Governance of AI Ethics: Perspective from the Global South (Africa)

Emile Ormond[1] [0000-0003-3200-2652]

[1]University of South Africa, Graduate School of Business Leadership

Emile.Ormond@gmail.com

Abstract. The literature on artificial intelligence (AI) ethics is dominated by a Global North outlook, despite AI ethics not being uniform across societies. The aim of this exploratory study is to contribute to the literature on AI ethics by offering a governance perspective from the Global South. The study uses a qualitative methodology and semi-structured interviews to explore the views and proposals of South African-based AI practitioners and associated experts on mechanisms, methods, and measures to govern AI. The study identifies themes relevant to the governance of AI ethics for organisations in South Africa, focusing on external and internal measures that influence AI ethics. The study finds that organisations can take a range of measures to address AI ethics. Additionally, the study highlights differences between how AI ethics is approached in South Africa and the Global North.

Keywords: artificial intelligence (AI), governance, ethics, Global North, Global South

1 Introduction

The field of artificial intelligence (AI)1 is rapidly advancing, with the potential to revolutionise various industries. As AI systems become increasingly integrated into society, they raise ethical questions and have been at the centre of scandal and controversy [1]–[10]. A recent example of this is the development of generative AI tools, such as

[1] While there are many definitions of "artificial intelligence," this article broadly considers AI as a human-designed system that decides on the best actions to achieve a given complex goal. It does so by acquiring data through various means, such as machine learning algorithms, sensors, or other sources, interpreting the collected structured or unstructured data, reasoning on the knowledge derived from the data, or processing the information.

ChatGPT [11]–[14]. As the technology is increasingly rolled out at pace [15], it is crucial for organisations and policymakers to consider the ethical implications of AI's development and deployment. However, the field of AI ethics and policy-related work is dominated by the Global North. The lack of input and perspectives from the Global South affects the issues that receive attention and, more importantly, those that do not [16] [17]. This article explores AI ethics from a South African perspective – the country serving as a representative of Africa and, more broadly, the Global South. It provides empirical research that explores South African-based AI practitioners and associated experts' views and proposals on mechanisms, methods, and measures to govern AI ethics. It builds on previous research that compared universal AI ethics risks to those specific to South Africa [18].

2 Trends in the Literature

There are still notable gaps and shortcomings in the AI ethics literature, despite the rapid expansion of the body of work in recent years [4], [16], [19]. There are still relatively few studies providing an empirical account of how the AI fraternity perceive and govern ethics in practice. The literature mostly takes an outside-in view, where findings and recommendations are not explicitly based on empirical data. There is, for instance, a plethora of non-empirical, normative guides and proposals for how organisations should manage AI ethics [20]. Whereas, studies that provide empirical data from practitioners or associated experts are relatively rare and frequently anecdotal [21], resulting in calls for the creation of more practical AI governance insights and proposals, which would have utility to organisations [22].

This gap in empirical AI ethics research is slowly being filled, albeit with varied focus areas, populations, and methodological approaches. Examples of this include Orr and Davis [23] who interviewed a sample of 21 Australian AI practitioners on how they attribute ethical responsibility with AI systems. Moss and Metcalf's [24], who conducted an ethnographic study with 24 "ethics owners" in digital technology companies in Silicon Valley. Morley et al.'s [25] conducted a mixed method qualitative study on UK-based AI practitioners' understanding, motivation, barriers, and application of AI ethics principles and practice. Rakova et al. [26] conducted 26 interviews with practitioners working in the AI industry in a handful of Global North countries to investigate common challenges for "responsible AI" initiatives. Ryan et al. [27] held workshops with 19 primarily Global North AI practitioners to explore the tensions between AI individual ethical values versus organisational values. Stahl et al. [21] present empirical findings, based on ten case studies in the Global North, of how companies approach AI ethics.

The field of AI ethics has made considerable progress in recent years, producing a variety of guides, policies, frameworks, principles, and values for improving the ethical (or responsible) development and utilisation of AI [16]. These resources ostensibly in-

fluence how the public and private realms will design, develop, deploy, utilise and regulate AI [19]. However, the production of knowledge related to AI ethics mimics reflects the composition of the predominant AI industry and workforce – lacking diversity and centred in a few hubs [20], [28][29] [30]–[33]. This lack of diversity is also reflected in the creation of AI ethics codes, policies, and frameworks [34]. The most prevalent literature on AI ethics is primarily produced and published by North American and European-based entities [4], [19], [35], [36]. Specifically, Africa's contribution to the AI ethics literature has been "very weak" [37], although recent efforts have been made to reverse this trend [38]. Closely linked to the Global North dominance, AI ethics is mostly approached and portrayed in universalistic sense [17]. This often implies an assumption that the constituent elements of AI ethics (e.g., values, concepts, risks, morality, norms) are the same everywhere. However, there is little research that considers and explores how AI ethics, both in theoretical and pragmatic terms, are unlikely to be the same for disparate groups at an inter-regional or country level [33], [39]–[42]. This is a significant gap because the perception, practice, and impact of AI ethics will vary among entities with disparate conditions [18], [43]–[45], [46], [47]. This presents a problem, as the ethical frameworks and standards for AI reflect the value systems and power structures of the Global North [19]. These Western values are not necessarily universal, and may differ from other cultures, including the heterogeneous populations in Africa [37], [38]. Despite this, most literature on AI ethics does not make a meaningful attempt to explore the differences between how the dominant narrative in the Global North and developing regions view and approach AI ethics [38], [48].

3 Methodology

The research explored the views and proposals of the AI community members in South Africa on mechanisms, methods, and measures to govern AI ethics. The research is exploratory in nature, and consequently, the empirical research used an inductive and qualitative approach [49]. The unit of analysis is South Africa's AI industry, which includes organisations that specialise in AI-related products or services and the individuals who constitute these organisations. The unit of observation is on three corresponding levels to allow for source triangulation. These levels include individuals involved in an AI-driven organisation (industry participants), individuals in ancillary areas such as academia and research (expert participants), and individuals who have elements of both categories (hybrid participants). Sixteen study participants were identified using a combination of purposive and snowball sampling, and they included seven industry participants (44 percent), five expert participants (31 percent), and four hybrid participants (25 percent). Data collection took place via semi-structured interviews, using a novel research instrument, and was part of a broader research undertaking on AI ethics in South Africa [18]. The instrument was reviewed by two subject matter experts and two qualitative methodology experts for inter alia relevance and validity. Data collection occurred via online interviews during January and May 2022, and interviews

lasted on average 45-60 minutes. The data was iteratively reviewed and analysed via a hybrid inductive-deductive approach to identify codes and themes via ATLAS.ti. The study received ethics approval from the university's research ethics committee.

4 Findings

The findings were divided into two overarching themes according to the level of applicability. The first at the macro level and, the second, at an organisational level. In other words, how participants see the external environment in which organisations exist and operate, on the one hand, and, how participants see internal, organisational dynamics, on the other hand.

4.1 External

The participants highlighted several high-level, themes in the macro environment that affect the perception and treatment of AI ethics. Table 1 provides an overview of the themes.

Table 1. Overview of External Themes

Theme	Description
Self-regulation insufficient	Self-regulation may benefit unscrupulous actors
Government oversight required	State best positioned to set and enforce regulations
Multi-stakeholder dialogue	Regulation should involve multi-actor dialogue
Horses for courses	Bespoke regulations for different sectors
Existing governance code lacking	AI corporate governance guidance needed
International obligations	South Africa's global AI ethics obligations

Self-Regulation Insufficient. There was nearly unanimous scepticism across all participant categories for only having self-regulation measures in place to govern AI, either at an industry or enterprise-level. Participants noted that self-regulation inherently brings forth significant challenges and limitations. Many raised questions over whether an organisation that is fundamentally driven by profit should self-regulate and to whose

benefit this would be. There was a broad consensus across all participant categories that there should be some form of external, mandatory regulation. Industry participants noted that this would create an equal playing field and set clear expectations and requirements. Whereas now, organisations with lower ethical standards could benefit relative to ones that have higher ethical standards. Moreover, all participants noted that organisations generally take existing mandatory measures seriously and suggested that organisation would follow suit if there were similar requirements for AI.

Government Oversight Required. There were various views regarding the form and function of AI regulation, including who should be responsible, what it should entail, and what should be included or excluded. The most common view, across participant categories, was for government, in some shape or form, to be responsible for regulating AI. There was, however, scepticism over the South African government's political will, resources, capacity, and technical competence to effectively play such a role. Although it was noted that many governments, including those in the Global North, were also grappling with regulating fourth industrial revolution (4IR) technology. A handful of expert participants, however, noted that AI regulations would, even with limited implementation or enforcement, at least set expectations for acceptable behaviour and help shape the ethical milieu.

Multi-Stakeholder Dialogue. Participants indicated that regardless of the form or source of oversight, regulation cannot be merely a top-down implementation of laws, rules, or requirements. Instead, it should include a multi-stakeholder dialogue that involves enterprises, government, civil society, and the public. Through this dialogue, the different stakeholders can determine what measures need to be put in place to govern AI. Furthermore, any governance framework for AI should strike a balance between proper oversight and regulation on one end and encouraging innovation and growth in the industry on the other.

Horses for Courses. A handful of industry participants remarked that regulation should not consist of a one-size-fits-all approach, framework, or requirement. Instead, different sectors should have different requirements based on their risk levels, particularly with regards to people's rights. Additionally, it is worth considering that regulation may be more necessary or appropriate in certain already regulated sectors, such as financial services and healthcare. For example, healthcare and financial services often involve sensitive personal information, which requires a higher level of security and privacy protection than other industries. Similarly, autonomous vehicles or drones pose different risks than other AI applications.

Existing Governance Code Lacking. Participants noted that the prevailing corporate governance mechanisms is lacking in terms of its relevance in governing AI ethics risks. In South Africa, the preeminent corporate governance code is the King Report on Corporate Governance (King Code). However, most industry participants made remarks suggesting that they did not consider the King Code relevant to AI governance. Moreover, sections in the King Code that addresses technology were said to be too

generic and did not sufficiently account for AI's unique features.

International Obligations. Expert participants noted that South Africa is directly or indirectly influenced by existing global agreements, such as the UNESCO Recommendations on the Ethics of AI and the OECD Principles on AI, as well as foreign statutory requirements such as the General Data Protection Regulation (GDPR), which are relevant to AI and related areas such as data governance. However, at an African level, there are currently no AI or data management obligations in place by regional bodies such as the African Union or the Southern African Development Community.

4.2 Internal

The participants identified an eclectic range of internal governance and management measures that are relevant to organisations with regards to AI ethics. Table 2 provides an overview of the themes.

Table 2. Overview of Internal Themes

Theme	Description
Awareness of ethics	Starting point of AI ethics is awareness, conceptual clarity
Bottom-up consultation	Organisations must consult with stakeholders
Diverse and informed staff	Plurality in workforce composition, leaders knowledgeable of AI
Develop existing frameworks	Adjust existing (Global North) ethics frameworks for local context
Tailored path	Universal approach not feasible, desirable
Expand existing structures	Build upon strategy, vision, values and enterprise risk governance

Awareness of Ethics. Participants noted that recognizing AI ethics as a critical issue is the starting point for governance. In other words, awareness is a meta-measure and a necessary precondition for organisations to manage AI ethics effectively. Furthermore, AI ethics should be regarded as a holistic and interdisciplinary phenomenon – it is not solely a technical issue, nor limited to ethicists.

Bottom-up Consultation. Expert and industry participants called for organisations to

adopt a bottom-up, consultative approach with stakeholders. In other words, they recommended engaging with those who are directly affected by the technology and not simply imposing it on them from the top-down. This includes considering ethics at all stages of the AI life cycle. Organisations must recognize the importance of engaging with stakeholders in the development and deployment of AI. This includes engaging with end-users, customers, and other stakeholders who are directly impacted by the technology. By involving these stakeholders in the development of AI, organisations can ensure that the technology is designed and deployed in a manner that is ethical, transparent, and inclusive.

Diverse and Informed Staff. Expert and hybrid participants noted that the composition of an organisation's workforce and leadership needs to be diverse in terms of disciplinary approach (e.g., social science, computer science, humanities) and demographics (e.g., age, gender, race, and ethnicity). This diversity would facilitate the responsible and ethical development and utilization of AI by providing staff with a broader variety of perspectives and backgrounds to assess ethically relevant decisions. Moreover, senior leaders and governing bodies need to be more astute in understanding the technology and its socio-technical nature and consequences. This includes developing a deeper understanding of the ethical implications of AI and the broader societal impact of the technology. This understanding should be integrated into the organisation's decision-making processes and governance structures to ensure that ethical considerations are considered at all levels of an organisation.

Develop Existing Frameworks. Many participants across categories mentioned that local companies have access to a number of AI ethics frameworks. However, all of these frameworks were developed by entities in the Global North, including big multinational consulting firms, large technology companies, and professional associations. A handful of expert and hybrid participants noted that these frameworks cannot simply be cut-and-pasted into the Global South or South African context. In addition, there are practical challenges for organisations in operationalizing these frameworks in their daily workflows.

Tailored Path. Industry and expert participants indicated that a one-size-fits-all approach to AI ethics is unsuitable because different organisations, or potentially even different business units within an enterprise, may face dissimilar ethical questions. Therefore, organisations should adopt a tailored approach to AI ethics governance and management. The appropriate measure will be influenced by variables such as industry, size, operating model, maturity level, and organisational culture.

Expand Existing Structures. Industry participants noted that AI ethics governance need not be an entirely new initiative, but rather can build on existing organisational features and dynamics. For example, an organisation's current mission, vision, strategy, values, governance structures, and management measures can be leveraged to support AI ethics governance. By building on existing features, organisations can create a more integrated and sustainable approach to AI ethics governance.

5 Discussion

This section discusses the relevance of the external and internal findings, among South African participants. It ends by briefly contrasting some of the major South African themes with the Global North.

5.1 External

The first two themes, 'self-regulation insufficient' and 'government oversight required', relate to the necessity and nature of regulation and supervision. The findings reiterate the existing body of literature that is critical and sceptical of exclusive industry self-regulation [1], [3], [50]–[52]. Notably, industry participants expressed misgivings about self-regulation and were unanimous in the need for some form of external regulation. This stands in contrast to what one might expect, as organisations would ostensibly have the most to gain from not having any supervision. The South African participants' calls for regulation echo similar appeals from leading technology companies in the Global North [53]. Industry participants argued that regulation would establish an equal playing field by demarcating acceptable conduct for all organisations. Instead of inhibiting innovation, a pervasive risk with regulation, it may allow the industry to act more freely by demarcating a fence of acceptable conduct and result in a net gain of innovation [54]. There is self-reported evidence to suggest that organisations would comply with mandatory external regulation. The participants, for instance, claimed that their organisations are adhering to existing South African legal requirements on the collection, storage, and treatment of data. Notwithstanding, until there is some form of external oversight, the de facto position will be a continuation of the status quo where organisations self-define acceptable and ethical conduct. South African practitioners could, in the absence of external regulation, adopt a voluntary model, similar to some US-based organisations [55].

The most common view was that the government is best positioned to provide external oversight of AI, given that it has the mandate and authority to formulate and enforce regulations in the interest of all societal stakeholders. This view, however, seems to stand in contrast to the South African government's approach, which has shown limited appetite to guide and regulate AI. Rather, the South African government appears almost exclusively focused on the technology as a socio-economic tool. There is, based on prevailing policy papers and officials' remarks, not much focus per se on the responsible and ethical use of 4IR technologies [56]–[58]. Moreover, the country still does not have a national AI strategy, which puts it in a minority among middle-income countries [59]. In contrast, other African states such as Egypt, Kenya, Mauritius, and Rwanda have

more elaborate AI policies and bodies [60]–[62]. The South African government's relative indifference diverges even more from the Global North where countries such as the EU, UK, and US have comprehensive national strategies [63]–[66].

There is little evidence in the literature that the commercial use of AI, either in South Africa or elsewhere, takes place in the context of meaningful consultation between organisations and its stakeholders [24]. A 'multi-stakeholder dialogue', which dovetails with the internal theme of 'bottom-up consultation', consists of key stakeholders such as government, industry, civil society, and citizens having an in-depth consultation on the technology. Such a dialogue would ostensibly provide more legitimacy and transparency to the development and use of AI, which is currently a top-down, elite-driven, and imposed endeavour [67]. A multi-stakeholder dialogue and bottom-up consultations could also pre-emptively forestall AI-related ethical scandals as the population would have been consulted on, for instance, how and where AI is used. A multi-stakeholder dialogue would present both representation questions (e.g., who gets to participate, what is their mandate) and logistical challenges (e.g., how is it constituted? How do you get wide-spread participation?). However, participants did not delve into these broader and practical aspects of this proposal. Nonetheless, there are existing outreach models that could be emulated for this type of consultation. For instance, parliamentary consultations.

There is practical merit in the proposal of having different regulations for different sectors, as outlined by the 'horses for courses' theme. This ties up with the literature, which has indicated that organisations in diverse sectors would be affected and need to have different approaches to the responsible use of AI [68]–[70]. There is little overlap between how AI may be used, for instance, in the agriculture, financial services, health, and tourism sectors. This view is reflected in the UK's proposed AI regulations, which involve various oversight actors [65]. Moreover, guidance or regulations would need to have a certain level of granularity and applicability to be practically useful in different sectors. Putting aside questions over who would be responsible for regulation, it would nonetheless be exceedingly cumbersome to formulate different requirements for each sector, especially in a government with constrained resources – which is the case for most Global South countries. Similarly, it would be equally challenging to monitor and evaluate compliance. Indeed, it may be more feasible for external governance to only apply to a handful of key areas, which are associated with fundamental human rights. Such an approach would broadly correspond with the regulatory model (i.e., AI Act) that the EU has opted for, where the level of regulation and oversight of AI systems corresponds to its potential harm and negative impact on fundamental rights [71]. This theme interlinks with the 'tailored path', where each organisation adopts an ethics strategy that is fit-for-purpose to its, inter alia, industry, maturity-level, and place on the AI value chain. Meaning that there are no off-the-shelf approaches or solutions that would likely be applicable to all organisations. Rather, leadership will need to tailor organisational approaches, albeit on the template of pre-existing frameworks and regulatory requirements.

Regarding existing corporate governance frameworks, the King Code was found to be inadequate to deal with the specific ethical and governance challenges presented by

AI. The King Code provides a generic approach to technology governance, lacking nuance especially on 4IR technologies [72]. The King Code's shortcoming means that organisations, in terms of corporate governance best practice, have no specific obligations nor guidance with regards to AI, or 4IR more broadly. The Institute of Directors South Africa, the body responsible for the King Code, could issue a supplementary guidance paper on AI – similar to what it did for the issue of climate change [73]. The supplementary guidance contextualises climate change within South Africa's existing corporate governance requirements and environment and spells out governing bodies obligations and responsibilities. Such supplementary guidance would provide South African-specific guidance to local governing bodies. However, the industry and hybrid participants gave little indication that the AI industry gave much consideration to the King Code, even as it relates to general corporate governance requirements. This suggests that an AI-related update or addition to the code may not filter through into practice, at least not for non-listed, SMEs that are not obliged to adhere to the King Code. Whereas an update to existing corporate governance guidance is more likely to affect larger, stock exchange listed companies, who have a track record of implementing the guidance [74].

Moving now to the emergence of a variety of international approaches that touch on AI ethics, as identified in the 'international obligations' theme. AI is transnational in nature – a model can, for instance, be developed in one jurisdiction but exported and used in another. Additionally, many corporates scale their AI models to operate at a global level – recognisable examples are Google's search engine, OpenAI's generative AI applications, and social media feeds. Consequently, there should ideally be international standards and governance. There have been recent developments on this front. There is no African-centred AI approach, but there are several international-level efforts that are applicable to South Africa on a voluntary basis or may indirectly influence it. In the former category, is the UNESCO recommendations on AI that include an AI impact assessment [75]. In the latter camp, is the OECD AI Principles and the EU's AI Act, which observers have labelled the "the General Data Protection Regulation (GDPR) for AI" [63], [76]. The EU's proposed AI-focused legislation will almost certainly have an impact on South Africa [62], [77], [78]. A by-product of the GDPR was that customers beyond Europe become more empowered in how their data is collected and stored [78], [79]. It is not clear whether organisations give any consideration to these international efforts, and how they would translate these into practice unless they are formalised and codified into South African regulations or law. Notwithstanding, any South African domiciled organisation that wants to operate or serve customers in a foreign jurisdiction would need to account for the transnational requirements of that area.

5.2 Internal

The starting point for governing AI ethics is to acknowledge and understand the topic. In other words, organisations need to have 'awareness of ethics'. The overwhelming amount of literature on AI ethics often implicitly assumes this underlying assertion

– organisations need to recognise that AI ethics is worthy of time and resource commitments. There is some evidence that this is often not the case [21]. Artificial intelligence ethics cannot be another compliance tick box that is obfuscated and merely buried within broader and more generic processes and procedures. Or alternatively only dealt with in a reactive, crisis-born manner, which seems to be a typical approach [26]. Employees need to be aware of ethical issues in order to raise relevant concerns in a consistent, iterative manner [80]. Rank-and-file staff are unlikely to take AI ethics seriously if an organisation's executive leadership or governing body does not view it as important or is not cognisant of its scope and dynamics. Moreover, ethics cannot be seen merely as a technical problem with technical solutions, which obscures the socio-technical nature of AI. Rather, according to this view, AI needs to be understood and approached holistically and interdisciplinary, which is also a growing call in the literature[4],[41],[81]–[84].

Closely linked to awareness is the theme of 'diversity & informed staff'. This theme reenforces a reoccurring idea, that there is currently a gap in leaders' knowledge of AI and, consequently, there is a lack of governance on this front. At a governing body-level, it suggests that organisations need to incorporate expertise beyond the traditional general business management domain. A governing body could include a combination of more technically savvy and social science-orientated individuals. Alternatively, governing bodies and senior leadership can consult independent experts to advise them on AI and its ethical implications. Which are all calls that have also been made by others [85]. At an operational-level, a diverse workforce is more likely to be cognisant of the broader social-ethical impact of an organisation's output [27], [80]. This theme also echoes existing literature that calls for diverse AI workforces [28], [86]. However, the constraint to this is that the AI workforce globally tends to be predominantly males from a computer science or statistics background [20], [87]. Meaning that organisations may find it challenging to hire more diverse teams, due to a limited pool of diverse talent. Similarly, organisations, especially SMEs with constrained resources, would find it challenging to justify hiring non-technical staff in order to have a more representative, ethically orientated workforce.

There was a wide-spread awareness among participants of ethical frameworks and ethical codes from the Global North, as noted in the 'elaborate on existing frameworks' theme. None of the participants, however, indicated awareness of any local organisations that utilise these codes. Moreover, there is little evidence in the literature to suggest that the use of these frameworks or codes is widespread, either in South Africa or beyond [25], [88]–[91][1], [25], [32], [88], [90]–[92]. AI ethics frameworks and codes seem to primarily be in place among large US-based technology companies, for example Alphabet/Google, Microsoft, and IBM [93]–[95]. While the research did not explore the reasons for the lack of utilisation, there was no suggestion that it was because of a flaw or deficiency in these resources. Indeed, participants generally praised the quality of the frameworks and codes. However, shortcomings that make it challenging are that they tend to be either quite abstract – leaving questions of how-to operationalise it – or technically orientated and not accounting for AI's social impact [24], [86], [96]–[98]. Besides, the frameworks may not be ideally positioned for South Africa, given that it was created from a Global North perspective with different values and

cultural assumptions [19], [37]. Additionally, the absence of these frameworks and codes are possibly a function of how organisations see AI (i.e., it does not need special resources) and what they use it for. For instance, one participant did not see his organisation's use of AI as posing any noteworthy ethical risks. Furthermore, it could also be a function of an organisation's maturity level, with SMEs less likely to have a formal approach to ethics in general [99]. A company focused on survival is unlikely to adopt specialised frameworks or codes for AI. Another factor may be that organisations do not have the necessary expertise or resources to convert these to the local context. These codes and frameworks may have greater adaptation if concrete regulatory requirements are introduced, which would incentivise formal ethics governance.

The governance and management of AI ethics does not require the reinvention of the wheel. That is, an AI ethics structure does not necessarily need to be developed from scratch, but organisations can 'build onto existing structures', as noted by several authors [22], [70], [80], [100]. Indeed, AI ethics structures can be derived from and erected on top of an organisation's existing vision, mission, values, strategy, policies, structures, and workflows. Some participants' organisations were doing this in the sense that their AI work flowed strongly from their vision and raison d'etre – this while they did not have a formal AI strategy or approach. In other words, there was an alignment between their organisational purpose (i.e., values, mission, vision), customer value proposition, and utilisation of AI. This latter type of approach is probably more manageable and sustainable for smaller organisations that have an aligned workforce but lack the resources or capacity to have a formal ethics approach. A structured approach would be suited to larger, complex organisations. Furthermore, AI governance can be incorporated into existing corporate governance structures [22].

5.3 Differences Between South Africa, Global North

Based on the findings and discussion, a handful of notable differences between South Africa, as a representative of the Global South, and the Global North's approach vis-à-vis AI ethics can be noted in order to contrast the approaches and potentially serve as a basis for further exploration in future studies. Firstly, there appears to be little pressure or expectation on South African organisations to demonstrate a commitment to utilising AI in an ethical, stakeholder-centric manner. This deduction can be made based on several themes. In contrast, many organisations in the Global North, especially large technology companies, have to show some cognizance and sincerity towards using AI ethically. This difference may be due to the Global North having a longer track record of working on AI, and civil society and populations being more attuned to their rights in relation to digital products and services. For instance, South Africa has not had any AI-related public scandals, whereas there are regular controversies in the Global North. Secondly, related the previous assertion, there are more official constrains, regulations, and laws in the Global North on AI. For instance, the EU's efforts to regulate AI at a transnational level, and more than a dozen individual states in the US have passed legislation on AI. In contrast, in South Africa

there is no overt regulation and only limited legal frameworks that have nominal relevance to AI. Moreover, there is evidence of established, formal cooperative partnerships on AI in the Global North, where companies band together to advise and discuss ethical AI. There is little evidence of similar efforts in South Africa. Thirdly, the South African government's policy documents seem concerned with AI almost exclusively as an economic growth tool and fails to give much recognition of its sociotechnical nature. In contrast, the majority of national policy papers and strategies from Global North countries (in addition to a growing number of states from the Global South) incorporate elements on the responsible and ethical use of AI and its potential fall-out. Lastly, the Global North has produced a substantial number of codes, values, and frameworks on AI ethics produced by inter alia academia, technology companies, civil society, consultancies, international organisations, and think tanks [96]. Indeed, it could be postulated that there are an excess codes and frameworks. In contrast, there is a complete absence of this in South Africa, according to the findings. Local organisations would need to create their own or import/alter it from the Global North repository.

6 Limitations

The study has several limitations, primarily trade-offs related to the research strategy and methodology. Firstly, the qualitative design limits the study's transferability, albeit a common feature of studies focused on ethical issues [101], [102]. Secondly, the sample size is relatively small and limited to South Africa, which also narrows the transferability of the findings. It must be assumed that the results represent only a part of the overall AI ethics landscape in the country and the Global South more broadly. Furthermore, the sample is not necessarily representative of the diversity of the AI-related workforce. It cannot be ruled out that a different composition of participants would yield different results. Lastly, there is a risk of participants demonstrating self-reporting bias and social desirability bias [101], [103], [104].

7 Conclusion

The research identified several factors that influence the external and internal governance and management of AI ethics in South Africa, a representative of the Global South. It was determined that there are numerous outside factors that shape the environment in which organisations act and respond to ethics in relation to AI. Accordingly, organisational leaders can take a range of intra-company measures to address AI ethics. The article makes the following contributions to the literature and practice: (1) it adds a Global South/African perspective on mechanisms, methods, and measures to govern AI ethics. These findings, while being exclusively focused on one country, could be relevant to other countries in the Global South that share broadly similar features to South Africa; (2) it provides a counterweight to the dominance and predominant focus

on the Global North; (3) it relates the findings to existing literature and practice, especially to that of the Global North; and (4) it provides policymakers and organisations with empirical data to govern AI.

References

[1] A. Campolo, M. Sanfilippo, M. Whittaker, and K. Crawford, "AI Now 2017 Report," 2017.

[2] D. Fagella, "What is Artificial Intelligence? An Informed Definition," *EmerJ*, 2018. [Online]. Available: https://emerj.com/ai-glossary-terms/what-is-artificial-intelligence-an-informed-definition/. [Accessed: 19-Jun-2019].

[3] M. Whittaker *et al.*, "AI Now Report 2018," 2018.

[4] S. Larsson, M. Anneroth, A. Fellander, L. Fellander-Tsai, F. Heintz, and R. Cedering Angstorm, "Sustainable AI: An inventory of the state of knowledge of ethical, social, and legal challenges related to artificial intelligence," Lund, 2019.

[5] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," *Science (80-.).*, vol. 366, no. 6464, pp. 447–453, 2019.

[6] Z. Tufecki, "Machine intelligence makes human morals more important," *TED*, 2019. [Online]. Available: https://www.ted.com/talks/zeynep_tufekci_machine_intelligence_makes_human_morals_more_important. [Accessed: 24-Aug-2019].

[7] G. Burke, M. Mendoa, J. Linderman, and M. Tarm, "How AI-powered tech landed man in jail with scant evidence," *Associated Press*, 2021. [Online]. Available: https://apnews.com/article/artificial-intelligence-algorithm-technology-police-crime-7e3345485aa668c97606d4b54f9b6220. [Accessed: 01-Sep-2021].

[8] C. Q. Choi, "7 Revealing Ways AI Fails," *IEEE Spectrum*, Sep-2021.

[9] S. Ho and G. Burke, "An algorithm that screens for child neglect raises concerns," *Associated Press*, Apr-2022. [Online]. Available: https://apnews.com/article/child-welfare-algorithm-investigation-9497ee937e0053ad4144a86c68241ef1. [Accessed: 04-Jun-2022].

[10] R. Waelen, "The struggle for recognition in the age of facial recognition technology," *AI Ethics*, no. 0123456789, 2022.

[11] K. Roose, "An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy.," *The New York Times*, 02-Sep-2022.

[12] K. Huang, "Alarmed by A.I. Chatbots, Universities Start Revamping How They Teach," *The New York Times*, 16-Jan-2023.

[13] N. Tiku, "AI can now create any image in seconds, bringing wonder and danger," *The Washington Post*, 28-Sep-2022.

[14] K. Roose, "A Conversation With Bing's Chatbot Left Me Deeply Unsettled," *New York Times*, 16-Feb-2023.

[15] N. Tiku, G. De Vynck, and W. Oremus, "Big Tech was moving cautiously on AI. Then came ChatGPT.," *The Washington Post*, 27-Jan-2023.

[16] O. Bakiner, “What do academics say about artificial intelligence ethics? An overview of the scholarship,” *AI Ethics*, no. 0123456789, 2022.
[17] R. Dotan, “Global AI Ethics: Examples, Directory, and a Call to Action,” 2022.
[18] E. Ormond, “Global To Local: South African Perspectives on AI Ethics Risks,” in *The 3rd Southern African Conference for Artificial Intelligence Research*, 2022.
[19] C. Roche, P. J. Wall, and D. Lewis, “Ethics and diversity in artificial intelligence policies, strategies and initiatives,” *AI Ethics*, no. 0123456789, 2022.
[20] D. Zhang *et al.*, “2021 AI Index Report,” 2021.
[21] B. C. Stahl, J. Antoniou, M. Ryan, K. Macnish, and T. Jiya, “Organisational responses to the ethical issues of artificial intelligence,” *AI Soc.*, vol. 37, no. 1, pp. 23–37, 2022.
[22] M. Mäntymäki, M. Minkkinen, T. Birkstedt, and M. Viljanen, “Defining organizational AI governance,” *AI Ethics*, no. 0123456789, 2022.
[23] W. Orr and J. L. Davis, “Attributions of ethical responsibility by Artificial Intelligence practitioners,” *Inf. Commun. Soc.*, vol. 23, no. 5, pp. 719–735, 2020.
[24] E. Moss and J. Metcalf, “Ethics Owners: A New Model of Organizational Responsibility in Data-Driven Technology Companies,” 2020.
[25] J. Morley, L. Kinsey, A. Elhalal, F. Garcia, M. Ziosi, and L. Floridi, “Operationalising AI ethics: barriers, enablers and next steps,” *Ai Soc.*, no. Villarreal 2020, 2021.
[26] B. Rakova, J. Yang, H. Cramer, and R. Chowdhury, “Where Responsible AI meets Reality: Practitioner Perspectives on Enablers for Shifting Organizational Practices,” *Proc. ACM Human-Computer Interact.*, vol. 5, no. CSCW1, pp. 1–23, 2021.
[27] M. Ryan, E. Christodoulou, J. Antoniou, K. Iordanou, and M. Ryan, “An AI ethics ‘ David and Goliath ’: value conflicts between large tech companies and their employees,” *AI Soc.*, no. 0123456789, 2022.
[28] B. Chakravorti, A. Bhalla, R. S. Chaturvedi, and C. Filipovic, “50 Global Hubs for Top AI Talent,” *Harvard Business Review*, 2021. [Online]. Available: https://hbr.org/2021/12/50-global-hubs-for-top-ai-talent. [Accessed: 03-Jan-2022].
[29] J. McKendrick, “Nine Companies Are Shaping The Future Of Artificial Intelligence,” *Forbes*, 2019. [Online]. Available: https://www.forbes.com/sites/joemckendrick/2019/04/10/nine-companies-are-shaping-the-future-of-artificial-intelligence/#25b7184e2cf1. [Accessed: 25-May-2019].
[30] P. R. Daugherty, H. J. Wilson, and R. Chowdhury, “Using Artificial Intelligence to Promote Diversity,” *MIT Sloan Management Review*, 2018. [Online]. Available: https://mitsmr.com/2DQz2XT.
[31] F. Agrafioti, “How to Setup an AI R&D Lab,” *Harvard Business Review*, 2018. [Online]. Available: https://hbr.org/2018/11/how-to-set-up-an-ai-rd-lab.
[32] A. Winfield, “My top three policy and governance issues in AI/ML,” *Alan*

Winfield's Web Log, 2019. [Online]. Available: http://alanwinfield.blogspot.com/2019/05/my-top-three-policy-and-governance.html. [Accessed: 20-Jun-2019].

[33] C. M. Gevaert, M. Carman, B. Rosman, Y. Georgiadou, and R. Soden, "Fairness and accountability of AI in disaster risk management: Opportunities and challenges," *Patterns*, vol. 2, no. 11, p. 100363, 2021.

[34] M. Hickok, "Lessons learned from AI ethics principles for future actions," *AI Ethics*, no. 0123456789, 2020.

[35] C. Cath, S. Wachter, B. Mittelstadt, M. Taddeo, and L. Floridi, "Artificial Intelligence and the 'Good Society': the US, EU, and UK approach," *Sci. Eng. Ethics*, vol. 24, pp. 505–528, 2018.

[36] J. Alsever, C. Cooney, and M. Blake, "2022 Tech Trends Report: Artificial Intelligence," 2022.

[37] S. M. A. Kiemde and A. D. Kora, "Towards an ethics of AI in Africa: rule of education," *AI Ethics*, vol. 2, no. 1, pp. 35–40, 2022.

[38] D. O. Eke, K. Wakunuma, and S. Akintoye, "Introducing Responsible AI in Africa," in *Responsible AI in Africa*, D. O. Eke, K. Wakunuma, and S. Akintoye, Eds. Palgrave Macmillan, 2023, pp. 1–11.

[39] F. A. Raso, H. Hilligoss, V. Krishnamurthy, C. Bavitz, and L. Kim, "Artificial Intelligence & Human Rights : Opportunities & Risks," Boston, 2018.

[40] M. . Smith and S. Neupane, "Toward a research agenda Artificial intelligence and human development," 2018.

[41] M. Carman and B. Rosman, "Applying a principle of explicability to AI research in Africa: should we do it?," *Ethics Inf. Technol.*, vol. 23, no. 2, pp. 107–117, 2021.

[42] M. Madianou, "Nonhuman humanitarianism: when 'AI for good' can be harmful," *Inf. Commun. Soc.*, vol. 24, no. 6, pp. 850–868, 2021.

[43] A. Gwagwa, E. Kraemer-Mbula, N. Rizk, I. Rutenberg, and J. De Beer, "Artificial Intelligence (AI) Deployments in Africa: Benefits, Challenges and Policy Dimensions," *African J. Inf. Commun.*, no. 26, pp. 1–28, 2020.

[44] S. Sedola, A. J. Pescino, and T. Greene, "Artificial Intelligence for Africa," 2021.

[45] Ipsos, "Global Opinions and Expectations About Artificial Intelligence," 2022.

[46] R. L. Sims, A. E. Gegez, and L. Popova, "Attitudes towards business ethics: A five nation comparative study," *J. Bus. Ethics*, vol. 50, no. 3, pp. 253–265, 2004.

[47] B. Scholtens and L. Dam, "Cultural Values and International Differences in Business Ethics," *J. Bus. Ethics*, vol. 75, pp. 273–284, 2007.

[48] S. T. Segun, "Critically engaging the ethics of AI for a global audience," *Ethics Inf. Technol.*, vol. 23, no. 2, pp. 99–105, 2021.

[49] M. Saunders, P. Lewis, and A. Thornhill, *Research Methods for Business Students*, 8th ed. Pearson Education, 2019.

[50] F. Pasquale, "Odd Numbers," *Real Life*, 2018. [Online]. Available: https://reallifemag.com/odd-numbers/. [Accessed: 19-Mar-2019].

[51] T. Ferretti, "An Institutionalist Approach to AI Ethics: Justifying the Priority

of Government Regulation over Self-Regulation," *Moral Philos. Polit.*, 2021.
[52] M. Ryan and B. C. Stahl, "Artificial intelligence ethics guidelines for developers and users: clarifying their content and normative implications," *J. Information, Commun. Ethics Soc.*, vol. 19, no. 1, pp. 61–86, 2021.
[53] B. Smith, "Facial recognition : It 's time for action," *Microsoft*, 2018. [Online]. Available: https://blogs.microsoft.com/on-the-issues/2018/12/06/facial-recognition-its-time-for-action/?ranMID=24542&ranEAID=je6NUbpObpQ&ranSiteID=je6NUbpObpQ-AlSgAi22jukIDhg4pFcWfA&epi=je6NUbpObpQ-AlSgAi22jukIDhg4pFcWfA&irgwc=1&OCID=AID681541_aff_7593_1243925&tduid.
[54] P. Aghion, A. Bergeaud, and J. van Reenen, "The Impact of Regulation on Innovation," 2021.
[55] G. Banavar, "What It Will Take for Us to Trust AI," *Harvard Business Review*, 2016. [Online]. Available: https://hbr.org/2016/11/what-it-will-take-for-us-to-trust-ai. [Accessed: 16-Mar-2019].
[56] South African Government, "Report of the Presidential Commission on the 4th Industrial Revolution," 2020.
[57] South African Government, "ICT and Digital Economy Masterplan for South Africa Draft for discussion (DRAFT)," 2020.
[58] Department of Communications and Digital Technologies, "Draft National Policy on Data and Cloud," Pretoria, 2021.
[59] A. Vats and N. Natarajan, "G20. AI National Strategies, Global Ambitions," 2022.
[60] ALT Advisory, "AI Governance In Africa," 2022.
[61] J. Steyn, "SA lags several African countries on AI policy," *Business Day*, Johannesburg, 22-Jun-2022.
[62] R. Adams, "AI in Africa: Key Concerns and Policy Considerations for the Future of the Continent," 2022.
[63] European Union Commission, "Proposal for a Regulation Of The European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence Ameding Certain Union Legislative Acts COM/2021/206 final," 2021.
[64] United States Government Accountability Office, "FACIAL RECOGNITION TECHNOLOGY: Privacy and Accuracy Issues Related to Commercial Uses," 2020.
[65] M. & S. Department of Digital, Culture and D. Collins, "UK sets out proposals for new AI rulebook to unleash innovation and boost public trust in the technology," *United Kingdom Goverment*, 2022. [Online]. Available: https://www.gov.uk/government/news/uk-sets-out-proposals-for-new-ai-rulebook-to-unleash-innovation-and-boost-public-trust-in-the-technology. [Accessed: 26-Jul-2022].
[66] The Office of Science and Technology Policy, "Blueprint For An AI Bill of Rights," Washington D.C., 2022.
[67] R. Y. Wong, M. A. Madaio, and N. Merrill, "Seeing Like a Toolkit: How

Toolkits Envision the Work of AI Ethics," vol. 1, no. 1, pp. 1–21, 2022.
[68] I. Munoko, H. L. Brown-Liburd, and M. Vasarhelyi, "The Ethical Implications of Using Artificial Intelligence in Auditing," *J. Bus. Ethics*, 2020.
[69] P. Tambe, P. Cappelli, and V. Yakubovich, "Artificial intelligence in human resources management: Challenges and A path forward," *Calif. Manage. Rev.*, vol. 61, no. 4, pp. 15–42, 2019.
[70] R. Blackman, "A Practical Guide to Building Ethical AI," *Harv. Bus. Rev.*, 2020.
[71] M. Schaake, "European commission's Artificial Intelligence Act," 2021.
[72] Institute of Directors South Africa, "King IV: Report on Corporate Governance for South Africa 2016," 2016.
[73] Institute of Directors South Africa, "Guidance paper: Responsiblities of Governing Bodies in Responding to Climate Change," 2021.
[74] M. Mpinganjira, M. Roberts-Lombard, G. Svensson, and G. Wood, "Measurement properties of the construct of the code of ethics content: The South African experience," *South African J. Bus. Manag.*, vol. 49, no. 1, pp. 1–8, 2018.
[75] UNESCO, "Recommendation on the ethics of artificial intelligence," *UNESCO*, 2021. [Online]. Available: https://en.unesco.org/artificial-intelligence/ethics#recommendation. [Accessed: 18-Dec-2021].
[76] OECD, "Recommendation of the Council on Artificial Intelligence," 2019.
[77] A. Engler, "The EU AI Act Will Have Global Impact, but a Limited Brussels Effect," *Brookings Institute*, 2022. [Online]. Available: https://www.brookings.edu/research/the-eu-ai-act-will-have-global-impact-but-a-limited-brussels-effect/. [Accessed: 26-Jul-2022].
[78] C. Siegmann and M. Anderljung, "The Brussels Effect and Artificial Intelligence : How EU regulation will impact the global AI market," Oxford, 2022.
[79] A. Petrova, "The impact of the GDPR outside of the EU," *Lexology*, 2019. [Online]. Available: https://www.lexology.com/library/detail.aspx?g=872b3db5-45d3-4ba3-bda4-3166a075d02f. [Accessed: 13-Jul-2022].
[80] R. Eitel-Porter, "Beyond the promise: implementing ethical AI," *AI Ethics*, vol. 1, no. 1, pp. 73–80, 2021.
[81] M. Coeckelbergh, "Technology Regulation Ethics of artificial intelligence : Some ethical issues and regulatory challenges," *Technol. Regul.*, pp. 31–34, 2019.
[82] L. Bartolo and R. Thomas, "Qualitative humanities research is crucial to AI," *fast.ai*, 2022. [Online]. Available: https://www.fast.ai/2022/06/01/qualitative/. [Accessed: 14-Jun-2022].
[83] L. Weinberg, "Rethinking Fairness: An Interdisciplinary Survey of Critiques of Hegemonic ML Fairness Approaches," *J. Artif. Intell. Res.*, vol. 74, pp. 75–109, 2022.
[84] E. Drage and K. Mackereth, "Does AI Debias Recruitment? Race, Gender, and AI's 'Eradication of Difference,'" *Philos. Technol.*, vol. 35, no. 4, pp. 1–25,

2022.

[85] M. Galligan, V. Katyal, M. Mohlenkamp, C. Parry, and C. Adkins, “AI ethics A new imperative for businesses , boards , and C-suites,” 2019.

[86] M. Ryan, J. Antoniou, L. Brooks, T. Jiya, K. Macnish, and B. Stahl, “Research and Practice of AI Ethics: A Case Study Approach Juxtaposing Academic Discourse with Organisational Reality,” *Sci. Eng. Ethics*, vol. 27, no. 2, pp. 1–29, 2021.

[87] IBM, “Responsibility for AI Ethics Shifts from Tech Silo to Broader Executive Champions, says IBM Study,” *IBM*, 2022. [Online]. Available: https://newsroom.ibm.com/2022-04-14-Responsibility-for-AI-Ethics-Shifts-from-Tech-Silo-to-Broader-Executive-Champions,-says-IBM-Study. [Accessed: 11-Aug-2022].

[88] J. Morley, L. Floridi, L. Kinsey, and A. Elhalal, “From What to How: An Overview of AI Ethics Tools, Methods and Research to Translate Principles into Practices,” 2019.

[89] A. Winfield, “An Updated Round Up of Ethical Principles of Robotics and AI,” *Alan Winfield’s Web Log*, 2019. [Online]. Available: http://alanwinfield.blogspot.com/2019/04/an-updated-round-up-of-ethical.html. [Accessed: 20-Jun-2019].

[90] J. Fjeld, N. Achten, H. Hilligoss, A. C. Nagy, and M. Srikumar, “Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI,” 2020.

[91] D. Baker and A. Hanna, “AI Ethics Are in Danger. Funding Independent Research Could Help,” *Stanford Soc. Innov. Rev.*, 2022.

[92] A. Winfield and M. Jirotka, “Ethical governance is essential to building trust in robotics and AI systems,” *Philos. Trans. A Math. Phys. Eng. Sci.*, vol. 376, no. 2133, p. 19, 2018.

[93] J. Perez, “IBM and Microsoft Have Integrated AI Ethical Standards into Their Operations, So Can You,” *IEEE*, 2021. [Online]. Available: https://spectrum.ieee.org/the-institute/ieee-products-services/ibm-and-microsoft-have-integrated-ai-ethical-standards-into-their-operations-so-can-you.

[94] B. Green, D. Lim, and E. Ratte, “Responsible Use of Technology_The Microsoft Case Study,” 2021.

[95] H. Field, “How Microsoft and Google use AI red teams to ‘stress test’ their systems,” *Emerging Tech Brew*, Jun-2022.

[96] B. Attard-Frost, A. De los Ríos, and D. R. Walters, “The ethics of AI business practices: a review of 47 AI ethics guidelines,” *AI Ethics*, no. 0123456789, 2022.

[97] D. Greene, A. L. Hoffmann, and L. Stark, “Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning,” in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019, p. 10.

[98] E. Prem, “From ethical AI frameworks to tools : a review of approaches,” *AI Ethics*, no. 0123456789, 2023.

[99] I. van Wyk and P. Venter, “Perspectives on business ethics in South African small and medium enterprises,” *African J. Bus. Ethics*, vol. 16, no. 1, pp. 81–104, 2022.

[100] J. Mökander and L. Floridi, “Operationalising AI governance through ethics-based auditing: an industry case study,” *AI Ethics*, no. 0123456789, 2022.

[101] P. Grant, S. Arjoon, and P. McGhee, “In Pursuit of Eudaimonia: How Virtue Ethics Captures the Self-Understandings and Roles of Corporate Directors,” *J. Bus. Ethics*, vol. 153, no. 2, pp. 389–406, 2018.

[102] D. Rossouw, *Developing Business Ethics as an Academic Field*. Johannesburg: RAU, 2004.

[103] A. Crane, “Are You Ethical? Please Tick Yes Or No: On Researching Ethics in Business Organizations,” *J. Bus. Ethics*, vol. 20, pp. 237–248, 1999.

[104] D. M. Randall and A. M. Gibson, “Methodology in Business Ethics Research: A Review and Critical Assessment,” *J. Bus. Ethics*, vol. 9, pp. 457–471, 1990.