technology integration in the company and exercise diligence when making decisions about AI technology within the company. The fiduciary duties include exercising supervision obligations, even in technical areas, even areas that present enterprise risk, which is already true for AI technologies in corporations (Gesser, Regner & Gressel, 2022). That does not imply that board members must become AI technology experts or that they should be involved in day-to-day AI technology operations or risk management. However, the board should have a deeper understanding of the implications of AI technology adoption and exercise diligence when making decisions to ensure effective oversight with respect to the growing opportunities and risks presented by AI technologies.

However, the participants were concerned about the breach of fiduciary duty due to the digital divide in the boards. They (participants) indicated that some board members do not understand AI technologies. The boards need to be better trained in AI technologies to close this knowledge gap. The board should also consider adding one or more directors with AI or technology experience or having an AI technology expert as an alternative director.

The boards are responsible for taking reasonable steps to gain deeper insights into the impact of AI technology in the corporation. Exercising care, skill, and diligence means that the board, individually and collectively, need to constantly obtain and act in accordance with adequate information and understanding to execute their oversight and control mandate effectively. In practice, working with various stakeholders to develop a basic understanding would help board members better fulfil their duty and ensure that necessary checks and balances are put in place at the various stages of AI development and adoption.

This study recognises that effective corporate governance is vital for the success of any organisation and confirms the realisation of such governance in terms of the board's role in governing AI technologies. The board's fiduciary responsibility ensures the corporation operates efficiently, effectively, and ethically. The rise in the adoption of AI technologies means that the board must also discharge its fiduciary responsibility in overseeing these technologies to ensure they align with the organisation's strategic goals and comply with laws and ethical standards.

To achieve this objective, the board must understand AI technologies and their potential impacts, such as employment risks, data privacy, security, and bias, and address these issues by putting in place processes and structures to ensure that AI technologies are well managed to avoid harm in society. Boards will, therefore, have to ensure that appropriate policies, guidelines and practices with regard to the use of AI are implemented to mitigate the risks involved. Such policies may include procedures for responding to a material AI-related incidents, responding to AI-related whistle-blower complaints, and risk management for any vendors that supply the corporation with critical AI-related resources (Gesser, Regner & Gressel, 2022). These policies should set out where AI technologies will be used, what checks and balances will be put in place in terms of oversight, and set standards for robust and safe operation. Finally, the board has to publish an integrated report on how the company governs AI ethics openly and transparently in line with the recommendations by the World Benchmarking Alliance (2021).

8.2.1. Sub-Theme 1: Recommendations

8.2.1.1. *AI training at the board level*

The participants were concerned about the breach of fiduciary duty due to the digital divide in the boards. At this level, the digital divide refers to the extent of the digital challenge and the digital awareness of boards members. With the rise of the digital economy, we are entering a new era of managerial innovation with both opportunities and risks (Bonnet, 2014). Digital transformation is a top-down leadership exercise, and boards cannot remain isolated from leading this fundamental change. However, some of the participants indicated that some board members do not understand AI technologies.

This study has argued that AI technologies can potentially transform many industries; therefore, boards can no longer afford to ignore these technologies' rapid advancement. AI technologies present a great opportunity for the corporation to have a competitive edge, but they also present risk challenges which have a high impact on society. Suppose the boards do not understand AI technologies and their implications. In that case, they will fail to provide oversight, control, and strategic direction on the impact of AI and related decisions regarding investments, risk, ethics, digitalisation, and its stakeholders.

To close this gap, board training is essential for board members to better comprehend AI risks and opportunities. Training helps the board ensure responsible design, development, deployment, and use of AI technologies by mitigating risks, taking advantage of opportunities, fostering ethical consideration and compliance, and building public trust. Such training should include comprehensive modules in algorithmic bias, safety, privacy, accountability, transparency, responsible AI, alignment to strategic objectives, AI governance and ethics, AI risk management, compliance with regulations, etc. The training should be multi-disciplinary in approach to include various subject matter experts such as AI researchers, ethicists, legal experts, policymakers, industry representatives, and civil society advocates.

In conclusion, the approach to the training of the board should be proactive, adaptive, and aligned with strategic objectives and the interests of all stakeholders. It should balance promoting AI technologies with ethical considerations and mitigating risks to ensure that AI

technology adopted or designed in the company benefits all stakeholders and society and abides with Human Rights Law. As difficult and complex as the technical concepts may be, boards must nevertheless increase their engagement with the risks and opportunities of AI technologies if they are to fulfil their fiduciary responsibilities. Therefore, companies must invest in board training and development, i.e., IODSA board training, to help directors comprehend and get up to speed on the latest developments of these complex and constantly evolving AI technologies.

8.2.1.2.   Board composition - Tech expert on the board

The findings reveal that most boards do not have tech specialists or experts with technology knowledge at the board level. This is also backed up by the global study by Gordon (2021), which shows that most boards do not have sufficient technology depth and knowledge expertise. The majority of the board still have traditional knowledge, particularly in finance and the law. The participants indicated that it would be imperative for boards to have one or more professionals at a board level as well as in a specialised committee of the board who will understand the dynamics, complexities, implications, and outcomes of using the technology.

To address this gap, the board nomination committee should evaluate the corporation's preparedness to handle AI technologies' integration by mapping the technological expertise required to their current non-executive directors' knowledge pool. This should be done to determine if anyone from the current board qualifies to be an expert, or if there is a need to bring in alternative director(s)[51]. At a basic minimum, the board nominations committees should consider having one or more directors with AI technology experience or having an AI technology expert(s) as an alternative director(s). The board nomination committees must ensure the right mix of knowledge, skills, and experience to effectively oversee AI technologies. For companies highly reliant on AI technologies, such as banks, insurance and financial industries, this study recommends that at least two board members be AI or IT experts; all other JSE-listed companies board at least one AI or IT expert at the board level. Where a company struggles to find an AI or IT expert with the broader business expertise

---

[51] 'Alternate director' means a person elected or appointed to serve, as the occasion requires, as a member of the board of a company in substitution for a particular elected or appointed director of that company" (Companies Act 76 of 2008).

required of a board member, the study recommends an AI or IT expert alternative director(s) who can serve at the board level as and when needed.

Experts are recruited to bridge the knowledge and communication gaps in the executive management. The role of the expert(s) will be to lead the board in asking the right technical questions to the executive management and help monitor and oversee internal processes and the entire governance infrastructure for developing and deploying AI technologies. The expert must also thoroughly understand the underlying dynamics governing changes in AI technologies, and potential risks to be mitigated and opportunities which can be exploited. AI technologies impact every aspect of a company's operations risk, remuneration, cybersecurity, finance, and strategy. Thus, AI expertise needs to be widespread so that the full board and all its committees can properly consider the implications of engaging with AI technology in the company.

8.2.1.3. *AI risk governance at the board level*

The general agreement amongst the participants was that the directors need a basic understanding of AI risks, which may emanate from biases, job displacement and privacy & security violations. AI technologies bring new risks and opportunities which most boards have not previously dealt with. This means that the board needs to understand the risks associated with AI technologies and ensure that all the audits and governance processes are aligned to identify any risks upfront and manage them accordingly. They also need to understand the opportunities that AI technology offers the corporation and be able to tailor strategies towards realising this potential in the best way in their corporation.

It has to be underscored that as corporations take advantage of AI technologies, boards and board-level risk committees need to pay attention to the potential risks of AI and the various ethical issues that may arise. King IV principles 11 & 12 give clear direction that the board is responsible for risk and technology governance in the corporation. To achieve this, the board must understand AI technologies and their potential impacts by putting in place processes and structures to ensure that AI technologies are well managed to avoid harm in society. The board structures to deal with AI risks should include a board audit committee, social and ethics committee, AI steering committee, investment or finance committee, and effective management level structures such as a management risk committee, and IT or AI

steering committees to continuously ensure risk assessment, management and monitoring of AI technologies risks.

In terms of the processes, the board should ensure the assessment of AI risks and proper risk management processes are in place. This should involve identifying potential risks, evaluating their impact, developing mitigation strategies, and taking advantage of opportunities. As such, the board should ensure an AI risk management framework which will prescribe risk mitigation and monitoring mechanisms. The AI risk framework ensures the following at minimum:

- AI alignment with strategic objectives
- Acceptance levels of risk tolerance and appetite
- AI risk impact assessment processes
- AI risk management processes
- Processes of integrating other frameworks such as COBIT, ISO, COSO enterprise risk management framework, etc.
- AI risk oversight structures
- Continuous self-learning and change processes.

The board needs to ensure that the risk management framework is a subset of the AI governance framework and that the two are integrated to ensure efficiency and effective AI governance.

8.3. SUMMARY AND RECOMMENDATION OF THE BROADER RESEARCH QUESTION THEME TWO (2) – WHAT IS THE BOARD'S ROLE IN ENSURING THAT AI TECHNOLOGIES ARE ALIGNED WITH THE ORGANISATION'S ETHICAL CULTURE.

For this sub-theme, three questions aimed to explore different aspects of the board's role in aligning AI technologies with the organisation's ethical culture, including the board's role in setting the ethical tone and understanding the potential impact of AI technologies on the organisation's ethical culture.

In essence, the finding reveals that the board must ensure that the organisation's ethical standards are integrated into developing and deploying AI technologies, which means that

the board should govern ethics to ensure that the ethical culture within the corporation aligns with the tone set by the board. To ensure such, the board must set the ethical tone from the top and ensure it trickles down to the entire organisation. The board sets the tone by (i) exemplary leadership; (ii) cultivating ICRAFT (integrity, competence, responsibility, accountability, fairness and transparency) principles; (iii) establishing governance structure, i.e., audit committee and social and ethics committee; (iv) putting policies such as a code of conduct in place: and (v) ensuring alignment of organisational strategy and ethical culture to AI technologies.

The research acknowledges that AI technologies have the capacity to transform various aspects of organisational operations, decision-making processes, and interactions with stakeholders. The findings reveal that the rapid advancements and changes in AI technologies necessitate a combination of performance improvement and a willingness to learn from everyone involved, which can lead to a corporation's competitive advantage. AI technologies can improve organisational productivity (financial, marketing, and administrative) and streamline organisational processes for effective decision-making and tasks (Arslan et al., 2021). In the right setting, AI technologies can enhance the performance of individuals and teams, streamline workflows, and even complete some tasks to enable employees to focus on more meaningful work and work-life balance.

However, the findings also reveal the negative impact of AI technologies on the organisational culture. The participants highlighted issues such as unemployment, which is exacerbated by AI's potential to replace certain jobs, risk of abdicating ethical responsibilities, and privacy, transparency, accountability, and ethical dilemmas. As such, the board needs to consider the potential impact of these technologies on the organisation's ethical values, norms, and practices – and even the organisation's ability to comply with national legislation and International Law. Understanding the potential impact of AI technologies on the organisation's ethical culture is crucial because it enables the board to proactively address ethical concerns and ensure that AI technologies are developed, deployed, and utilised in a manner that aligns with the organisation's ethical values and standards. It also highlights the need for the board to provide guidance and oversight to management in navigating the ethical complexities associated with AI technologies.

Thus, the findings also reveal that incorporating AI technologies into an organisation's operations and decision-making processes requires careful consideration of ethical

implications. The strategies employed to develop, deploy, and utilise AI technologies should be aligned with the organisation's ethical values, principles, and goals. AI ethics should be a logical extension of a corporate strategy and objectives. An AI ethics statement should serve as a blanket declaration that the design, adoption, and use of AI technologies are consistent and ethically applied. For the purpose of enhancing transparency, the statement should include the hierarchy of ethical values of the corporation.

The alignment of organisational cultural norms with AI technology development and deployment is key to the uptake and efficacy of explicit rules and regulations. Individuals are more likely to adhere to the organisational culture and norms they believe in (Seger, 2022). Aligning AI technology strategies with an organisational culture and norms is crucial to ensure that AI technologies and applications are designed, developed, deployed, and used in ways that are consistent with the organisation's values and principles. It further promotes responsible and ethical AI practices, enhancing customer trust and stakeholder confidence and ultimately leading to sustainable business success. The board plays a crucial role in this alignment by providing oversight and guidance. As stated before, they are responsible for setting the ethical tone at the top, establishing clear ethical guidelines, and holding the CEO and executive management accountable for AI-related decisions and actions. The board's involvement ensures that ethical considerations are integrated into the corporation, making ethical AI technologies a fundamental component of its long-term vision and success.

8.3.1. Theme 2: Recommendations

8.3.1.1.   Tone from the top – practical steps for setting ethical tone

In terms of King IV, the board is tasked with governing the organisation's ethics in a way that supports the establishment of an ethical culture. However, the study reveals some gaps wherein some boards are found to be "just window dressing with framed pictures of values" (P16, P9s) and "slogans and posters on the walls" (P12) and are not leading the establishment of fundamental ethical values for corporations. Boards that merely give the appearance of an ethical culture and not setting an ethical tone or fail to adhere to ethical principles can have a negative impact on the corporation. High ethical standards are in the corporation's long-term interests as a means for brand reputation, not only in day-to-day operations but also with respect to longer-term success.

The board oversight of AI technologies can drive the right tone and set the guiding principles for an ethical framework that can be operationalised at the executive management level. The study recommends the following for setting an ethical AI tone at the board level.

- **Board Ethical Leadership**

The board must provide leadership and guidance on how the organisation should exploit AI technologies. It should closely examine the executive management team to ensure that they are committed to the responsible use of AI technologies. This includes hiring executives who cultivate integrity, competence, responsibility, accountability, fairness, and transparency, providing ethical leadership training, and holding the executives accountable. Lastly, it is necessary to ensure the establishment of AI governance structures at the board and executive levels to support responsible AI deployment and development.

- **Implement Robust AI Ethics Policies and Procedures**

Boards should work with management to establish clear, comprehensive, and well-communicated ethics policies and procedures that guide the responsible development and deployment of AI technologies. The content and principles embodied in the Code of Conduct must be integrated into employee training. The board must regularly review and update these policies to ensure they remain effective.

- **Transparency, Explainability, and Accountability**

Boards should demonstrate a commitment to transparency regarding AI's ethical impact on the corporation and society at large. The board should ensure that members of society are aware if AI technologies are making decisions that impact their lives, and to understand how those decisions are reached. Transparency leads to trust in AI technologies when humans are able to understand, interpret, and explain the AI decision-making process. Such ability to provide interpretable explanations of AI decision-making processes is also linked to the accountability of AI technologies as well as their explainability. Transparency and explainability of AI processes play a very important role in the overall strive to develop more trustworthy AI technologies.

The board should also ensure human control is linked to accountability and that AI technologies are developed or deployed in such a way that humans can intervene in their actions or decision-making. The capability for human intervention during both the design and implementation, as well as the use of AI technology, including the possibility to override a decision made by the AI technologies when it violates the law, is paramount to ensure greater accountability of AI technologies.

Lastly, the board must ensure that AI is a standing agenda item in board meetings, openly discussing progress and holding the executive team accountable for fostering AI ethical culture.

- **AI alignment to organisational culture**

Ethical culture requires boards to state the values and AI ethics the organisation adheres to, and adherence to these ethical values and principles across the organisation's AI technologies must be ensured. The alignment of organisational cultural norms and values with AI technology development, deployment, and use, is key to the uptake and efficacy of explicit rules and regulations. Therefore, the board must ensure the alignment of AI technologies with corporate strategy, ethical values, and overall corporate objectives to explicitly translate relevant ideas into specific norms that influence concrete design and governance to shape the development and use of AI technologies. To align AI technologies with corporate values and business ethics implies that explicability, transparency, accountability, fairness, and responsible use of data are all essential pillars of ensuring the responsible use of AI technologies.

8.4. SUMMARY AND RECOMMENDATIONS OF THE BROADER RESEARCH QUESTION FOR THEME THREE (3) – THE BOARD'S ROLE IN ENSURING THAT THE RESEARCH OR DESIGN, DEVELOPMENT, OR DEPLOYMENT, OR USE OF AI TECHNOLOGIES DOES NOT NEGATIVELY IMPACT SOCIETY.

The third theme focused on the board's role in ensuring that AI technologies do not negatively impact society. Also already outlined under chapter 7 (7.6.1), the theme assessed the board's responsibility in mitigating potential societal harm arising from the various stages of AI technologies within the corporation. Only one broader question was asked on boards understanding their responsibility in mitigating societal harm from AI technologies. The

context of this question aimed to assess the perceptions of the board's awareness of the potential risks, challenges, and unintended consequences that may arise from the organisation's involvement with AI technologies. It also emphasised the board's role in ensuring that the corporation operates in a way that aligns with ethical standards and avoids or minimises harm to society.

The findings revealed that it is essential for boards to adopt the triple bottom approach (balancing financial, environmental, and societal priorities) to mitigate against potential societal harm. Therefore, the board should ensure that the corporation is seen as a corporate citizen (operating responsibly and ethically towards society). To be considered responsible corporate citizens, the boards need to consider the responsible corporate use of AI technologies and the potential impact on key stakeholders such as employees, customers, society, and the environment. The board must ensure that AI technologies comply with the constitution, laws and standards, company policies and procedures, strategy, codes of conduct and ethics, and that potential risks are monitored and reviewed regularly.

The findings further revealed that the board needs to ensure partnerships with suppliers of artificial intelligence if outsourcing and engage with key stakeholders, including employees and customers, to better understand the potential risk and impact of AI technologies and identify ways to minimise any negative impact. This may involve working with multi-disciplinary experts and engaging in dialogue with diverse groups to ensure that the voices of those most affected by AI technology are heard and considered.

The board should establish a social and ethics committee to ensure that AI technology development, deployment and adoption do not impact people and the environment. To avoid environmental harm and minimise the release of dangerous compounds, the board must adopt a cradle-to-grave approach, which includes proper e-waste management and recycling to ensure safe disposal to avoid the risk of harm to society or the environment. It should also ensure data wiping of personal information in line with Protection of Personal Information Act 4 of 2008 ("POPIA"), which places an obligation on responsible corporation to delete or destruct personal information under its control. POPIA section 14(5) state that personal information must be destroyed or deleted in a manner that prevents its reconstruction in an intelligible form.

Big data centres have both enormous water and electricity footprint (Thompson, 2023). Specially large ones like GPT-3 and GPT-4 for language services use an enormous amount of water for both on-site cooling and off-site electricity generation, collectively accounting for about 1-2% of the global electricity usage. For example, Google, Microsoft, and Meta reached an estimate of 2.2 billion cubic meters in 2022, equivalent to the total annual water withdrawal of municipal, industrial, and agricultural usage (Li et al., 2023). Thus, the board should ensure that should also ensure effective use of electricity and water if there is engagement with big data analytics.

The board plays a critical role in promoting transparency and accountability around the design, deployment, and development of AI technologies and establishing mechanisms for addressing any concerns or risks related to AI. The board's responsibility in mitigating societal harm from AI technologies is to ensure that the company is acting responsibly and ethically. The board's fiduciary oversight obligations extend to the use of AI technologies; thus, it should understand how AI impacts the corporation and its obligations, opportunities, and risks, and apply the same general oversight approach as it applies to other management, compliance, risk, and disclosure topics.

8.4.1. Theme 3: Recommendations

8.4.1.1. Board stakeholder engagement – Development, deployment, and overall adoption of AI technologies

The findings reveal gaps in stakeholder engagement in the development, deployment, and overall adoption of AI technologies driven from the board level. The JSE-listed companies in South Africa have adopted the stakeholder-inclusive approach in line with the King IV report. This means that the board should identify material stakeholders and their interests and expectations relevant to the corporation's strategic objectives and long-term sustainability. Therefore, before the development, deployment, and overall adoption of AI technologies, the board should consider the impact, both positive and negative, on all legitimate stakeholders, such as employees (threat of job displacement), society (bias or harm), customers (privacy and security), environment (safe disposal), etc. The impact on different stakeholders should be determined on a case-by-case basis.

Incorporating stakeholders in the AI environment would result in collaboration among various stakeholders during the planning, design, development, implementation, and use of AI technologies. Stakeholder participation throughout the AI technologies life cycle is crucial to achieving AI governance AI ethics, improves efficiency, and contributes to the long-term sustainability of the corporation.

The social and ethics committee must ensure that the decisions that are made are inconsideration of communities that the corporation is servicing. The social and ethics committee needs to ensure stakeholder engagement to the extent of open communication around interventions being affected to ensure remedial action is executed. Stakeholder engagement should take into account (in line with UNESCO and World Benchmarking Alliance recommendations) shifts in technologies and the emergence of new groups of stakeholders and allow for meaningful participation by marginalised groups, communities and individuals and, where relevant, in our case, previously disadvantaged groups, in respect of data protection and fairness.

As already indicated, the World Benchmarking Alliance (2021) recommend that corporations (i) should be transparent about how they are using AI technologies and should ensure that their AI systems are explainable and accountable to both their stakeholders and the wider public; (ii) be proactive in promoting the responsible development and use of AI technologies, and should engage with relevant stakeholders and the wider community, to ensure that the impact of their AI systems is positive; (iii) support the development of ethical frameworks and guidelines for the development and use of AI, and should ensure that their AI systems are designed and operated in accordance with these frameworks.

In dealing with the overall global challenge of AI ethics and governance, this study recommends that the JSE listing requirement should be mandatory for stakeholder engagements before the development, deployment, and overall adoption of AI technologies. Therefore, the listing requirement should expand the role of the social and ethics committee to include AI technologies. For all other corporations, the future King V corporate governance report should have an additional principle for AI governance, which should include transparency, fairness, safety, explainability, integrity, accountability, robustness, and human in the loop. Incorporating responsible AI principles in the King's report will act as an opportunity for boards to gain a deeper understanding of AI technologies' capabilities and

ensure responsible governance of AI technologies, as well as integration with institutional governance structures and culture.

8.4.1.2.   A cradle-to-grave approach

In order to avoid the risk of harm to society or the environment from AI technologies, a cradle-to-grave approach is recommended. This approach refers to a comprehensive and environmental perspective that considers the entire AI-powered hardware lifecycle, from its development (cradle) to its ultimate disposal (grave). This approach aims to promote sustainability, minimise harm to society, reduce environmental impact, and ensure that technology is disposed of responsibly. Therefore, as part of the long-term sustainability, the board should ensure that the materials used for AI-powered hardware are environmentally friendly and energy and water efficient (see 8.4.1.4 below). At the end of the product life cycle, the board should ensure components are recycled properly, and proper e-waste management processes are in place to prevent hazardous materials from entering the environment.

8.4.1.3.   Data wiping

Data wiping is increasingly important as the rapid pace of technological advancement leads to higher turnover rates for electronic devices. Data wiping is basically a process of securely deleting all information from a storage device (Park & Kim, 2020), such as a hard drive or solid-state drive, to ensure that the data cannot be recovered. This process is crucial when decommissioning or repurposing hardware to prevent unauthorised access to sensitive information. To protect the information, especially sensitive personal information, the board should have processes to ensure that the hardware is securely wiped, erased, destroyed, or repurposed. This should be in line with POPIA section 14(5), which states that personal information must be destroyed or deleted to prevent its reconstruction in an intelligible form.

8.4.1.4.   Electricity and Water Usage by Big Data analytics

The energy, water, and carbon implications of AI technologies can no longer be ignored (Martin, 2023), especially those using big data analytics. Thus, the board must ensure responsible water and electricity usage as part of addressing environmental sustainability, i.e., implementing energy-efficient technologies, sourcing renewable energy, and employing

water conservation measures. The board must oversee and monitor compliance with environmental laws and encourage transparent reporting on water and electricity usage. The board should ensure future data centres are eco-friendly and geared towards a more sustainable and responsible approach to resource utilisation, aligning business operations with broader environmental and societal goals.

8.5. SUMMARY OF ANALYSIS OF THE BROADER RESEARCH QUESTION THEME FOUR (4) – THE BOARD'S ROLE IN ENSURING THAT ADEQUATE AI GOVERNANCE STRUCTURES ARE IN PLACE

As AI technologies become increasingly integrated into various aspects of business operations, it is crucial for corporations to have robust governance structures in place to ensure responsible and ethical development, deployment, and adoption. AI governance encompasses structures, rules, practices, processes, procedures, and technological tools employed to ensure AI technologies align with corporate strategies, objectives, values, and ethical principles (Mäntymäki, 2022). The thematic question aims to explore the extent to which board members recognise their role in overseeing and establishing these governance structures.

1. Governance Structures - Board Subcommittees

The findings reveal that most boards have subcommittees at the board level that deal with broader technology issues. The board with subcommittees uses them to provide additional technical expertise, to help with the balance of power, and to bring in additional independence or diversity of thought. Committees, such as the audit committee, investment committee, technology or ICT committee, social and ethics committee, and compensation committee, enable boards to handle this complexity through subgroup-focused responsibility and expertise while limiting the demands placed on individual directors.

The study observed that some participants were of the view that technology and emerging technologies such as AI should have a standalone board committee. In contrast, others suggest a multi-disciplinary approach as AI technology now affects every facet of the organisation, i.e., finance, ethics, risks, operations, etc. A single board committee could be indicative of how a board has identified the need for and prioritised technology within the

corporation. From other respondents' perspective, not all corporations have the capacity or need of a standalone committee.

Some participants indicated that technology and emerging technological issues such as AI were at the purview of the audit committee in their corporations. This research has observed that the current composition of audit committees in South African firms is mainly people with financial backgrounds, i.e., Chartered Accountants, not technologists. Thus, the study suggests a need for a dedicated AI and technology board committee supplemented by other board committees (multi-disciplinary approach) such as audit, risk and social and ethics sub-committees. Boards should contemplate the formation of a specialised board committee focused on AI and technology governance, considering factors such as the corporation's scale, the intricacy of its technological landscape, and the extent to which its business and operations rely on technology.

This approach allows for more effective committees and encourages the board to be all-encompassing and effective in their committee creation. The effective collaboration through cross-membership can also be cascaded down to management-level sub-committees that support the board committees, i.e., ethics steering committees, AI steering committees, risk committees, audit steering committees, etc. At both the board levels, the cross-membership approach can be a key driver in introducing an AI Centre of Excellence (CoE), wherein AI technologies projects and governance are not only the responsibility of IT units but of everyone in the corporation. The study views the social and ethics committee as viable to fulfil the role of an AI Centre of Excellence (CoE) at the board level. Thus, the study proposes an additional role for the social and ethics committee to oversee issues of ethics, data safety and robustness, fairness, accountability, and transparency. It would oversee the smooth adoption of AI technology overall.

2. AI Governance Framework

The findings reveal that most companies the participants represent do not have a responsible AI governance framework. However, most of these participants who responded "no" indicated a need for a responsible use of the AI governance framework in their organisation. The framework sets out a good mechanism to outline tangible control around AI policies; it sets a starting point and the baseline for people to refer to in terms of how the technology should be adopted, executed, and set considered roles and responsibilities. The

framework prescribes how the board discharges responsibility and delegates it to the responsible board committees. The AI governance framework is a picture that sets out how the board delegates authority for artificial intelligence and then holds accountable those responsible, i.e., management or board subcommittees. AI governance framework serves as a mechanism for risk management, promotes transparency and accountability, facilitates compliance with legal and regulatory requirements, and is a comprehensive tool for navigating ethical and societal implications. An effective AI governance framework must encompass key principles such as transparency, explainability, accountability, fairness, bias mitigation, safety, and security.

The AI governance framework provides oversight and supervision at the board level while entrusting executive management with core technical aspects, such as data engineering, building models, and choosing best-fit AI platforms. AI governance frameworks assist corporations in governing, monitoring, and providing oversight over design, development, and deployment of AI technologies. The framework also ensures AI technologies' alignment with the corporate strategic objective and ethical culture. An AI framework should instil accountability, responsibility, and oversight throughout the AI development, deployment, and use processes. Thus, the board should also establish clear policies and guidelines within the framework, ensuring that ethical considerations are embedded in decision-making processes related to AI technologies. Regular monitoring and assessment of AI technologies, along with addressing any ethical concerns that may arise, are crucial aspects of the board's responsibility in implementing and sustaining an effective AI governance framework.

3. AI value delivery and performance

Findings reveal that AI performance and value delivery should be part of the AI strategy that ensures alignment of AI investment with business objectives and demonstrates the value of AI technologies to stakeholders in the corporation. The board must set out key performance indicators (KPIs) for performance and value delivery. KPIs are measurable values that indicate how the corporation is achieving its goals and objectives. KPIs can be used as indicators that reflect on the contribution of AI technologies to the corporation's strategic objectives, such as growing revenue, increasing performance, improving customer satisfaction, increasing market share, reducing cost, enabling new products/or services, etc. This means that KPIs can measure AI technologies' success by demonstrating a concrete

return on investment (ROI), expressed as time, money, or labour. The KPIs indicators should be aligned with the strategic objectives of the corporation.

The study suggested a balanced scorecard as another indicator to measure AI technologies' performance and value delivery. BSC can facilitate the link between AI activities and outputs to the business outcomes and value metrics and provide a holistic view of AI performance. The BSC can help monitor and communicate the performance of AI technologies across four perspectives: financial, customer, internal, and learning and growth. Each of these four perspectives should be transformed into matching metrics and measures to assess the current environment (Grembergen, 2001). The assessments need to be continuously conducted and aligned with the pre-determined objectives and benchmarks with best practices. A well-developed scorecard contains a good mix of outcomes, measures, and performance drivers.

8.5.1. Theme 4: Recommendations

8.5.1.1.   Technology Subcommittee

The findings reveal that even though most boards deal with technology at the board level, the delegated subcommittee, in some cases, does not have in-depth knowledge and experience of AI technologies or at least has an independent expert advising at the committee and board level. Mostly, the audit committee is tasked with the responsibility and by design, most companies still prefer members with a financial background to serve on the audit committee.

Thus, the study recommends a dedicated AI and technology board committee supplemented by other board committees (multi-disciplinary approach) such as audit, risk, and social and ethics sub-committees. The specialised board committee should focus on AI and technology governance, potential risks, and challenges of new technological tools, considering factors such as the corporation's scale, the intricacy of its technological landscape, and the extent to which its business and operations rely on technology. The committee is crucial for ensuring that current and future business plans effectively unfold without concern over cyber threats, investment feasibility, or regulatory problems such as discrimination.

The establishment of a dedicated committee will add technical expertise at the board level and ensure that the board is to advise appropriately on technology matters. The committee should be composed of a majority of non-executive directors who are skilled and experienced in understanding the impact of technology in business and society to bridge the knowledge and communication gaps between the executive management and the board of directors. The volatile nature of technology has transformed how corporations do business. To remain viable in a competitive global economy, the boards must improve their technological knowledge and skills to understand and strategically respond to market demands.

8.5.1.2. Social and Ethics Committee as the new AI Center of Excellence

To the extent that AI technologies positively and negatively impact the stakeholders, society, and the environment, i.e., unemployment, AI bias, the risk of abdicating ethical responsibilities, privacy, transparency, accountability, and a number of ethical dilemmas. Firstly, the study recommends that the social and ethics committee oversees the overall development, deployment and adoption of AI technology, particularly to oversee issues of ethics, data safety and robustness, fairness, accountability and transparency. Additionally, the role of the committee would be to provide oversight on responsible use of the AI framework, risk assessment, impact assessment and evaluation of the transparency of the AI's decision-making process. The committee must recommend to the board whether the AI implementation may have a beneficial or potential negative impact on the environment, health, public safety, or society in general. The ethics committee must examine the risk of bias becoming embodied and amplified in the AI technology at issue and must propose measures to prevent and mitigate bias. As per the Companies Act requirement (Section 72(8)(e)), the committee should report annually to the shareholders at the company's annual general meeting.

Secondly, the study recommends the social and ethics committee as a viable committee to fulfil the role of an AI Centre of Excellence (CoE) at the board level. Under the social and ethics committee role, the AI Center of Excellence (CoE) will oversee AI technologies' deployment, development, and overall adoption. The role of the social and ethics committee is to deal with matters relating to good corporate citizenship, the promotion of equality, the prevention of unfair discrimination, environment, health, and public safety issues, as well as

consumer relationships, all of which can be impacted by the deployment and development of AI technologies.

To fulfil its mandate, the CoE should have the power to critically examine, approve or deny any development, deployment and adoption of AI technologies in terms of legal compliance, potential risks, and opportunities. The CoE should have the following role and responsibilities.

- To create a vision for AI and AI ethics and identify business-driven use cases, both of which would define the data platform and tools needed to deliver products and services with the most beneficial social impact and the least harmful ethical impact.
- To develop and maintain a network of AI and AI ethics champions in the corporation.
- To identify ethical concerns and establish ethical standards related to any risk associated with relevant AI systems through multi-stakeholder engagement and cross-disciplinary impact assessment.
- To focus on establishing a fair, accountable, and transparent practice to address AI concerns from stakeholders.
- To ensure the building of talent and applicable upskilling on the tech side and capacity building on the social (ethical) side.
- To ensure understanding of the ethical implications of the company or department's processes and products and translation of those into meaningful insights.
- To affect and maintain communication and implementation of best practices across the company.

Lastly, the study recommends that the board also establishes a mirror CoE at the executive management level to support the social and ethics committee. The CoE at the management level should be composed of all executives from different departments in the decision-making process to ensure a comprehensive and well-aligned approach to AI implementation. The role of this CoE is to steer the corporation's AI strategy, ensuring its alignment with business objectives, setting the strategic vision for AI initiatives, making informed decisions on resource allocation, and fostering a culture of innovation and continuous improvement. Additionally, it will guide the integration of AI into various business functions, facilitate collaboration between technical and non-technical teams, and oversee the ethical and responsible use of AI technologies. On a quarterly basis, the CoE at the management level will report to the social and ethics committee.

8.5.1.3. Cross-membership collaboration between committees (multi-disciplinary approach) in AI governance.

P5 has argued that in their case, AI technologies are a topic of three board sub-committees: the operations and IT risk committee, the social and ethics committee, and the risk capital, compliance and conduct sub-committee, which is a cross-membership. The study is of the view that the concept of cross-membership in board committees can significantly contribute to a multi-disciplinary approach to AI governance. King IV (2018) encourages a holistic and integrated approach when the board establishes board committees for effective collaboration and ensures complementary work is produced. AI governance involves addressing a range of issues, including ethical considerations, legal compliance, technological challenges, and societal assessments and impacts. Thus, having board members with diverse expertise serving on multiple committees can foster a holistic and integrated approach to AI governance.

The study recommends a cross-membership to support a multi-disciplinary approach. In this context, the board members with cross-membership in social and ethics and technology (IT) committees can ensure that ethical considerations are integrated into technological decision-making processes, helping to align AI development, deployment and adoption with strategic business objectives and societal values. The cross-membership will allow for a comprehensive understanding of the potential risks associated with AI implementations, enabling the development of effective risk mitigation strategies by other board committees. Lastly, the combined technical experts with individuals from strategic planning or business development committees can provide a more strategic and business-oriented perspective on AI development and deployment.

Breaking down silos and encouraging cross-membership can allow corporations to leverage the diverse expertise of board members to create a more integrated and effective approach to AI governance. This approach is particularly important given the inter- and multi-disciplinary nature of AI and the need to navigate the complex ethical, legal, technical, and societal dimensions associated with its development, deployment, and adoption. Effective collaboration through cross-membership should also be cascaded down to management-level sub-committees that support the board committees, i.e., ethics steering committees, AI steering committees, data committees (necessary to ensure integrity of data, fairness,

and security, robustness and accuracy, thus safety), risk committees, audit steering committees, etc.

8.5.1.4.   AI governance framework

The finding revealed that only one company had interaction with an AI governance framework. To address this gap, the study recommends that all JSE-listed companies be required as part of the listing requirement to responsibly use an AI governance framework. The AI governance framework should have the following key principles at the basic minimum.
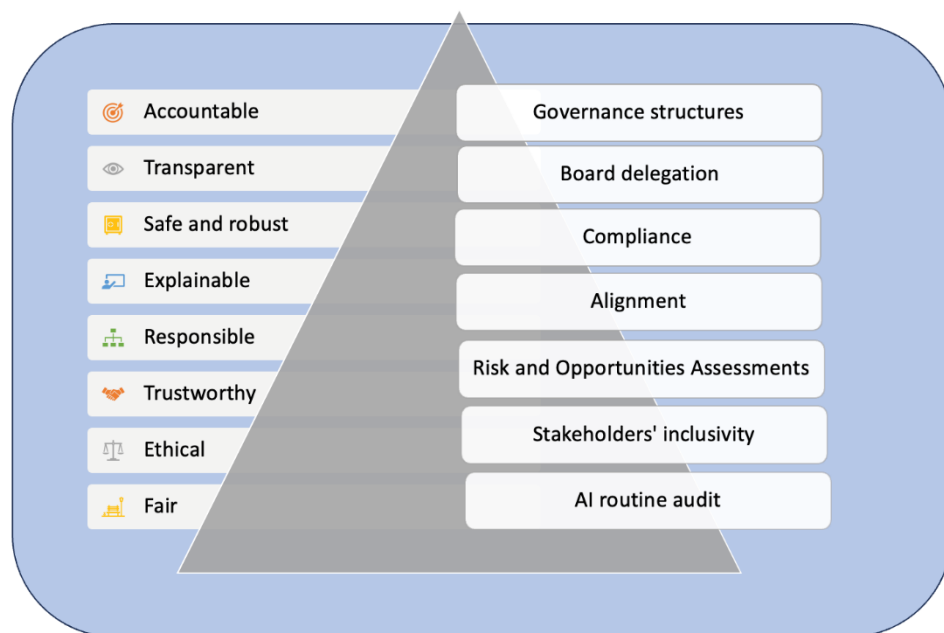

Figure 16: Figure 16: AI Governance Framework, Own Deductions

1. **Ethical Considerations:** AI governance frameworks should incorporate ethical principles, including fairness, transparency, accountability, safety and robustness, as well as respect for privacy. They should guide the board and executive management in setting the ethical tone of the organisation and making ethical decisions.
2. **Governance structures:** The board ensures that governance structures are established at the board and management level to govern, monitor, and provide oversight over the design, development, and deployment of AI technologies against the framework of ethical principles.

3. **Compliance:** The board ensures that the design, development, and deployment of AI technologies adhere to relevant laws and regulations, such as the POPI act, the constitution on non-discrimination, etc.

4. **Alignment:** The board ensures that AI technology initiatives are aligned with the corporation's values, mission, and long-term goals.

5. **Risk and Opportunities Assessments**: The board ensures continuous risk assessments and mitigation strategies for potential negative impacts. The board should also leverage opportunities resulting in AI investments.

6. **Board delegation**: The Board delegates authority for artificial intelligence and then holds accountable those responsible, i.e., management or board subcommittees.

7. **Human on the loop:** The board maintains human oversight and control over AI technologies to prevent undue automation and decision-making processes that could have significant impacts on individuals or society. Human intervention during both development and deployment (and use) should include the possibility of overriding a decision made by AI technologies when it violates the law (Kumar et al.,2020) to ensure greater accountability of AI technologies.

8. **Stakeholders' inclusivity:** The board must ensure a stakeholder-inclusivity approach so that everyone impacted has a voice and is adequately represented. Further, it should ensure that all stakeholders understand the importance of data ethics and AI governance and their role in upholding these principles.

9. **AI routine audit:** The board should ensure processes and policies for AI technologies routine audit by both internal and external auditors to ensure that datasets or algorithms are free from any form of biases which may impact society.

Implementing AI governance frameworks would require leadership at the board and management level to ensure coordinated efforts across the corporation to ensure that the AI framework is put into practice in order to be effective. Responsible use of the AI governance framework provides a structured approach to ensure that AI technologies are deployed and utilised in an ethical, accountable, and transparent manner. Thus, the board would be responsible for implementation, oversight and monitoring of AI governance framework.

8.6. CONCLUSION

The recommendations in this study which included: AI training at the board level; having a tech expert on the board; AI risk governance at the board level; tone from the top – practical steps for setting ethical tone; board stakeholder engagement; a cradle-to-grave approach for disposal of AI powered hardware; data wiping; protection of the environment; establishment of a technology governance committee; the social and ethics committee as an AI Center of Excellence; and cross-membership collaboration between committees (multi-disciplinary approach) in AI governance, are grounded in the stakeholder theory approach informed by a relational ethics framework. The proposed governance framework advocates that the board must ensure a stakeholder-inclusive approach so that everyone impacted has a voice and is adequately represented. Further, it should ensure that all stakeholders understand the importance of data ethics and AI governance and their role in upholding these principles for the sake of each participant in the engagement of AI technology. A relational ethical framework driving a stakeholder theory approach ensures that the collective ownership of AI technology by all stakeholders can be actualised into collective responsibility for the welfare of everyone involved. There is a lot of scope for future elaboration of the strength that relational ethics brings to a stakeholder approach to corporate governance, specifically in the context of AI governance, given that AI technology impacts each member of each community as well as the rights of communities as a whole in terms of for instance data representation.

CHAPTER 9: CONCLUSIONS

9.1. INTRODUCTION

The chapter will highlight the synopsis of all the chapters, provide an overview of the research recommendations, novelty of the study, and recommendations for future studies, and set out the study's final conclusion.

9.2. OVERVIEW OF CHAPTERS

Chapter 1: Introduction

In Chapter One (1), the researcher gave an overview of business reliance on artificial intelligence (AI) technologies for efficient operations and support functions to achieve competitive advantage, its achievements in different industries, ethical concerns raised, and the spotlight on the board's role in governing AI ethics. The chapter further outlines the research problem, which examines the boards' lack of AI awareness and overseeing and monitoring roles in terms of AI technologies. More precisely, this research aimed to investigate how boards perform their functions in the governance of AI technologies through a case study of JSE-listed companies. The chapter further highlights the qualitative research design using phenomenological methodology to understand the lived experiences of board members in governing AI ethics. Furthermore, this chapter provides an overview of the significance of the study due to the rapid increase in the significance of AI ethics and governance due to cases of algorithm bias, privacy, and safety concerns.

Chapter 2: Artificial Intelligence

In chapter two (2), the researcher explored different definitions from the early figures of AI, such as Turing, John McCarthy, Marvin Minsky, and Simon Newell, who laid the foundation to recent scholars. The chapter detailed the history of AI from the 1950s, transitioned from theory to practice: the rise of AI in the 60s, the first AI winter in the 70s, the age of renewed optimism in the 80s, the expansion of AI applications such as deep learning architectures in the 2000s, and explosive growth to date which include the wide adoption of deep learning techniques and generative pre-trained transformer models. The chapter further clarifies the key concepts used interchangeably, such as machine learning, deep learning, and generative AI, which are layers of data-driven AI technologies. Furthermore, the chapter highlighted the importance of AI in health, education, business, environment, and

agriculture, while acknowledging risks and ethical concerns. Finally, the chapter discusses global spending on AI and the current status of South Africa's incorporation of AI technologies into business operations and product development.

Chapter 3: Corporate Governance and Artificial Intelligence

In this chapter, the researcher reviewed the definition of corporate governance from Cadbury Report (1992), Shleifer and Vishny (1997), Zingales (1998), Denis and McConnell (2003), OECD (2004), Solomon (2007), L'Huillier (2014), amongst others. In the end, the study aligned with the King IV report on corporate governance, which elevated the definition by focusing its version not only on "directing and controlling" the company but also highlighting the principles of ethical leadership, transparency, accountability, fairness, and responsibility to achieve a strong and viable ethical culture, good performance, value creation, legitimacy, and effective control of the organisation which is more relevant to the study. This chapter also discusses the theoretical framework of corporate governance, which is located within agency theory. The chapter also discusses the different board models, that is, the one- and two-tier board models and the control, strategic, stakeholder, and advisory and fiduciary duties of the board. This chapter further considers AI technologies as board members and argues that AI technologies can contribute significantly to data-driven decision-making. Core fiduciary duties, such as loyalty and care, and a human understanding of ethical considerations, especially the ethics of AI, is required for good AI governance in a corporation. The chapter concludes with a conceptual framework that focuses on the role of the board as a focal point of CG, meaning that the board is the overarching structure responsible for governance and oversight through a mix of processes, structures, and rational mechanisms of the organisation's AI ethics approach.

Chapter 4: Artificial Intelligence Ethics

This chapter focuses on a broad discussion of ethics (metaethics, normative ethics, and applied ethics) and concludes, based on a literature review, that AI ethics is a sub-discipline of the field of applied ethics, which is concerned with two things: the moral behaviour of humans that create AI technologies and the impact of AI technologies on society, which are both related to the governance of these technologies. The chapter further highlights the most common AI ethics principles, namely accountability and responsibility, fairness, transparency, explainability, and privacy, and agrees that principles can be used to guide the development of governance practices, international standards, and further regulation. The chapter further discusses organisational ethics and the alignment of organisational

cultural norms with AI technologies and AI ethics. The chapter concludes by arguing that the board of directors sets the tone for any corporate ethics and compliance programme. Therefore, it is the responsibility of the board and executive management to ensure that technological advancements in AI align with the organisation's ethical culture.

Chapter 5: Artificial Intelligence Governance

In this chapter, the researcher discussed the relationship between AI, corporate, and IT governance. By definition, AI governance closes the gap between accountability and ethics in technological advancement, and the chapter highlights the general consensus that AI governance concerns AI technologies that are explainable, transparent, free from bias, and ethical, as many scholars allude to their definitions. This chapter further highlights the importance of AI governance in managing the risks posed by AI technologies. The chapter highlights the different IT frameworks and AI governance as part of the organisational corporate governance structure, including AI governance frameworks that boards can utilise to govern AI technologies. The researcher concluded that boards should ensure that there are effective management-level AI compliance and reporting structures to facilitate board oversight, which includes continuous AI technology risk assessments and monitoring of high-risk AI systems and written AI policies, procedures, and training.

Chapter 6: Research Approach and Methodology

The chapter presents the methodology and design adopted for the study to help the researcher achieve its envisioned objectives. The methodology is one of the essential chapters of the research process as it guides the reader on the coherent steps adopted by the researcher while embarking on the study. This study adopts a qualitative methodology. To gain direct knowledge of the experience, the phenomenological approach was followed to analyse and reflect on the lived experiences and perceptions of board members and specialists of the role of the board in the governance of AI technologies in an organisation. This chapter further provides a detailed description of the research approach, tools and applications, primary and secondary data collection, and data analysis process methods selected for this study. The sampling procedures, validity and reliability, limitations, feasibility, and positionality are also examined in detail. Ethical considerations are discussed to demonstrate that the rights and welfare of the participants were protected and that the research was conducted responsibly and transparently.

Chapter 7: Data Collection and Analysis

In this chapter, the researcher presents data collected from semi-structured interviews. These interviews, which were conducted from twenty-four research participants (board members and specialist participants), yielded insights that are pivotal to the research. As a result, the collected data were infused with qualitative depth from the participants' lived experiences and personal narratives as board members. This chapter further highlights the analysis of data collected from the participant interviews. After familiarising himself with participant data, the researcher developed significant statements that were then grouped into themes. The findings and analysis from the semi-structured interviews are presented in sub-theme sections 1 to 4. The insights gleaned from these interviews were instrumental in furthering the researcher's understanding of the role of the board in AI ethics enabling effective governance of AI technologies, interpreting each facet of the research sub-questions, and setting the stage for the discussions and recommendations.

Chapter 8: Summary and recommendations

This chapter provides a comprehensive summary of the thematic analysis and offers key recommendations, which include AI training at the board level; having a tech expert on the board; AI risk governance at the board level; tone from the top – practical steps for setting ethical tone; board stakeholder engagement; a cradle-to-grave approach for disposal of AI-powered hardware; data wiping; protection of the environment; establishment of a technology governance committee; the social and ethics committee as an AI Centre of Excellence; cross-membership collaboration between committees (multi-disciplinary approach) in AI governance; and being guided by an AI framework. The chapter further recommend that JSE-listed companies be required as part of the listing requirement to responsibly use an AI governance framework.

9.3. OVERALL FINDINGS AND RECOMMENDATIONS

This research investigated how boards perform their functions in the ethics and governance of artificial intelligence through a case study of JSE-listed companies and a focused literature review. The purpose of the study was to explore the board's role as a focal point of corporate governance, leading and setting an ethical tone relating to the ethically responsible governance of AI technologies and ensuring that emerging AI technologies support the setting of strategic objectives and goals that are ethically acceptable.

The following figure 18 summarises the recommendations of this study. These recommendations are based on the role of the board in establishing governance structures, processes, and rational mechanisms to govern artificial intelligence ethics in a private organisation (see Chapter 8 for detailed recommendations of the study).
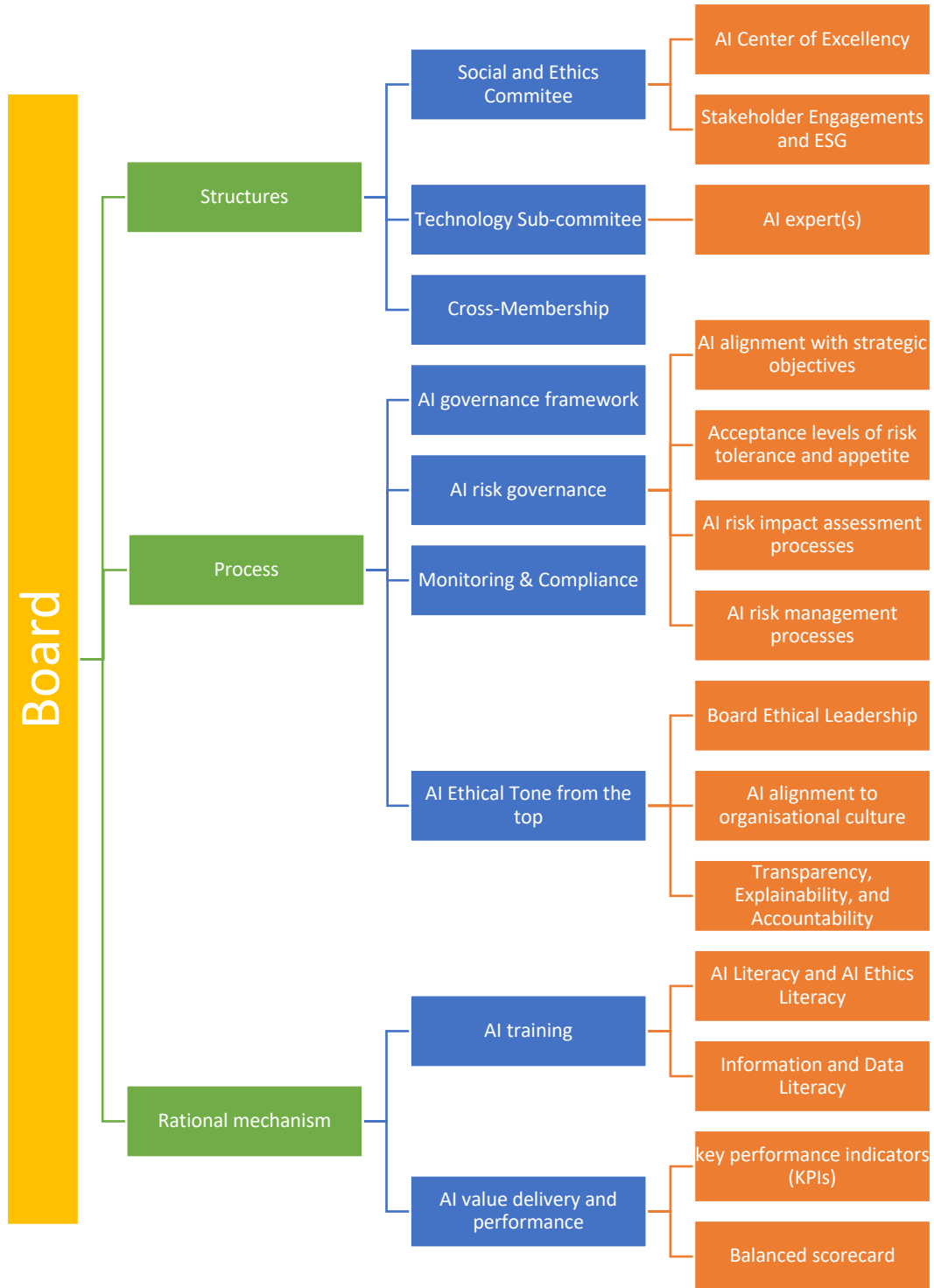

Figure 17: summarises the recommendations of this study

9.4. THE NOVEL CONTRIBUTION OF THE STUDY

There is a distinct gap in the literature in terms of a multidisciplinary understanding of the business world's role in mitigating the ethical challenges faced by society as a result of the research, design, development, deployment, and use of AI systems and applications (Floridi, 2015). This study aimed not only to address this gap in general, but also to highlight the much-neglected role of the private sector in AI ethics, especially at the leadership or board level (Kambil and Lucas, 2002), by bringing together different disciplines such as corporate governance, applied ethics, and AI governance in a novel manner, given the study's primary situatedness in AI ethics, rather than exclusively in corporate governance.

As the first scholarly study of its kind, this study provides answers on the role of the board in governing artificial intelligence ethics. The study unpacks the inter-disciplinary complexities determining the role of boards in terms of AI ethics and governance and provides recommendations to ultimately improve governance in private sector corporations.

The study has significant real-world implications for boards, as they govern corporations which design, develop, deploy, and adopt AI technologies. As boards steer their corporations in the future, they must understand the ethical implications of AI technologies and learn about mechanisms that can be utilised to govern them effectively. This study assists boards by providing practical recommendations for establishing governance structures, processes, and rational mechanisms that can be used to enhance the effectiveness of the AI ethics governance process.

9.5. LIMITATION

The methodology involved conducting interviews exclusively with board members and specialists from these listed companies, which presents a limitation on perspective diversity. While this study adopted a stakeholder-inclusive approach, this approach informed the *recommendations* of the study, rather than the choice of participants in the study's survey. Interviews were conducted with board members rather than broader stakeholders of JSE-listed companies, as the aim was to achieve an understanding of the state of the art of knowledge and engagement with AI governance and ethical concerns among *board members*. Recommendations were driven by stakeholder theory given the arguments in

chapter 3 for a stakeholder approach to corporate governance, and informed by the information on the extent of familiarity of board members with AI ethics principles and concerns. Secondly, including broader stakeholders would have been too resource-intensive. While a broader stakeholder engagement might have enriched the study, practical limitations necessitated a more focused approach. Future research could expand on this work by incorporating a more comprehensive stakeholder perspective.

However, the study encompassed interviews with 24 board members and specialists, who collectively represented 58 companies through both current and past board memberships. It is important to note that it is a common practice for JSE-listed company board members to hold multiple board positions simultaneously. While this provides a breadth of insight, it also introduces a limitation regarding the representation. The JSE comprises approximately 300 to 400 listed companies, so the study's findings might not encapsulate the full spectrum of governance practices across all organisations.

Therefore, research findings must be interpreted in light of these limitations. These limitations represent avenues for future research.

Despite these constraints, the findings and recommendations of this study have theoretical generalisability, which is enhanced by the verification of findings that the literature review provides. They can be applied to organizational settings beyond those specifically investigated, contributing valuable insights into the governance of AI ethics across various business contexts.

9.6. FUTURE RESEARCH

9.6.1. Role of Audit Committee in AI technologies

The study observed that most companies have assigned the oversight responsibility of AI technologies to their audit committees. However, this study also found that most audit committees lack technical expertise in this area. Assigning oversight responsibility to audit committees is consistent with its overarching role in risk oversight. It is also worth noting that it is common for agenda items taken on by the audit committee at the outset to eventually be overseen by other committees. Some aspects of AI technology oversight seem to align more with committees other than audit committees. For example, the impact of generative AI on corporate culture may be aligned with the oversight of human resources, social and

ethics committees, or the full board. Thus, the study recommends future studies should examine (i) the roles and interactions between Audit Committees and other committees in the realm of AI technologies, (ii) assessment of the current effectiveness of Audit Committees in overseeing AI-related risks, (iii) research on enhancing corporate governance in the Age of AI with a specific focus on Audit Committees and (iv) research on including data officer in the audit committee.

9.6.2. Board Composition

The study acknowledges that the recommendations in chapter 8 would necessitate a shift on the "traditional" board composition in South Africa, which is "still considered to be an 'old boys club' whose membership favours white males and a new elite group of black males" (Nyirenda, 2010). There is no doubt that "AI technology is changing the world in which future generations will live, and therefore responsible AI practices should surely demonstrate interaction with issues of intergenerational justice' (Ruttkamp-Bloem, 2023). Therefore, future board composition in the age of AI should consider an intergenerational mix.

This study recommends further studies on future board composition in South Africa at the age of AI.

9.6.3. Board governance and oversight

To navigate the ever-evolving landscape of AI technologies and its impact on corporate governance, the researcher recommends future studies should be conducted on the efficacy of existing board structures in addressing AI-technology-related challenges and recommend adjustments to ensure boards and board committees are equipped to provide effective oversight. Further studies should examine successful AI integration cases, and multidisciplinary collaboration with industry and academic experts will contribute valuable insights, enabling boards to proactively guide corporations toward responsible and innovative AI adoption.

9.7. CONCLUSION

Corporations around the world recognise that the development and deployment and adoption of AI technologies in private sector companies can lead to a competitive advantage. AI technologies can improve organisational productivity (financial, marketing,

and administrative) and streamline organisational processes for effective decision making and tasks and can bring significant benefits not only to their internal operations but, more importantly, to their bottom lines. However, to ensure that these investments provide the expected returns, effective governance mechanisms need to be implemented. This study provides boards with a set of AI governance recommendations that they can customise for specific corporations. These recommendations will allow boards to discharge their governance responsibilities more effectively regarding AI technologies and ensure the sustainability of the AI technology adopted in the company.

REFERENCES:

1.   Aguilera, R. V. (2005). Corporate governance and director accountability: An institutional comparative perspective. British Journal of Management, 16, S39-S53.

2.   Anand, V., Manz, C. C., & Glick, W. H. (1998). An organizational memory approach to information management. Academy of management review, 23(4), 796-809.

3.   Asaro, P. M. (2020). What should we want from a robot ethic?. In Machine Ethics and Robot Ethics (pp. 87-94). Routledge.

4.   Aydın, Ö., & Karaarslan, E. (2023). Is ChatGPT leading generative AI? What is beyond expectations?. What is beyond expectations.

5.   Bagnoli, C. (Ed.). (2013). Constructivism in ethics. Cambridge University Press.

6.   Bailey, J. (2008). First steps in qualitative data analysis: transcribing. Family practice, 25(2), 127-131.

7.   Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2021). Privacy issues of AI. In An introduction to ethics in robotics and AI (pp. 61-70). Springer, Cham

8.   Battersby, M. (1989). Critical thinking as applied epistemology: Relocating critical thinking in the philosophical landscape. Informal Logic, 11(2).

9.   Baxt, R. (2005). Duties and responsibilities of directors and officers. AICD.

10.   Berkowitz, M. (ed.) (2020). 'AI, Machine Learning and Big Data' (2nd edition). The Global Legal   Group. https://www.globallegalinsights.com/practice-areas/ai-machine-learning-and-big-data-laws-and-regulations/south-africa

11.   Bogdan, R., & Biklen, S. K. (1997). Qualitative research for education. Boston, MA, USA:: Allyn & Bacon.

12.   Borenstein, J., & Howard, A. (2021). Emerging challenges in AI and the need for AI ethics education. AI and Ethics, 1(1), 61-65.

13.   Brand, V. (2020). Artificial Intelligence and Corporate Boards-Some Ethical Implications. Vivienne Brand, 'Artificial Intelligence and Corporate Boards: Some Ethical Implications', in Andrew Godwin, Pey Woan Lee and Rosemary Langford (eds) Innovation, Technology and Corporate Law (Edward Elgar Publishing, 2021, Forthcoming).

14.   Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative research in psychology, 3(2), 77-101.

15.   Broccolo, B. M., & Peregrine, M. W. (2021). Fiduciary Engagement in Artificial Intelligence Innovation: A Governance Imperative.

16.   Bruyn, S. T. H. (1966). The human perspective in sociology: The methodology of participant observation.

17.   Bughin, J., Seong, J., Manyika, J., Chui, M., & Joshi, R. (2018). Notes from the AI frontier: Modeling the impact of AI on the world economy. McKinsey Global Institute.

18. Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In Conference on fairness, accountability and transparency (pp. 77-91). PMLR.

19. Butcher, J., & Beridze, I. (2019). What is the State of Artificial Intelligence Governance Globally?. The RUSI Journal, 164(5-6), 88-96.

20. Butvinik, D. (2022). Foundations of Ethical Artificial Intelligence: Concepts and Principles of Explainability and Trust. Medium. https://medium.com/analytics-vidhya/foundations-of-ethical-artificial-concepts-and-principles-of-explainability-and-trust-ccfec94765f2

21. Canca, C. (2020). Operationalizing AI ethics principles. Communications of the ACM, 63(12), 18-21.

22. Carpenter, C. (2021). New Complexities in Governance: How Board of Directors are Managing the Rise of Artificial Intelligence. Women Get On Board. Retrieved January 31, 2022, from https://womengetonboard.ca/new-complexities-in-governance-by-denise-carpenter/

23. Cheong, B. C. (2024). Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making. Frontiers in Human Dynamics, 6, 1421273.

24. Chiao, V. (2019). Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice. International Journal of Law in Context, 15(2), 126-139.

25. Clarke, R. (2009). Privacy impact assessment: Its origins and development. Computer law & security review, 25(2), 123-135.

26. Colaizzi, P. F. (1978). Psychological research as the phenomenologist views it.

27. Cope, D. G. (1969). Methods and meanings: Credibility and trustworthiness of qualitative research. Number 1/January 2014, 41(1), 89-91.

28. Coppin, B. (2004). Artificial intelligence illuminated. Jones & Bartlett Learning.

29. Cossin, D., & Metayer, E. (2015). How strategic is your board?. MIT Sloan Management Review Sloanselect Collection, 45-51.

30. Crane, A., Matten, D., Glozer, S., & Spence, L. J. (2019). Business ethics: Managing corporate citizenship and sustainability in the age of globalization. Oxford University Press, USA.

31. Crawford, K. (2017), The Trouble with Bias - NIPS 2017 Keynote, https://www.youtube.com/watch?v=fMym_BKWQzk.

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

33. Cridland, E. K., Jones, S. C., Caputi, P., & Magee, C. A. (2014). Being a girl in a boys' world: Investigating the experiences of girls with autism spectrum disorders during adolescence. Journal of autism and developmental disorders, 44, 1261-1274.

34.   Cubric, M. (2020). Drivers, barriers and social considerations for AI adoption in business and management: A tertiary study. Technology in Society, 62, 101257.

35.   Dafoe, A. (2018). AI governance: a research agenda. Governance of AI Program, Future of Humanity Institute, University of Oxford: Oxford, UK.

36.   Davenport, T.H. & S. Dasgupta. (2019). How to set up an AI Centre of Excellence. Harvard Business Review. https://hbr.org/2019/01/how-to-set-up-an-ai-center-of-excellence

37.   De Cremer, D., & Kasparov, G. (2022). The ethical AI—paradox: why better technology needs more and not less human responsibility. AI and Ethics, 2(1), 1-4.

38.   de Laat, P. B. (2021). Companies Committed to Responsible AI: From Principles towards Implementation and Regulation?. Philosophy & technology, 34(4), 1135-1193.

39.   Debreceny, R.,S. (2006) "Re-engineering IT Internal Controls: Applying Capability Maturity Models to the Evaluation of IT Controls", Proceedings of the 39th Hawaii International Conference on System Sciences, Hawaii.

40.   Dilmegani, C. (2022). Top 9 Ethical Dilemmas of AI and How to Navigate Them in 2022. AIMultiple. Retrieved March 29, 2022, from https://research.aimultiple.com/ai-ethics/

41.   du Plessis, J. J., Großfeld, B., Saenger, I., & Sandrock, O. (2017). An overview of German business or enterprise law and the one-tier and two-tier board systems contrasted. German corporate governance in International and European context, 1-16.

42.   Eccles, R. G., & Vogel, M. (2022). Board Responsibility for Artificial Intelligence Oversight. The Harvard Law School Forum on Corporate Governance. https://corpgov.law.harvard.edu/2022/01/05/board-responsibility-for-artificial-intelligence-oversight/

43.   Eisenhardt, K. M. (1989). Agency theory: An assessment and review. Academy of management review, 14(1), 57-74.

44.   Ertel, W. (2018). Introduction to artificial intelligence. Springer.

45.   Falco, G., Shneiderman, B., Badger, J., Carrier, R., Dahbura, A., Danks, D., ... & Yeong, Z. K. (2021). Governing AI safety through independent audits. Nature Machine Intelligence, 3(7), 566-571.

46.   Ferrell, O. C., Crittenden, V. L., Ferrell, L., & Crittenden, W. F. (2013). Theoretical development in ethical marketing decision making. AMS review, 3(2), 51-60.

47.   Ferrell, O. C., Harrison, D. E., Ferrell, L., & Hair, J. F. (2019). Business ethics, corporate social responsibility, and brand attitudes: An exploratory study. Journal of Business Research, 95, 491-501.

48.   Fischer, G. (2017). Exploring design trade-offs for quality of life in human-centered design. Interactions, 25(1), 26-33.

49.   Fixler, A., Oliaro, B., Frieden, M., Girardo, C., Winterbottom, F. A., Fort, L. B., & Hill, J. (2023). Alert to action: implementing artificial intelligence–driven clinical decision support tools for Sepsis. Ochsner Journal, 23(3), 222-231.

50.   Free, R. (n.d.). What is an AI Ethics Board? https://cms.law/en/rou/publication/what-is-an-ai-ethics-board. Retrieved September 27, 2022, from https://cms.law

51.   Freeman, R. E., & Reed, D. L. (1983). Stockholders and stakeholders: A new perspective on corporate governance. California management review, 25(3), 88-106.

52.   García-Peñalvo, F., & Vázquez-Ingelmo, A. (2023). What do we mean by GenAI? A systematic mapping of the evolution, trends, and techniques involved in Generative AI.

53.   Gavrilova, Y. (2021). Ethics in Artificial Intelligence: From A to Z. Serokell Software Development Company. Retrieved August 30, 2022, from https://serokell.io/blog/ai-ethics-guide

54.   Gesser, A., Regner, B., & Gressel, A. (2022). AI Oversight Is Becoming a Board Issue. The Harvard Law School Forum on Corporate Governance. Retrieved May 4, 2022, from https://corpgov.law.harvard.edu/2022/04/26/ai-oversight-is-becoming-a-board-issue/?utm_content=buffer731dd&utm_medium=social&utm_source=linkedin.com&utm_campaign=buffer

55.   Gilman, S. C. (2005). Ethics codes and codes of conduct as tools for promoting an ethical and professional public service: Comparative successes and lessons. Prepared for the PREM, the World Bank.

56.   Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. Organizational research methods, 16(1), 15-31.

57.   Giorgi, A. P., & Giorgi, B. (2008). Phenomenological psychology. The SAGE handbook of qualitative research in psychology, 165-178.

58.   Given, L. M. (Ed.). (2008). The Sage encyclopedia of qualitative research methods. Sage publications.

59.   Gordon, C. (2021). Why Board Directors And CEOs Need To Learn AI Knowledge Foundations: Building AI Leadership Brain Trust Is A Business Imperative: Are You Ready? Forbes. https://www.forbes.com/sites/cindygordon/2021/05/10/why-board-directors-and-ceos-need-to-learn-ai-knowledge-foundations-building-ai-leadership-brain-trust-is-a-business-imperative-are-you-ready/?sh=318307b46c05

60.   Gregor, K., & LeCun, Y. (2010, June). Learning fast approximations of sparse coding. In Proceedings of the 27th international conference on international conference on machine learning (pp. 399-406).

61.   Gregory, H. J. (2022). Board Oversight: Key Focus Areas for 2022. The Harvard Law School Forum on Corporate Governance. https://corpgov.law.harvard.edu/2022/01/05/board-oversight-key-focus-areas-for-2022/

62.   Guerra, S. (2021). The Black Box of Governance: Boards of Directors Revealed by Those who Inhabit Them. Routledge.

63.   Gupta, A. (2020). AI ethics groups are repeating one of society's classic mistakes. MIT Technology Review. Retrieved May 24, 2021, from https://www-

technologyreviewcom.cdn.ampproject.org/c/s/www.technologyreview.com/2020/09/14/1008
323/ai-ethics-representation-artificial-intelligence-opinion/amp/

64. Gupta, A., & Heath, V. (2020). AI ethics groups are repeating one of society's classic mistakes. MIT Technol. Rev.

65. Gyekye, K. (2011). African Ethics. Https://Plato.Stanford.Edu. Retrieved June 15, 2021, from https://plato.stanford.edu/archives/fall2011/entries/african-ethics/

66. Hagendorff, T. (2022). AI ethics and its pitfalls: not living up to its own standards?. AI and Ethics, 1-8.

67. Hallamaa, J. (2022). AI Ethics as Applied Ethics. Frontiers. Retrieved May 24, 2022, from https://www.frontiersin.org/articles/10.3389/fcomp.2022.776837/full

68. Hallamaa, J., & Kalliokoski, T. (2022). AI Ethics as Applied Ethics. Frontiers in Computer Science.

69. Hallamaa, J., & Kalliokoski, T. (2022). AI Ethics as Applied Ethics. Frontiers in Computer Science.

70. Häußermann, J. J., & Lütge, C. (2022). Community-in-the-loop: towards pluralistic value creation in AI, or—why AI needs business ethics. AI and Ethics, 2(2), 341-362.

71. Held, V. (2006). The ethics of care: Personal, political, and global. Oxford university press.

72. Hickok, M. (2021). Lessons learned from AI ethics principles for future actions. AI and Ethics, 1(1), 41-47.

73. Hillman, A. J., & Dalziel, T. (2003). Boards of directors and firm performance: Integrating agency and resource dependence perspectives. Academy of Management review, 28(3), 383-396.

74. Hind, M., Tucker, E., & Vaidyanathan, R. (2021, September 13). Foundations of trustworthy AI: How mature is your AI governance? Watson Blog. Retrieved February 15, 2022, from https://www.ibm.com/blogs/watson/2021/06/ai-governance-maturity/

75. Holzinger, A., Goebel, R., Fong, R., Moon, T., Müller, K. R., & Samek, W. (2022). xxAI-Beyond Explainable Artificial Intelligence. In International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers (pp. 3-10). Springer, Cham. Https://Mcdermott-Will-Emery-2793.Docs.Contently.Com/v/Fiduciary-Engagement-in-Artificial-Intelligence-Innovation-a-Governance-Imperative. https://mcdermott-will-emery-2793.docs.contently.com/v/fiduciary-engagement-in-artificial-intelligence-innovation-a-governance-imperative

76. https://searchenterpriseai.techtarget.com/definition/AI-governance
https://www-forbes-com.cdn.ampproject.org/c/s/www.forbes.com/sites/michaelperegrine/2022/04/26/the-looming-board-challenge-oversight-of-artificial-intelligence/amp/
https://www-techtarget-com.cdn.ampproject.org/c/s/www.techtarget.com/whatis/definition/AI-code-of-ethics?amp=1

77.   Hu, J. (2008). Introduction to fabric testing. In Fabric testing(pp. 1-26). Woodhead Publishing.

78.   Hughes, A. (2022). ChatGPT: Everything you need to know about OpenAI's GPT-3 tool. BBC Science Focus Magazine. https://www.sciencefocus.com/future-technology/gpt-3/

79.   Idowu, S. O. (2013). Encyclopedia of corporate social responsibility (Vol. 21). N. Capaldi, L. Zu, & A. D. Gupta (Eds.). New York: Springer.

80.   Ingley, C. B., & Van der Walt, N. T. (2001). The strategic board: The changing role of directors in developing and maintaining corporate capability. Corporate Governance: An International Review, 9(3), 174-185.

81.   Jecker, N. S. (2023). Ubuntu and Bioethics. In Handbook of African Philosophy (pp. 161-180). Cham: Springer International Publishing.

82.   Jecker, N. S., Jecker, N., Jonsen, A., & Pearlman, R. (1997). Introduction to the Methods of Bioethics. Bioethics. An Introduction to the History, Methods, and Practice, 113-125.

83.   Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399.

84.   Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. Science, 349(6245), 255-260.

85.   Jupp, V., & Sapsford, R. (Eds.). (2006). Data collection and analysis. Sage Publications.

86.   Kamalnath, A. (2019). The Perennial Quest for Board Independence: Artificial Intelligence to the Rescue?. Alb. L. Rev., 83, 43.

87.   Kaneko, T., Yoshioka, N., & Sasaki, R. (2020). STAMP S&S: Safety & Security Scenario for Specification and Standard in the society of AI/IoT. In 2020 IEEE 20th International Conference on Software Quality, Reliability and Security Companion (QRS-C) (pp. 168-175). IEEE.

88.   Kerry, C. (2020). Protecting privacy in an AI-driven world. https://www. brookings. edu/research/protecting-privacy-in-an-ai-driven-world/.

89.   Khatry, S., Kwartler, T., Firth-Butterfield, K,. & Caine, M. (2021, September 23). What would it take to make AI 'greener'? World Economic Forum. https://www.weforum.org/agenda/2021/09/make-ai-greener-climate-solution-cop26-technology/

90.   King, P. (1999). 10 Ockham's Ethical Theory. The Cambridge Companion to Ockham, 227.

91.   Kok, J. N., Boers, E. J., Kosters, W. A., Van der Putten, P., & Poel, M. (2009). Artificial intelligence: definition, trends, techniques, and cases. Artificial intelligence, 1, 270-299.

92.   Kompella, K. (2022). How AI ethics is the cornerstone of governance. SearchEnterpriseAI. Retrieved May 4, 2022, from https://www.techtarget.com/searchenterpriseai/tip/How-AI-ethics-is-the-cornerstone-of-governance

93.   Koolen, C., & van Cranenburgh, A. (2017). These are not the stereotypes you are looking for: Bias and fairness in authorial gender attribution. In Proceedings of the First ACL Workshop on Ethics in Natural Language Processing (pp. 12-22).

94. Kopelman, L. M. (1990). What is Applied About "Applied" Philosophy?. The Journal of medicine and philosophy, 15(2), 199-218.

95. Korstjens, I., & Moser, A. (2018). Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. European Journal of General Practice, 24(1), 120-124.

96. Kumar, A., Braud, T., Tarkoma, S., & Hui, P. (2020). Trustworthy AI in the age of pervasive computing and big data. In 2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops) (pp. 1-6). IEEE.

97. Kumar, S., & Choudhury, S. (2022). Normative ethics, human rights, and artificial intelligence. AI and Ethics, 1-10.

98. Kyrou, A. (2015). Robots are People like You and Me: Changing our View in order not to Suffer from Automation. Multitudes, 1(1), 94-104. https://doi.org/10.3917/mult.058.0094

99. Lafferty, W. M., Schmidt, L. A., & Wolfe Jr, D. J. (2011). A brief introduction to the fiduciary duties of directors under Delaware law. Penn St. L. Rev., 116, 837.

100. Lapach, Y. (2022, February 11). AI Governance – What are the KPIs? And who is accountable? 2021.AI. Retrieved March 29, 2022, from https://2021.ai/ai-governance-kpi/

101. Larson, B. N. (2017). Gender as a variable in natural-language processing: Ethical considerations.

102. Larsson, S., & Heintz, F. (2020). Transparency in artificial intelligence. Internet Policy Review, 9(2).

103. Laverty, S. M. (2003). Hermeneutic phenomenology and phenomenology: A comparison of historical and methodological considerations. International journal of qualitative methods, 2(3), 21-35.

104. Lawton, G., & Wigmore, I. (2021). AI ethics (AI code of ethics). TechTarget. Retrieved June 18, 2022, from

105. Lee, K. F. (2018). AI superpowers: China, Silicon Valley, and the new world order. Houghton Mifflin Harcourt.

106. Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. Available at SSRN 3403301.

107. Lewis, S. (2018, November 6). AI governance. SearchEnterpriseAI. (accessed May 2021).

108. Li, P., Yang, J., Islam, M. A., & Ren, S. (2023). Making AI Less" Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models. arXiv preprint arXiv:2304.03271.

109. Liao, S. M. (Ed.). (2020). Ethics of Artificial Intelligence. Oxford University Press.

110. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in neural information processing systems, 30.

111. Madhani, P. M. (2017). Diverse roles of corporate board: Review of various corporate governance theories. The IUP Journal of Corporate Governance, 16(2), 7-28.

112. Malherbe, S., & Segal, N. (2001). Corporate Governance in South Africa. Development Centre Discussion Paper.

113. Malik, N., Tripathi, S. N., Kar, A. K., & Gupta, S. (2021). Impact of artificial intelligence on employees working in industry 4.0 led organizations. International Journal of Manpower, 43(2), 334-354.

114. Manen, M. V. (1997). From meaning to method. Qualitative health research, 7(3), 345-369.

115. Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. AI and Ethics, 1-7.

116. Markel, G. A., Morduchowitz, D., & Catalano, M. C. (2022). A Director's Duty of Oversight after Marchand in "Caremark" Case. The Harvard Law School Forum on Corporate Governance. Retrieved January 24, 2022, from https://corpgov.law.harvard.edu/2022/01/23/a-directors-duty-of-oversight-after-marchand-in-caremark-case/?utm_content=buffer32a3b&utm_medium=social&utm_source=linkedin.com&utm_campaign=buffer

117. Marr, B. (2018, February 14). The Key Definitions Of Artificial Intelligence (AI) That Explain Its Importance. Forbes. https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/?sh=7bc534104f5d

118. Marr, B. (2018). 10 Amazing Examples Of How Deep Learning AI Is Used In Practice?.

119. Mars, C., Dès, R., & Boussard, M. (2020). The three stages of Explainable AI: How explainability facilitates real world deployment of AI. Research Gate.

120. Mars, C., Dès, R., & Boussard, M. (2020). The three stages of Explainable AI: How explainability facilitates real world deployment of AI. Research Gate.

121. Marwala, T. (2020). Closing the gap : the fourth industrial revolution in africa. Macmillan.

122. McCarthy, J. (2007). What is artificial intelligence?.

123. McMullin, C. (2021). Transcription and qualitative methods: Implications for third sector research. VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations, 1-14.

124. McNulty, T., Zattoni, A., & Douglas, T. (2013). Developing corporate governance research through qualitative methods: A review of previous studies. Corporate Governance: An International Review, 21(2), 183-198.

125. Merriam, S. B. (2002). Introduction to qualitative research. Qualitative research in practice: Examples for discussion and analysis, 1(1), 1-17.

126. Metz, T. (2007). Toward an African moral theory. Journal of Political Philosophy, 15(3), 321-341.

127. Metz, T. (2021). African reasons why AI should not maximize utility. In African values, ethics, and technology: Questions, issues, and approaches (pp. 55-72). Cham: Springer International Publishing.

128. Mhlambi, S. (2020). From rationality to relationality: ubuntu as an ethical and human rights framework for artificial intelligence governance. Carr Center for Human Rights Policy Discussion Paper Series, 9, 31.

129. Mhlanga, D., & Moloi, T. (2020). The Stakeholder Theory In The Fourth Industrial Revolution. International Journal of Economics and Finance Studies, 12(2), 198-214.

130. Michalos, A. C., & Poff, D. C. (Eds.). (2012). Citation classics from the journal of business ethics: Celebrating the first thirty years of publication (Vol. 2). Springer Science & Business Media

131. Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and 'the dark side'of AI. European Journal of Information Systems, 1-12.

132. Mittelstadt, B. (2019). AI Ethics–Too principled to fail. arXiv preprint arXiv:1906.06668.

133. Mittelstadt, B. (2019). AI Ethics–Too principled to fail. arXiv preprint arXiv:1906.06668.

134. Mizruchi, M. S. (1983). Who controls whom? An examination of the relation between management and boards of directors in large American corporations. Academy of management Review, 8(3), 426-435.

135. Mle, T. R. (2012). Professional and ethical conduct in the public sector. Africa's Public Service Delivery & Performance Review, 1(1), 26-37.

136. MMC Ventures (2019) The state of AI 2019: Divergence. Report. London, UK: MMC Ventures.

137. Moloi T., Marwala T. (2020) The Agency Theory. In: Artificial Intelligence in Economics and Finance Theories. Advanced Information and Knowledge Processing. Springer, Cham. https://doi.org/10.1007/978-3-030-42962-1_11

138. Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2019). From what to how. An overview of AI ethics tools, methods and research to translate principles into practices. arXiv preprint arXiv:1905.06876.

139. Mostert, L., & Tembedza, W. (2020). Artificial Intelligence has POPIA implications. Webber Wentzel.

140. Mulamula, X. R. (2018). Evaluating the impact of Board members' appointment processes on corporate governance in South African State-Owned Enterprises(Master's dissertation). University of South Africa.

141. Mulamula, X. R. (2021). The role of the board in information technology governance at the Airport Company South Africa. University of the Witwatersrand, Johannesburg.

142. Mulamula, X. R. (2022, May). The Role of the Board in Artificial Intelligence Ethics. Retrieved September 22, 2022, from https://www.iodsa.co.za

143. Mulamula, X. R., & Lushaba, S. (2020). Who is responsible? AI vs corporate governance and SA law. Bizcommunity. Retrieved October 1, 2020, from https://www.bizcommunity.com/Article/196/547/208888.html

144. Mulligan, D., Kroll, J., Kohli, N. & Wong, R. (2019). This thing called fairness: Disciplinary confusion realizing a value in technology. ACM Human-Computer Interaction, 3, 119. https://doi.org/10.1145/3359221.

145. Murdoch, B. (2021). Privacy and artificial intelligence: challenges for protecting health information in a new era. BMC Medical Ethics, 22(1), 1-5.

146. Murphy, S. A., & McIntyre, M. L. (2007). Board of director performance: a group dynamics perspective. Corporate Governance: The international journal of business in society.

147. Muswaka, L. (2014). The two-tier board structure and co-determination: Should South Africa follow the Germany example?. Mediterranean Journal of Social Sciences, 5(9), 142-142.

148. Naturalistic inquiry. Improving Human Performance Quarterly, 8(4), 268-76.

149. Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. In Algorithms of oppression. New York university press.

150. Ntim, C. G. (2009). Internal corporate governance structures and firm financial performance: evidence from South African listed firms (Doctoral dissertation, University of Glasgow).

151. O'Neal, D., & Thomas, H. (1995). Director networks/director selection: The board's strategic role. European Management Journal, 13(1), 79-90.

152. O'Neill, J. (2001). Meta-ethics. A companion to environmental philosophy, 163-176.

153. Pearl, J. (2019). The limitations of opaque learning machines. Possible minds: twenty-five ways of looking at AI, 13-19.

154. Peregrine, M. (2022). The Looming Board Challenge: Oversight Of Artificial Intelligence. Forbes. Retrieved May 4, 2022, from

155. Perifanis, N. A., & Kitsios, F. (2023). Investigating the influence of artificial intelligence on business value in the digital era of strategy: A literature review. Information, 14(2), 85.

156. Petrin, M. (2019). Corporate Management in the Age of AI. Colum. Bus. L. Rev., 965.

157. Petropoulos, G. (2018). The impact of artificial intelligence on employment. Praise for Work in the Digital Age, 119, 121.

158. Polkinghorne, D. E. (1989). Phenomenological research methods. In Existential-phenomenological perspectives in psychology: Exploring the breadth of human experience (pp. 41-60). Boston, MA: Springer US.

159. Purohit, A. (2023). AI, ML, DL, and Generative AI face off: A Comparative analysis | Synoptek. https://synoptek.com/insights/it-blogs/data-insights/ai-ml-dl-and-generative-ai-face-off-a-comparative-analysis/#:~:text=Generative%20AI%2C%20a%20branch%20of,might%20be%20created%20by%20humans.

160. Raab, C. D. (2020). Information privacy, impact assessment, and the place of ethics. Computer Law & Security Review, 37, 105404.

161. Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., ... & Barnes, P. (2020, January). Closing the AI accountability gap: Defining an end-to-end framework for

internal algorithmic auditing. In Proceedings of the 2020 conference on fairness, accountability, and transparency (pp. 33-44).

162. Raso, F. A., Hilligoss, H., Krishnamurthy, V., Bavitz, C., & Kim, L. (2018). Artificial intelligence & human rights: Opportunities & risks. Berkman Klein Center Research Publication, (2018-6).

163. Renn, O. (2017). Risk governance: coping with uncertainty in a complex world. Routledge.

164. Richardson, R., Schultz, J. M., & Crawford, K. (2019). Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. NYUL Rev. Online, 94, 15.

165. Rockness, H., & Rockness, J. (2005). Legislated ethics: From Enron to Sarbanes-Oxley, the impact on corporate America. Journal of Business Ethics, 57(1), 31-54.

166. Rotman, L. (2017). Understanding Fiduciary Duties and Relationship Fiduciarity. McGill Law Journal/Revue de droit de McGill, 62(4), 975-1042.

167. Russell, S., & Norvig, P. (2002). Artificial intelligence: a modern approach.

168. Ruttkamp-Bloem, E. (2021). The Quest for Actionable AI Ethics. In Southern African Conference for Artificial Intelligence Research (pp. 34-50). Springer, Cham.

169. Rwegasira, K. (2000). Corporate governance in emerging capital markets: Whither Africa?. Corporate Governance: An International Review, 8(3), 258-267

170. Sahut, J. M., Dana, L. P., & Teulon, F. (2021). Corporate governance and financing of young technological firms: A review & introduction. Technological Forecasting and Social Change, 163, 120425.

171. Schiff, D., Biddle, J., Borenstein, J., & Laas, K. (2020). What's Next for AI Ethics, Policy, and Governance? A Global Overview. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (pp. 153-158).

172. Schneider, A., Wickert, C., & Marti, E. (2017). Reducing complexity by creating complexity: A systems theory perspective on how organizations respond to their environments. Journal of Management Studies, 54(2), 182-208.

173. Schneider, J., Abraham, R., & Meske, C. (2020). AI Governance for Businesses. arXiv preprint arXiv:2011.10672.

174. Schroff, F., Kalenichenko, D., & Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 815-823).

175. Seger, E. (2022). In Defence of Principlism in AI Ethics and Governance. Philosophy & Technology, 35(2), 1-7.

176. Sharma, N., Sharma, R., & Jindal, N. (2021). Machine Learning and Deep Learning Applications-A Vision. Global Transitions Proceedings, 2(1), 24-28.

177. Shehabi, A., Smith, S., Sartor, D., Brown, R., Herrlin, M., Koomey, J., ... & Lintner, W. (2016). United states data center energy usage report.

178. Siau, K., & Wang, W. (2020). Artificial intelligence (AI) ethics: ethics of AI and ethical AI. Journal of Database Management (JDM), 31(2), 74-87.

179. Sidorova, A., & Rafiee, D. (2019). AI Agency Risks and their Mitigation through Business Process Management: a Conceptual Framework. In Proceedings of the 52nd Hawaii International Conference on System Sciences.

180. Silverman, K. (2020). Why Your Board Needs a Plan for AI Oversight. MIT Sloan Management Review. https://sloanreview.mit.edu/article/why-your-board-needs-a-plan-for-ai-oversight/

181. Skilton, M., & Hovsepian, F. (2017). The 4th industrial revolution: Responding to the impact of artificial intelligence on business. Springer.

182. Skilton, M., & Hovsepian, F. (2017). The 4th industrial revolution: Responding to the impact of artificial intelligence on business. Springer.

183. Solomon, J. (2020). Corporate governance and accountability. John Wiley & Sons.

184. Solove, D. J. (2006). A taxonomy of privacy. University of Pennsylvania law review, 477-564.

185. Stahl, B. C. (2021). Ethical issues of ai. In Artificial Intelligence for a Better Future (pp. 35-53). Springer, Cham.

186. Stuart, R., & Peter, N. (2016). Artificial intelligence-a modern approach 3rd ed.

187. Suresh, H., & Guttag, J. V. (2019). A framework for understanding unintended consequences of machine learning. arXiv preprint arXiv:1901.10002, 2(8), 73.

188. Taulli, T. (2020). AI (Artificial Intelligence) Governance: How To Get It Right. Forbes. https://www.forbes.com/sites/tomtaulli/2020/10/10/ai-artificial-intelligence-governance-how-to-get-it-right/?sh=43220d3a745f

189. Taylor, S. J., & Bogdan, R. (1984). Introduction to qualitative research methods: The search for meanings. Wiley-Interscience.

190. The Lighthill Debate (1973) - part 3 of 6 https://youtu.be/RnZghm0rRlI?t=10m8s.

191. Theron, H. F., & Koornhof, P. G. J. (2016). Bow to the King (IV)? A new era for IT governance in South Africa. In African Cyber Citizenship Conference 2016 (ACCC2016) (p. 161).

192. Thomaz, F., Efremova, N., Mazzi, F., Clark, G., MacDonald, E., Hadi, R., ... & Stephen, A. T. (2021). Ethics for AI in Business. Available at SSRN.

193. Thuraisingham, B. (2020). Artificial Intelligence and Data Science Governance: Roles and Responsibilities at the C-Level and the Board. In 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI) (pp. 314-318). IEEE.

194. Tian, G. Y. (2016). Current Issues of Cross-Border Personal Data Protection in the Context of Cloud Computing and Trans-Pacific Partnership Agreement: Join or Withdraw. Wis. Int'l LJ, 34, 367.

195. Torre, F., Teigland, R., & Engstam, L. (2019). 7 AI leadership and the future of corporate governance. The Digital Transformation of Labor, 116.

196. Tóth, Z., Caruana, R., Gruber, T., & Loebbecke, C. (2022). The Dawn of the AI Robots: Towards a New Framework of AI Robot Accountability. Journal of Business Ethics, 1-22.

197. Towell, E., Thompson, J. B., & McFadden, K. L. (2004). Introducing and developing professional standards in the information systems curriculum. Ethics and information technology, 6(4), 291-299

198. Ugowe, A. O. (2016). Monitoring good corporate governance in developing countries: Evidence from Nigeria. Nnamdi Azikiwe University Journal of International Law and Jurisprudence, 7, 125-133.

199. Valle, R. S., King, M., & Halling, S. (1989). An introduction to existential-phenomenological thought in psychology. In Existential-phenomenological perspectives in psychology: Exploring the breadth of human experience (pp. 3-16). Boston, MA: Springer US.

200. Van Rijmenam, M., Schweitzer, J., & Williams, M. A. (2019). Overcoming Principal-Agent Problems When Dealing with Artificial Agents: Lessons for Governance from a Conversation with Tay.

201. Vassilopoulos, A. P., & Georgopoulos, E. F. (2010). Novel computational methods for fatig life modeling of composite materials. In Fatigue life prediction of composites and composite structures (pp. 139-173). Woodhead Publishing.

202. Veasey, E. N. (1997). The Defining Tension in Corporate Governance in America. The Business Lawyer, 393-406.

203. Velasquez, M., Andre, C., Shanks, T., & Meyer, M. J. (2010). What is Ethics? Markkula Center for Applied Ethics. Retrieved June 15, 2022, from https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/what-is-ethics/

204. Verbeke, A., & Hutzschenreuter, T. (2020). The dark side of digital globalization. Academy of Management Perspectives, (ja).

205. Vincent-Lancrin, S., & Van der Vlies, R. (2020). Trustworthy artificial intelligence (AI) in education: Promises and challenges.

206. Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the general data protection regulation. International data privacy law, 7(2), 76-99.

207. Waelen, R. A. (2022). The struggle for recognition in the age of facial recognition technology. AI and Ethics, 1-8.

208. Wagner, C., Kawulich, B., & Garner, M. (2012). EBOOK: Doing social research: A global context. McGraw Hill.

209. Wallach, W., & Allen, C. (2009). Moral machines: Teaching robots right from wrong. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195374049.001.0001

210. Weill, P., & Woerner, S. L. (2017). Is Your Company Ready for a Digital Future?. MIT Sloan Management Review.

211. Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., ... & Schwartz, O. (2018). AI now report 2018(pp. 1-62). New York: AI Now Institute at New York University.

212. Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). The role and limits of principles in AI ethics: towards a focus on tensions. In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (pp. 195-200).

213. World Economic Forum. (2020). Governance: Empowering AI Leadership. Https://Www.Weforum.Org. Retrieved May 11, 2022, from https://express.adobe.com/page/3MEmZh5EprOXJ/

214. Xu, Z., & Sun, J. (2018). Model-driven deep-learning. National Science Review, 5(1), 22-24.

215. Yang, J., Jin, H., Tang, R., Han, X., Feng, Q., Jiang, H., ... & Hu, X. (2023). Harnessing the power of llms in practice: A survey on chatgpt and beyond. arXiv preprint arXiv:2304.13712.

216. Yin, R. K. (2003). Designing case studies. Qualitative research methods, 5(14), 359-386.

217. Zhang, Y., Wu, M., Tian, G. Y., Zhang, G., & Lu, J. (2021). Ethics and privacy of artificial intelligence: Understandings from bibliometrics. Knowledge-Based Systems, 222, 106994.

ANNEXURE A: ETHICAL CLEARANCE

Faculty of Humanities
Fakulteit Geesteswetenskappe
Lefapha la Bomotho

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Humanities 100.
Since 1919

14 September 2021

Dear Mr XR Mulamula

Project Title:	The role of the board in artificial intelligence ethics and governance – a case for JSE listed companies
Researcher:	Mr XR Mulamula
Supervisor(s):	Prof EB Ruttkamp-Bloem
Department:	Philosophy
Reference number:	17306699 (HUM021/0621)
Degree:	Doctoral

I have pleasure in informing you that the above application was **approved** by the Research Ethics Committee on 26 August 2021. Data collection may therefore commence.

Please note that this approval is based on the assumption that the research will be carried out along the lines laid out in the proposal. Should the actual research depart significantly from the proposed research, it will be necessary to apply for a new research approval and ethical clearance.

We wish you success with the project.

Sincerely,

Prof Karen Harris
Chair: Research Ethics Committee
Faculty of Humanities
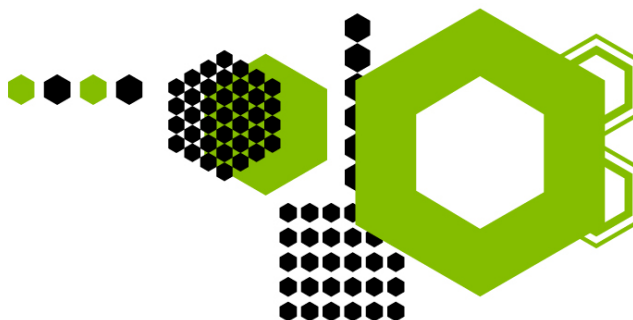**UNIVERSITY OF PRETORIA**
e-mail: tracey.andrew@up.ac.za

ANNEXURE B: IODSA CONSENT LETTER

144 Katherine Street
Grayston Ridge Office Park
Block B 1st floor
Sandown Sandton
2196
PO Box 521372
Saxonwold
2132

T +27 11 035 3000
E info@iodsa.co.za

INSTITUTE OF
DIRECTORS
SOUTH AFRICA

3 June 2021

Mr XR Mulamula
5 Fairview Thornhill Estate
Bendor
Polokwane
0699

Per email: mulamula@gmail.com

Dear Mr Mulamula,

RE: PERMISSION TO UTILISE THE IODSA'S GOVERNANCE RESEARCH PLATFORM

I acknowledge receipt of your request to conduct research through the IoDSA's Governance Research Platform for your Doctor of Philosophy studies, by means of interviews with JSE non-executive directors.

I am pleased to inform you that your request to use the IoDSA's Governance Research Platform has been approved for your topic *The Role of the Board in Artificial Intelligence Ethics and Governance – A Case for JSE Listed Companies.*

Whilst undertaking the research, the following terms and conditions apply:
1. You are to honour the appointments made with the interviewees.
2. All information gathered remains strictly confidential at all times.
3. To ensure compliance with the Protection of Personal Information Act.

Kindly note that a condition of the permission to utilise the IoDSA's Governance Research Platform, requires you, within 6 months of completing of your final thesis, to issue high level findings in an executive summary to IoDSA members and/or an article in Directorship magazine.

Yours sincerely.

Julie Dixon
Manager: Forums and Research

BETTER DIRECTORS.
BETTER BOARDS.
BETTER BUSINESS.

ANNEXURE C: PARTICIPANT(S) LETTER OF INTRODUCTION

XR Mulamula
5 Fairview
Thornhill Estate
Bendor
Polokwane, 0699

Date: 01 March 2022

RESEARCH PARTICIPANT LETTER OF INTRODUCTION AND INFORMED CONSENT

Dear Sir / Madam

My name is Xitshembhiso Russel Mulamula, and I am a Doctor of Philosophy student at the University of Pretoria. As part of my studies, I have to undertake a research project, and I am investigating the board's role in artificial intelligence ethics and governance (A case for JSE listed companies) under the supervision of Prof Emma Ruttkamp-Bloem. The research aims to investigate how boards perform their functions over the ethics and governance of Artificial Intelligence.

As part of the research project, I have selected participants serving in JSE listed boards and are members of the Institute of Directors South Africa (IODSA). Thus, I would like to invite you to participate in the interviews. The research will be conducted under the IODSA research platform, which aims to assist students working on governance-related projects to obtain practical research input from IoDSA members that can be used in their research findings and end deliverables.

The research will follow a qualitative research approach. Thus, this activity will involve semi-structured interviews questions, which will be centred around the board's role in AI ethics and governance. The interviews will take around 60 minutes. Interviews will be conducted through virtual platforms. I would also like to record the virtual interview; the recordings will be deleted directly after transcribing and analysing the data.

There will be no personal costs to you if you participate in this project. You will not receive any direct benefits from participation, but there are no disadvantages or

1

penalties if you choose not to participate or withdraw from the study. You may cancel or not answer any questions if you do not want to. The interview will be completely confidential and anonymous as I will not be asking for any identifying information. The information you provide will be held securely and not disclosed to anyone else in any identifying manner. All research data will be stored in a password protected format in the Department of Philosophy. I will be using a pseudonym (false name) to represent your participation in my final research report. If you experience any distress or discomfort in this process, we will stop the interview or resume another time.

If you have any questions during or afterwards about this research, feel free to contact me on the details listed below. This study will be written up as a research report online through the University of Pretoria library website. If you wish to receive a summary of this report, I will be happy to send it to you. The data collected from this research project will be stored in a safe and password protected cloud storage in the Department of Philosophy and kept for three years.

If you have any concerns or complaints regarding the ethical procedures of this study, you are welcome to contact the Ethic Administrators for Humanities: tracey.andrew@up.ac.za or dorcas.mokalapa@up.ac.za.

Attached is the research approval letter from the Institute of Directors South Africa.

Yours sincerely,

Researcher:
Xitshembhiso Russel Mulamula, email u17306699@tuks.co.za, 081 432 0766

Supervisor:
Prof Emma Ruttkamp-Bloem, email emma.ruttkamp-bloem@up.ac.za
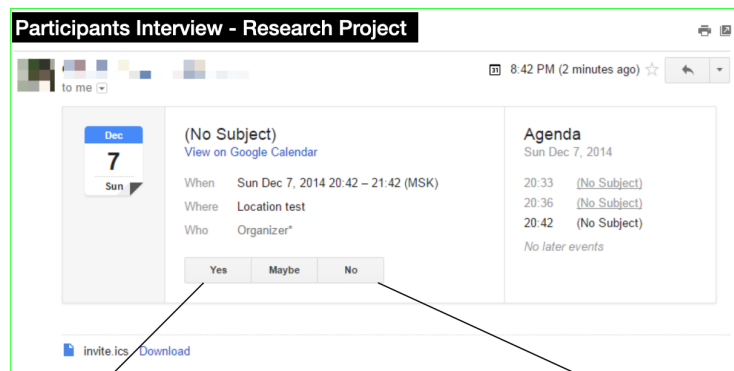
2

ANNEXURE D: PARTICIPANT(S) INFORMED CONSENT

RESEARCH PARTICIPANT CONSENT

The introductory email contains a research interview calendar invitation link with it:

- By selecting the **"Agree/ Yes"** option, I grant my permission to participate in this research interview. I hereby confirm that I have understood the purpose of this research.
- By selecting the **"Agree/ Yes"** option, I consent my responses to be used anonymously in this research report and in any subsequent academic publications that stem from this research report.
- By selecting the **"decline/ no"** option, I am excising my right not to participate in the research.

Example: Research Interview Calendar Invitation Link.

ANNEXURE E: RESEARCH QUESTIONS

The Role of the Board in Artificial Intelligence Ethics and Governance – A Case for JSE Listed Companies

General Questions for participants profiling.

What is your level of education?
What industry defines your skills and expertise?
How many years of experience in the board(s) of JSE?
How many board of JSE do you sit in?
Do you have background in technology?

Theme 1: How does the board discharge its fiduciary responsibility in overseeing AI technologies and its potential impact on the organisation?

- What is your understanding of artificial intelligence?
- What is your understanding of the general role of the board as far as fiduciary duties are concerned?
- What is your understanding of the role of the board as far as technology i.e. artificial intelligence is concerned?
- How does the board oversee artificial intelligence technologies in the organisation?
- What is your understanding of ethical concerns around artificial intelligence technologies in terms of possible societal harm ?

Theme 2: What is the board's role in ensuring that AI technologies are aligned with the organisation's ethical culture?

- What is your understanding of how the board play a key role in setting the ethical tone in the organisation?
- What is your understanding of how artificial intelligence technologies may impact on the ethical culture of the organisation?
- What is your understanding of how artificial intelligence technologies strategies are aligned to the organisation's ethical culture?

Theme 3: What is the board's role in ensuring that the research or design or development or deployment or use of AI technologies (whichever is applicable in your organisation's context) does not negatively impact on society?

- What is your understanding of the responsibility of the board in mitigating societal harm from AI technologies in play in your organisation?

Theme 4: What is the board's role in ensuring that adequate AI governance structures are in place?

- Does your board have a governance subcommittee for technology or AI?
    - o If yes, what is your understanding of the role of the AI or technology governance subcommittee?
    - o If no, do you think such a subcommittee is needed? Why or why not?
- Does your board have responsible use of AI governance framework?
    - o If yes, what is your understanding of the role of the framework?
    - o If no, do you think such a framework is needed? Why or why not?
- How does the board monitor the performance and value delivery of artificial intelligence technologies?