**South African Journal of Information Management**
ISSN: (Online) 1560-683X, (Print) 2078-1865

# Examining the applicability of the Protection of Personal Information Act in AI-driven environments

**Authors:**
Vicent Mbonye[1]
Marlini Moodley[2,3]
Farai Nyika[2]

**Affiliations:**
[1]Academic, School of Education, MANCOSA, Durban, South Africa.

[2]Academic, MANCOSA, Durban, South Africa

[3]Department of Management Sciences, Faculty of Marketing and Retail, Durban University of Technology, Durban, South Africa

**Corresponding author:**
Vicent Mbonye,
vicentmbonye@gmail.com

**Dates:**
Received: 24 Nov. 2023
Accepted: 28 May 2024
Published: 30 Aug 2024

**How to cite this article:**
Mbonye, V., Moodley, M. & Nyika, F., 2024, 'Examining the applicability of the Protection of Personal Information Act in AI-driven environments', *South African Journal of Information Management* 26(1), a1808. https://doi.org/10.4102/sajim.v26i1.1808

**Read online:**
Scan this QR code with your smart phone or mobile device to read online.

**Background:** Technological advancements have heightened the importance of safeguarding individual privacy and data. In response to these challenges, South Africa introduced the *Protection of Personal Information (POPI) Act*. This legislation established robust legal frameworks aimed at protecting confidential information and upholding individuals' right to anonymity. However, there is a significant research gap regarding the *POPI Act's* direct implications and effectiveness in the context of artificial intelligence (AI) adoption and utilisation. Understanding the interplay between the *POPI Act* and AI technologies is crucial for ensuring regulatory compliance, safeguarding personal data and fostering responsible AI deployment in South Africa.

**Objectives:** This study investigates the *POPI Act's* applicability in addressing privacy issues related to AI adoption in various sectors.

**Method:** The research uses a document review methodology to analyse the documents and synthesise the results. This approach offers efficiency, accessibility, cost-effectiveness and non-intrusiveness benefits, making it a valuable tool for qualitative research across various disciplines.

**Results:** Despite the *POPI Act's* guiding principles aligning with key concepts of personal information protection, there are several gaps in its applicability to AI advancements across various sectors.

**Conclusion:** The study emphasises the need for a dynamic legal framework that evolves with AI advancements, advocating for the incorporation of more stringent measures to address emerging privacy concerns.

**Contribution:** The research contributes to the ongoing discourse on data protection and AI by highlighting the need for a forward-thinking legal framework that balances innovation and privacy, ensuring that the *POPI Act* remains effective in the face of evolving technologies.

**Keywords:** artificial intelligence; data privacy; personal information; *Protection of Personal Information Act (POPI Act)*; data protection.

## Introduction

Artificial intelligence (AI) has gained widespread use in today's digital environment, revolutionising business operations and client interactions (Haleem et al. 2022). Moreover, several industries, such as healthcare, finance, and marketing, have adopted AI technologies to enhance efficiency and effectiveness (Dwivedi et al. 2021). Despite AI's progress in many industries, there are valid privacy concerns associated with its widespread use (Zheng & Cai 2020). These concerns include losing control over personally identifiable information, unauthorised data access, consent issues, and potential biases (Bharadiya 2023; Bleier, Goldfarb & Tucker 2020; Sallam 2023; Shah et al. 2020). The increasing sophistication and autonomy of AI systems exacerbate these worries, necessitating stricter regulation and control (Buiten 2019; Gianni, Lehtinen & Nieminen 2022; Wirtz, Weyerer & Sturm 2020). Thus, there is a need for tighter regulation and oversight because of the growing autonomy and sophistication of AI systems (Buiten 2019; Gianni et al. 2022; Wirtz et al. 2020). Because of the rapid evolution of technology, privacy and data security have become increasingly important over the years. The South African government responded by enacting the *Protection of Personal Information (POPI) Act*, a comprehensive law that aims to safeguard personal information and individual privacy (Adams et al. 2021; Ramcharan 2020).

The *POPI Act* was established to guide data protection and privacy, including processing restrictions, clarification of the purpose, quality information, participation of data subjects, and responsibility, which can serve as a foundation for organisations that handle many sensitive data types, including AI-generated ones (Ramcharan 2020). However, other parts of smart technology adoption regulation are patchy and poorly positioned to handle some of the most important legal and ethical concerns that have arisen because of this (Townsend & Botes, 2023).

While the *POPI Act* provides a strong foundation for ensuring personal information privacy in the digital age, the literature has not sufficiently examined its effectiveness in resolving the privacy issues arising from the use of AI. Moreover, South Africa presently does not have any explicit legislation, regulations, or official policies that provide guidance on the ethical use of AI (Adams 2021; Ka Mtuze & Morige 2024). This gap is exacerbated by a lack of research on the application of AI technologies across different sectors. Hence, the goal of this study is to investigate ethical issues associated with the integration of AI across multiple sectors. The aim is to evaluate the effectiveness of the principles stated in the *POPI Act* in addressing these ethical difficulties. The objective is to provide significant insights into the existing conversations on AI ethics and legal frameworks in South Africa.

## Overview of South Africa's *Protection of Personal Information Act* (2013)

By outlining standards for how both government and businesses can handle citizens' private data, South Africa's *POPI Act of 2013* aims to protect people's privacy (Da Veiga & Ophoff 2020; Naidoo 2021). This Act, which became effective in 2019, aims to secure the regulation of personal information in compliance with international norms, supporting cross-border collaboration (Ramcharan 2020). The United States (US), Canada (Canada), Australia, and the United Kingdom (UK) have all implemented data privacy regimes that served as inspiration for the *POPI Act* (Ramcharan 2020).

### A person's right to privacy and personal information is defined

According to the *POPI Act*, an individual's right to privacy is defined as the freedom to conduct oneself without interference with one's private affairs (Government of South Africa 2013; Staunton et al. 2020). Individual identifiers fall under the umbrella term 'personal information'.

### De-identify

The *POPI Act* protects the personal information of individuals known as data subjects. The Act also states that individuals have the right to control how their personal data is used by organisations. Among these protections is the option to have one's data updated or deleted (Naidoo 2021; Swales 2021). The Act uses the term 'de-identification' to describe the process of erasing information that could potentially identify an individual.

## *Protection of Personal Information Act* application principles

The board of directors, branches, business units, and divisions of the responsible entity are all bound by the policy and its guiding principles. The policy extends to everyone associated with the company, including workers, vendors, contractors, volunteers, and representatives (Government of South Africa 2013; Staunton et al. 2020). However, processing personal information as part of routine household operations is exempt from the *POPI Act*.

### Accountability (*Protection of Personal Information Act* principle 1)

According to this principle, the entity utilising personal information is accountable for maintaining privacy (Adams et al. 2021). Organisations are obligated to adhere to the guiding principles outlined in the act to promote responsible behaviour. The *POPI Act* particularly highlights the importance of addressing compliance risks linked to personal information protection. This principle proposes that organisations may designate an Information Officer who is responsible for overseeing the organisation's compliance with *POPI Act* and acting as a point of contact for data subjects and the Information Regulator (Government of South Africa 2013).

### Processing limitation (*Protection of Personal Information Act* principle 2)

This concept demands that businesses process personal information in a manner that is fair, lawful, and not excessive, with the data subject's consent (Government of South Africa 2013, Naidoo 2021, Netshakhuma 2020, Thaldar & Townsend 2021). Businesses should only acquire personal information for legitimate, specified, and necessary purposes.

### Purpose specification (*Protection of Personal Information Act* principle 3)

The idea stresses the importance of being open and honest about data processing. It stresses that organisations should only acquire and use personal data for specified, explicit purposes (Adams et al. 2021; Government of South Africa 2013; Staunton et al. 2020).

### The *Protection of Personal Information Act* principle 4 limits further processing

This principle reinforces the idea that organisations shouldn't process personal information for purposes that diverge from the original intent without additional consent. It draws attention to the need for organisations to obtain consent before repurposing personal data (Adams et al. 2021; Government of South Africa 2013; Netshakhuma 2020; Staunton et al. 2020; Thaldar & Townsend 2021).

### Information quality (*Protection of Personal Information Act* principle 5)

The principle compels companies to guarantee that the gathered personal information is accurate, full, and not misleading. We include data from outside sources here. The Act (Adams et al. 2021; Government of South Africa 2013; Mbonye, Subramaniam & Padayachee 2021) gives special attention to important data, such as the beneficiary details of a life insurance policy.

### Openness (*Protection of Personal Information Act* principle 6)

Under this tenet, businesses must give individuals notice before collecting or using their data. It stresses the importance of designating points of contact for questions and requests from data subjects. Data subjects can lodge complaints about the processing of their personal data, request access to their data, and request updates or corrections (Adams et al. 2021; Government of South Africa 2013; Mbonye et al. 2021).

### Security safeguards (*Protection of Personal Information Act* principle 7)

To prevent data loss, unauthorised access, or other security breaches, it is crucial to adhere to this principle. The environment in which we use security controls dictates the need for more stringent controls to safeguard sensitive data (Adams et al. 2021; Government of South Africa 2013; Mbonye et al. 2021).

### The *Protection of Personal Information Act* principle 8 involves the participation of data subjects

According to the Government of South Africa (2013), this concept ensures that individuals have access to and control over their data.

The principle recommends that organisations appoint Information Officers (or Deputy Information Officers) to ensure compliance with the privacy policy (Adams et al. 2021; Government of South Africa 2013; Mbonye et al. 2021).

## Methodology

The study used a structured document review methodology to make sure that all relevant literature and government documents about AI and data privacy in South Africa's *POPI Act* were looked at in depth and critically. The process involved several key steps:

### Document retrieval

In the first stage, we conducted a thorough search for publications in well-established databases such as Google Scholar and EBSCOhost. We used precise terms that directly address privacy concerns in the AI field, as well as the *POPI Act* in South Africa. The researchers additionally conducted searches on online government repositories and other sources to compile a complete collection of articles pertaining to the safeguarding of personal information. We limited the article search to articles published between 2019 and 2023. The government materials needed to prioritise a South African context, ensuring that they are relevant to the latest developments in AI and data protection regulations in South Africa.

### Criteria for selection

We selected documents for further investigation based on specific inclusion criteria. The criteria encompassed aspects of topic relevancy, publication in the English language, and concentration on AI applications across several sectors. The government records needed to be specific to South Africa, in accordance with the study's scope. The analysis included only peer-reviewed research articles and government documents, eliminating non-English publications and items that did not pertain to the chosen topics. We conducted the document evaluation procedure according to predetermined criteria to ensure the genuineness, reliability, inclusivity, and significance of the selected documents (Dunne, Pettigrew & Robinson 2016; Flick 2018; Kridel 2015; Mogalakwe 2009).

The selection of documents was based on their genuineness and reliability. We prioritised primary sources to ensure the validity of the information, adhering to Mogalakwe's (2009) emphasis on the fundamental genuineness of research. We exhaustively evaluated criteria such as authorship, publication date, and source reliability to carefully assess the legitimacy and trustworthiness of the selected documents (Dunne et al. 2016; Flick 2018).

We conducted an evaluation to ascertain how accurately and comprehensively the chosen literature represented the research topic. We made a deliberate effort to avoid including content that is unique to a certain individual and instead focussed on selecting items that reflect the wider discussion on the topic. When considering the selection of documents, factors that affect how representative they are, such as the condition of the documents and any constraints on accessing them, were considered based on Kridel's (2015) observations about the problems of document selection.

We thoroughly analysed the documents' content to determine its meaning and importance. We analysed the context and relevance of the content in the selected books by considering both the literal and interpretive meanings. Mogalakwe's (2009) emphasis on establishing a connection between the literal meaning and the document's creation environment aligns with this approach.

We employed a purposive sample strategy to select articles that most effectively addressed the research questions and objectives. The selection procedure entailed a methodical and repetitive process of screening and evaluation, guided by theoretical considerations and thematic significance. This process followed the sampling techniques for document analysis recommended by Bowen (2009) and Flick (2018).

### Procedure for evaluating and extracting data

The screening process entailed a methodical examination of each document to extract crucial information that was pertinent to the study objectives. The document included topics such as the implementation of AI in different industries, measures to preserve data privacy, adherence to the regulatory requirements of the *POPI Act*, and suggestions for improving data security in AI-powered settings.

### Data synthesis and analysis

We subjected the selected materials to thematic analysis to identify recurring themes, trends, gaps, problems, and potential areas for improvement within the scope of the *POPI Act*. The investigation centred on comprehending the interaction between AI technology and data protection legislation, evaluating the efficacy of existing legal restrictions, and examining the consequences for stakeholders in different industries. The chosen methodology for theme analysis was based on Braun and Clarke's (2013) flexible method, which ensured a methodical and thorough assessment of the content of the document.

### Presenting the findings

The process of data synthesis consisted of three distinct stages: firstly, examination of the data through theme analysis; secondly, presentation of the discovered patterns and conclusions; and lastly, a comprehensive discussion of the outcomes and their potential consequences. The study utilised a narrative synthesis approach to assemble and categorise the primary themes through thematic analysis, following the framework proposed by Braun and Clarke. Through deliberation, the study team improved codes and developed distinct themes. The discussion section documented the findings.

### Limitations

While the methodology employed in this academic writing endeavour is designed to rigorously investigate the integration of AI in various sectors and its alignment with data privacy protection under the *POPI Act*, there are several limitations that warrant acknowledgement and consideration. The study's reliance on established databases such as Google Scholar and EBSCOhost may overlook valuable insights from alternative sources, potentially limiting the breadth of the literature review. Furthermore, the exclusion of non-English publications introduces a language bias that may obscure important contributions from non-English-speaking researchers. The focus on articles published between 2019 and 2023 could overlook foundational works and emerging trends beyond this timeframe, affecting the currency and completeness of the analysis. Finally, purposive sampling may limit the generalisability of the results. Transparently acknowledging these limitations is crucial for interpreting the study's outcomes accurately and understanding its scope.

## Literature review

### The use of artificial intelligence in different sectors

#### Artificial intelligence and the healthcare sector

Recent advances in AI tools for healthcare have the potential to increase healthcare delivery's efficacy and efficiency (Chew & Achananuparp 2022). There is a growing adoption of AI in the health sector. People experiencing common symptoms have adopted AI chatbots to provide direct health advice (Li 2023; Matulis & McCoy 2023; Sun, Gupta & Sharma 2022). Remote patient monitoring and digital health coaching can both benefit from AI's ability to efficiently manage large datasets generated by smart devices and extract relevant clinical insights. Using this data-driven method, AI-powered coaches can help patients self-manage disorders such as diabetes, obesity, hypertension, and depression (Lin 2022). Clinical Decision-Making AI Assistants can aid primary care physicians (PCPs) in making more informed decisions at the point of treatment (Lin 2022). Additionally, clinical decision-making AI assistants embedded in electronic health record (EHR) systems support PCPs in making well-informed clinical decisions at the point of care (Lin 2022; Abbasgholizadeh Rahimi et al. 2022). In addition, practice management AI solutions automate administrative chores such as billing, coding, and prior authorisations, allowing healthcare professionals to devote their time to more meaningful work (Firouzi et al. 2020; Keshta 2022).

#### Artificial intelligence and the finance sector

Artificial intelligence has facilitated the efficient analysis of large amounts of financial data to better inform strategic decisions and encourage the development of novel financial products (Cao 2022). According to Thowfeek, Samsudeen and Sanjeetha (2020), adopting AI will give the banking industry a competitive edge by enhancing operational efficiency and, consequently, profits.

Financial institutions can now examine large financial databases, leading to improved trading techniques and more well-informed investment decisions (Aziz & Andriansyah 2023; Cao 2022). Furthermore, AI enables financial institutions to anticipate future risks and proactively implement preventive measures. Real-time analysis of various data sources, integrated into risk management, achieves this (Cao 2022). Chatbots and virtual assistants powered by AI have changed customer service by making tailored help available 24/7. This not only boosts customer experience and engagement but also fosters customer happiness and loyalty (Cao 2022). Moreover, AI systems that are good at spotting unusual financial transactions have helped in the fight against fraud, protecting both financial institutions and their customers (Cao 2022). In addition, AI's ability to automate mundane jobs such as data entry and financial computations improves operational efficiency and frees up human workers to focus on more strategic endeavours (Cao 2022).

### Marketing and artificial intelligence

The integration of AI into marketing holds tremendous promise. Artificial intelligence solutions can boost customer happiness and loyalty by analysing customer behaviour and preferences to provide recommendations and products that are more likely to appeal to the individual (Cao 2022). Artificial intelligence has great potential to revolutionise the marketing industry by increasing the availability of data sources, expanding data management skills in software, and paving the way for the development of complex and cutting-edge algorithms. Studies such as Haleem et al. (2022) show how this revolutionary technology is changing the dynamic between brands and their consumers. Based on the data collected and created by AI algorithms, marketers can now prioritise consumer demands in real-time, using AI to quickly decide the most appropriate content to target customers and the ideal channel to utilise. Hence, users are more likely to make purchases after being exposed to customised experiences powered by AI (Haleem et al. 2022; Nagy & Hajdú 2021). Additionally, AI technologies are skilled at assessing the success of rival campaigns and identifying customer expectations, providing invaluable insights for marketing plans (Haleem et al. 2022). Artificial intelligence helps businesses analyse client data, determine preferences, and forecast behaviour, all of which leads to AI-based recommendations that boost the efficiency of marketing campaigns (Milan, Sahu & Sandhu 2023; Yaiprasert & Hidayanto 2023).

### Artificial intelligence and education

Artificial intelligence-driven programmes in education provide useful feedback to both students and teachers, giving them more agency over their learning (Singh & Mishra 2021). There are many upsides to adopting this AI technology. Notably, it facilitates the assessment of academic progress by providing detailed evaluations of programme efficacy (Arora 2021). Artificial intelligence in education plays a critical role in building a smart campus by integrating into institutional systems. Using AI, campus administrators can more easily keep tabs on things such as attendance, facility usage, parking lot management, alarm activation, room utilisation, and temperature control. These intelligent AI technologies thus improve institutional management (Arora 2021). This benefits students, faculty, and staff.

There is a fast advancement in the rate of technological growth and adoption (Perrault et al. 2019). While Moore's Law states that processing power would double every 2 years, AI development, especially with machine learning (ML) and deep learning (DL), has outpaced this rate (Perrault et al. 2019). For example, AI's skills double every 3–4 months. The development of chess-playing AI is a striking illustration of this trend. Modern AI, such as AlphaZero, which learned the game from the start using DL, outperformed traditional AI, such as Stockfish, which relied on already existing information. AlphaZero not only outperformed Stockfish in a matter of hours but also showed off its incredible ability to quickly learn and master chess concepts. AlphaZero won 28 games, lost none, and tied for 72 others against Stockfish in a timed 100-game tournament (Silver et al. 2018). While there are certainly advantages of using AI in a variety of settings, there are also serious privacy problems that should be carefully considered and regulated.

### The privacy concerns of artificial intelligence

Many privacy concerns arise with the widespread adoption of AI across industries. There are serious privacy problems associated with using AI in the healthcare industry. Patients' records become increasingly vulnerable to breaches when AI handles sensitive medical data (Murdoch 2021; Rahman et al. 2023). In addition, the deployment of AI-driven chatbots and virtual assistants raises worries regarding the confidentiality of patient-provider interactions and the potential exposure of personal health information. Furthermore, there are privacy concerns associated with the use of AI in medical monitoring devices and sensors (Li 2023; Matulis & McCoy 2023; Sun et al. 2022).

The use of AI in the financial industry raises similar privacy concerns, particularly regarding the safety of financial transactions and individual data (Truby, Brown & Dahdal 2020). Awotunde et al. 2021, have raised concerns about the accidental access of private financial records by AI algorithms intended for fraud detection, and the potential revelation of people's financial habits through data collected for risk assessment and investment strategies. There are privacy concerns about the content and context of conversations that involve AI-powered chatbots and their human counterparts.

Marketing AI's role in data analysis and personalised suggestions has justified concerns about consumer privacy (Bleier et al. 2020; Shah et al. 2020). Consumers' privacy could be at risk if AI algorithms were to collect and store detailed information about their habits and preferences (Davenport et al. 2020; Kaličanin et al., 2019). The convergence of many sources of real-time consumer data exacerbates the risk of privacy breaches (Huang & Rust 2021).

The significant data gathering on students' learning behaviours, preferences, and progress by AI-driven platforms in the education industry raises privacy concerns (Köbis & Mehner 2021; Kooli 2023). Data breaches, unauthorised access, and inappropriate use of such data for unexpected purposes bring the protection of student data to the forefront (Kasneci et al. 2023).

## Discussion of findings

The integration of AI into various sectors raises significant ethical concerns about privacy. While the South African *POPI Act* aims to safeguard personal information, it faces certain limitations in addressing these concerns within the context of AI. Next, we discuss how each principle of the *POPI Act* intersects with AI ethics and propose recommendations to bridge the gaps.

## Accountability (*Protection of Personal Information Act* principle 1)

The appointment of an information officer can address concerns about ethics in AI development. When hired, the person in charge of AI ethics should have a clear mandate to define, implement, and enforce the organisation's AI ethics policy. This person's watchful eye can ensure transparency, reduce bias, and make well-considered decisions. This is because businesses may not prioritise openness in algorithmic decision-making without clear instructions, leading to a lack of knowledge and understanding among stakeholders about data collection, analysis, and use (Aysolmaz, Müller & Meacham 2023). However, the *POPI Act* does not obligate entities to designate someone as the 'Information Officer'. The Act does not legally require businesses to have a designated 'Information Officer', despite its recommendation. In the context of AI, where ethical considerations play a key role in data processing, this position is essential for ensuring compliance with the *POPI Act*.

South Africa has not adopted an AI legislative framework to further enhance the protection of personal information in the context of AI and the *POPI Act* does not contain any explicit AI rules.

## The *Protection of Personal Information Act* principles 2 and 4 address processing limitation

This article suggests that principles 2 and 4 of the *POPI Act* may not go far enough in addressing the challenges of AI. This is because entities may have limited control over how AI uses the information it collects. Artificial Intelligence algorithms, for instance, re-construct data as they learn, meaning that the organisation may not oversee the modifications made to this data. Individuals may find it challenging to comprehend the processing of data within these algorithms, thereby complicating the process of informed consent. Unwanted outcomes may arise when individuals use personal information for purposes beyond its original collection without first considering the ethical implications.

## Purpose specification (*Protection of Personal Information Act* principle 3)

The principle may not explicitly address the potential risks associated with deploying AI algorithms in data processing, but it does entail alerting data subjects about the aim of gathering their personal information. Complex AI algorithms are possible (Ingrams, Kaufmann & Jacobs 2022; Kuziemski & Misuraca 2020). This could result in individuals being oblivious to the processing of their data and the potential repercussions. People may not be able to make educated judgements or give meaningful consent, for instance, if they do not know how their data are being used to develop predictive models. The main difficulty here is that AI applications may exploit the data in unexpected ways when trying to find novel answers to a problem. Consequently, the signed individual purpose specification undergoes alterations.

## Information quality (*Protection of Personal Information Act* principle 5)

While the concept highlights the responsible party's duty to ensure correct and up-to-date personal information, it may not expressly address the issues of preserving data quality in the context of AI. It might be difficult to strike a balance between data privacy and accuracy. People may be less likely to volunteer complete or accurate information if they fear misuse of their personal information. Given the importance of data to the success of AI solutions, this poses a challenge in the context of AI innovation management. For deep learning applications, lengthy and reliable datasets are often required, but they might be challenging to acquire for smaller businesses (Prem 2019). Unfair treatment, implicit bias, and a general sense of injustice may result from the use of algorithms in decision-making (Köchling & Wehner 2020; Köchling et al. 2021; Marcinkowski et al. 2020). As a result, the data used for analysis may be inaccurate or biased, which could compromise the reliability of AI-based conclusions and choices. For instance, incomplete or inaccurate data from specific demographic groups may lead to biased results and unfair treatment of individuals. This happens because the AI analyses and draws conclusions based on data that are not of high quality.

## Openness (*Protection of Personal Information Act* principle 6)

Notifying the regulator and data subjects about personal data gathering may help with transparency; however, this approach may not directly address transparency concerns when it comes to the use of AI algorithms. This is because of the algorithm's automated nature. In this situation, the organisation's algorithmic decision-making processes might evolve after the introduction of AI software (Köchling & Wehner 2020). This can result in stakeholders being unaware of the inner workings of AI models and their possible effects on people. It also becomes more challenging to recognise and correct biases or inaccuracies that may affect people's learning environments.

## Security safeguards (*Protection of Personal Information Act* principle 7)

This principle can address AI privacy concerns. This principle mandates organisations to implement technical and organisational measures to ensure data security. If adopted in AI environments, the principle may address vulnerabilities in the storage, handling, or access control of data used in AI-driven processes.

## Data subject participation (*Protection of Personal Information Act* principle 8)

According to the principle 8, anybody whose personal data may be in a controller's possession has the right to obtain confirmation (at no cost) that the controller maintains such data, as well as a description of such data. Applying this approach to AI could potentially address privacy concerns.

## Recommendations

To effectively address the intricate intersection of AI technologies and privacy concerns within the existing *POPI Act* framework, we recommend the development of regulatory guidance tailored specifically to AI implementation. This guidance should provide comprehensive instructions and best practices that enable organisations to navigate the unique challenges posed by AI while upholding the core principles of the *POPI Act*.

By establishing this regulatory guidance, South Africa can proactively address the ethical and privacy implications of AI technologies. The guidance would offer practical insights on implementing measures to ensure accountability, mitigate biases, enhance transparency, and maintain data quality throughout the AI lifecycle. Furthermore, it would outline ways to incorporate data subject participation and uphold stringent security safeguards in AI-driven processes.

This approach to regulatory supplementation will empower organisations to harness the potential of AI while adhering to the principles of the *POPI Act*. It reflects a forward-looking stance, enabling the law to effectively govern emerging technologies, safeguard personal information, and maintain South Africa's commitment to privacy rights in the digital age.

This guideline can accommodate the following suggestions:

### Accountability (*Protection of Personal Information Act* principle 1)

It should be mandatory for organisations that design, develop, or implement AI applications to introduce AI Ethics Officers in order to ensure transparent, unbiased AI systems. Establish a dedicated role responsible for defining, implementing, and enforcing ethical considerations in AI development. These officers can provide clear guidelines and oversee transparency, bias mitigation, and informed decision-making within AI initiatives. Although not compulsory, organisations should strongly consider this role to uphold ethical practices in AI.

### Processing limitation (*Protection of Personal Information Act* principle 2)

Entities that design, develop, or implement AI applications should conduct mandatory regular audits for AI bias and align AI practices with ethical data use. Regularly assess AI systems for biases and fairness issues. Ensure that data collection practices adhere to ethical guidelines to prevent unintended discriminatory outcomes or the amplification of biases in AI-driven decision-making.

### Purpose specification (*Protection of Personal Information Act* principle 3)

The entities should improve AI algorithm usage transparency. Data subjects should receive information not only about the purpose of data collection but also about the analysis of their information using AI algorithms. Ensure that individuals are fully aware of the potential implications of AI-driven analysis and modelling when providing consent.

### The *Protection of Personal Information Act* principle 4 limits further processing

Entities must strive to use AI models and algorithms that are explainable and interpretable. They could provide users with explanations for algorithmic outcomes, enabling them to comprehend the reasoning behind specific decisions or recommendations. They may also implement mechanisms for ongoing and dynamic consent, allowing individuals to revisit and update their consent preferences as new data sets emerge.

### Information quality (*Protection of Personal Information Act* principle 5)

Entities should develop mechanisms to ensure data quality in AI training. Implement strategies to address underrepresented demographic groups and inaccuracies in training data. They should develop fairness-aware algorithms and validation techniques to mitigate biases and ensure more accurate and equitable AI outcomes.

### Openness (*Protection of Personal Information Act* principle 6)

Entities should establish transparency requirements for AI algorithm changes. Develop guidelines that ensure transparency in algorithmic decision-making processes, even as algorithms evolve. Inform stakeholders about changes to AI models and their implications to enhance understanding and detect potential bias.

### Security safeguards (*Protection of Personal Information Act* principle 7)

The entities should apply technical and organisational security measures to AI environments. Apply *POPI Act's* security principles to safeguard data used in AI-driven processes. Address AI-specific vulnerabilities, such as data breaches caused by poorly secured AI models or uncontrolled access to AI-generated insights.

### The *Protection of Personal Information Act* principle 8 involves the participation of data subjects

Entities should enable data subjects to understand AI-derived insights. When fulfilling data subject requests, provide clear and understandable explanations of AI-generated insights and predictions. Develop mechanisms to explain the basis of AI decisions in a way that data subjects can comprehend. This can include addressing these concerns by establishing strict regulations for data handling, access control, encryption, and consent mechanisms to safeguard information.

## Conclusion

The increasing use of AI in several sectors highlights significant ethical considerations, with a particular focus on privacy-related issues. This research examined the complex interaction of AI technology and privacy concerns within the scope of South Africa's *POPI Act*. The principles included in the *POPI Act*, although strong, demonstrate significant constraints in effectively tackling the dynamic issues presented by AI technology.

Enforcing the mandatory appointment of AI Ethics Officers is a critical step in ensuring openness, impartiality in AI systems, and ethical decision-making.

Furthermore, the research encourages the use of stringent procedures, such as mandatory audits to address bias in AI systems, and the alignment of AI practices with ethical data usage principles, namely adhering to the Processing Limitation principle outlined in the *POPI Act*.

This study emphasises the fluid nature of AI algorithms, which underscores the need for ongoing efforts to maintain transparency in the processes of algorithmic decision-making. The implementation of robust technical and organisational security measures in AI operations is essential for enhancing resilience against potential data breaches and unauthorised access.

Within a wider legislative framework, the study suggests the creation of specific guidelines to facilitate the integration of AI within the scope of the *POPI Act*. This proactive strategy aims to address the ethical and privacy problems connected with AI technology in a forward-looking manner. The proposed regulatory advice offers a strategic framework for companies, providing practical knowledge to effectively address the complex obstacles presented by AI, all while maintaining the fundamental principles outlined in the *POPI Act*.

The article makes a valuable contribution to the continuing academic discussion surrounding data protection and AI. It achieves this by presenting practical solutions that effectively address the complex relationship between AI and privacy, specifically within the legal framework of the *POPI Act*.

## Acknowledgements

### Competing interests

The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

### Authors' contributions

V.M. conceived of the presented idea. M.M. and V.M. wrote the literature section. V.M. wrote the discussion of the findings section. F.N. wrote the introduction and methodology. All authors revised the manuscript.

### Funding information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

### Data availability

The articles and government documents used in this study are available from the corresponding author, V.M. upon reasonable request.

### Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. It does not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are responsible for this article's results, findings, and content.

## References

Abbasgholizadeh Rahimi, S., Cwintal, M., Huang, Y., Ghadiri, P., Grad, R., Poenaru, D. et al., 2022, 'Application of artificial intelligence in shared decision making: Scoping review', *JMIR Medical Informatics* 10(8), e36199. https://doi.org/10.2196/36199

Adams, R., Adeleke, F., Anderson, D., Bawa, A., Branson, N., Christoffels, A. et al., 2021, 'POPIA code of conduct for research', *South African Journal of Science* 117(5–6), 1–12. https://doi.org/10.17159/sajs.2021/10933C

Adams, N.R., 2021, South African company law in the fourth industrial revolution: Does artificial intelligence create a need for legal reform? Master's dissertation, University of the Western Cape, Cape town.

Arora, M., 2021, 'Artificial Intelligence: New Pathways and Challenges in Higher Education', In S. Verma & P. Tomar (eds.), *Impact of AI technologies on teaching, learning, and research in higher education*, pp. 30–48, IGI Global, Pennsylvania.

Awotunde, J.B., Adeniyi, E.A., Ogundokun, R.O. & Ayo, F.E., 2021, 'Application of big data with fintech in financial services', in P.M.S. Choi & S.H. Huang (eds)., *Fintech with artificial intelligence, Big Data and Blockchain*, pp. 107–132, Springer.

Aysolmaz, B., Müller, R. & Meacham, D., 2023, 'The public perceptions of algorithmic decision-making systems: Results from a large-scale survey', *Telematics and Informatics* 79, 101954. https://doi.org/10.1016/j.tele.2023.101954

Aziz, L.A.-R. & Andriansyah, Y., 2023, 'The role artificial intelligence in modern banking: An exploration of AI-driven approaches for enhanced fraud prevention, risk management and regulatory compliance', *Reviews of Contemporary Business Analytics* 6(1), 110–132.

Bowen, G.A., 2009, Document analysis as a qualitative research method', *Qualitative Research Journal* 9(2), 27–40. https://doi.org/10.3316/QRJ0902027

Bharadiya, J., 2023, 'Artificial Intelligence in transportation systems. A critical review', *American Journal of Computing and Engineering* 6(1), a34–a45. https://doi.org/10.47672/ajce.1487

Bleier, A., Goldfarb, A. & Tucker, C., 2020, 'Consumer privacy and the future of data-based innovation and marketing', *International Journal of Research in Marketing* 37(3), 466–480. https://doi.org/10.1016/j.ijresmar.2020.03.006

Braun, V. & Clarke, V., 2013, *Successful qualitative research: A practical guide for beginners*, Sage, London.

Buiten, M.C., 2019, 'Towards intelligent regulation of artificial intelligence', *European Journal of Risk Regulation* 10(1), 41–59. https://doi.org/10.1017/err.2019.8

Cao, L., 2022, 'Ai in finance: challenges, techniques and opportunities', *ACM Computing Surveys (CSUR)* 55(3), 1–38.

Chew, H.S.J. & Achananuparp, P., 2022, 'Perceptions and needs of artificial intelligence in health care to increase adoption: Scoping review', *Journal of Medical Internet Research* 24(1), e32939. https://doi.org/10.2196/32939

Da Veiga, A. & Ophoff, J., 2020, 'Concern for information privacy: A cross-nation study of the United Kingdom and South Africa', in N. Clarke & S. Furnell (eds.), *Human Aspects of Information Security and Assurance: 14th IFIP WG 11.12 International Symposium*, Springer Mytilene, Lesbos, Greece, July 08–10, 2020, pp. 16–29.

Davenport, T., Guha, A., Grewal, D. & Bressgott, T., 2020, 'How artificial intelligence will change the future of marketing', *Journal of the Academy of Marketing Science* 48, 24–42. https://doi.org/10.1007/s11747-019-00696-0

Dwivedi, Y.K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T. et al., 2021, 'Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice, and policy', *International Journal of Information Management* 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Dunne, B., Pettigrew, J. & Robinson, K., 2016, 'Using historical documentary methods to explore the history of occupational therapy', *British Journal of Occupational Therapy* 79(6), 376–384. https://doi.org/10.1177/0308022615608639

Firouzi, F., Farahani, B., Barzegari, M. & Daneshmand, M., 2020, 'AI-driven data monetization: The other face of data in IoT-based smart and connected health', *IEEE Internet of Things Journal* 9(8), 5581–5599. https://doi.org/10.1109/JIOT.2020.3027971

Flick, U., 2018, *An introduction to qualitative research*, Sage, London.

Gianni, R., Lehtinen, S. & Nieminen, M., 2022, 'Governance of responsible AI: From ethical guidelines to cooperative policies', *Frontiers in Computer Science* 4, 873437. https://doi.org/10.3389/fcomp.2022.873437

Government of South Africa, 2013, *No.4 of 2013: Protection of personal information Act, 2013*, The Government of South Africa, viewed 11 November 2023, from https://www.gov.za/documents/protection-personal-information-act#.

Haleem, A., Javaid, M., Qadri, M.A., Singh, R.P. & Suman, R., 2022, 'Artificial intelligence (AI) applications for marketing: A literature-based study', *International Journal of Intelligent Networks* 9, 119–132. https://doi.org/10.1016/j.ijin.2022.08.005

Huang, M.-H. & Rust, R.T., 2021, 'A strategic framework for artificial intelligence in marketing', *Journal of the Academy of Marketing Science* 49, 30–50. https://doi.org/10.1007/s11747-020-00749-9

Ingrams, A., Kaufmann, W. & Jacobs, D., 2022, 'In AI we trust? Citizen perceptions of AI in government decision making', *Policy & Internet* 14(2), 390–409. https://doi.org/10.1002/poi3.276

Ka Mtuze, S.S. & Morige, M., 2024, 'Towards drafting artificial intelligence (AI) legislation in South Africa', *Obiter* 45(1), 161–179. https://doi.org/10.17159/obiter.v45i1.18399

Kaličanin, K., Čolović, M., Njeguš, A. & Mitić, V., 2019, 'Benefits of artificial intelligence and machine learning in marketing', in *Sinteza 2019 - International scientific conference on information technology and data related research*, Singidunum University, Belgrade, Serbia, April 20, 2019, pp. 472–477.

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F. et al., 2023, 'ChatGPT for good? On opportunities and challenges of large language models for education', *Learning and Individual Differences* 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Keshta, I., 2022, 'AI-driven IoT for smart health care: Security and privacy issues', *Informatics in Medicine Unlocked* 30, 100903. https://doi.org/10.1016/j.imu.2022.100903

Kridel, C., 2015, 'The biographical and documentary milieu', in M.F. He. B.D. Schultz & W.H. Schubert (eds.), *The Sage guide to curriculum in education*, pp. 311–318, Sage, London.

Köbis, L. & Mehner, C., 2021, 'Ethical questions raised by AI-supported mentoring in higher education', *Frontiers in Artificial Intelligence* 4, 624050. https://doi.org/10.3389/frai.2021.624050

Köchling, A. & Wehner, M.C., 2020, 'Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development', *Business Research* 13(3), 795–848. https://doi.org/10.1007/s40685-020-00134-w

Köchling, A., Riazy, S., Wehner, M.C. & Simbeck, K., 2021, 'Highly accurate, but still discriminatory: A fairness evaluation of algorithmic video analysis in the recruitment context', *Business & Information Systems Engineering* 63, 39–54. https://doi.org/10.1007/s12599-020-00673-w

Kooli, C., 2023, 'Chatbots in education and research: A critical examination of ethical implications and solutions', *Sustainability* 15(7), 5614. https://doi.org/10.3390/su15075614

Kuziemski, M. & Misuraca, G., 2020, 'AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings', *Telecommunications Policy* 44(6), 101976. https://doi.org/10.1016/j.telpol.2020.101976

Li, L., 2023, 'Role of chatbots on gastroenterology: Let's chat about the future', *Gastroenterology & Endoscopy* 1(3), 144–149. https://doi.org/10.1016/j.gande.2023.06.002

Lin, S., 2022, 'A clinician's guide to artificial intelligence (AI): Why and how primary care should lead the health care AI revolution', *The Journal of the American Board of Family Medicine* 35(1), 175–184. https://doi.org/10.3122/jabfm.2022.01.210226

Marcinkowski, F., Kieslich, K., Starke, C. & Lünich, M., 2020, 'Implications of AI (un-) fairness in higher education admissions: The effects of perceived AI (un-) fairness on exit, voice, and organizational reputation', In M. Hildebrandt & C. Castillo (eds.), *Proceedings of the 2020 conference on fairness, accountability, and transparency*, Association for Computing Machinery, Barcelona, New York, 27–30th January, pp. 122–130.

Matulis, J. & McCoy, R., 2023, 'Relief in sight? Chatbots, in-baskets, and the overwhelmed primary care clinician', *Journal of General Internal Medicine* 38(12), 2808–2815. https://doi.org/10.1007/s11606-023-08271-8

Mbonye, V., Subramaniam, P.R. & Padayachee, I., 2021, 'POPIA compliant regulatory framework for smart grids to secure gaps in existing privacy laws', in S. Pudaruth & U. Singh (eds.), *International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*, IEEE, Durban, August 05–06, 2021, pp. 1–8.

Milan, A., Sahu, R. & Sandhu, J.K., 2023, 'Impact of AI on social marketing and its usage in social media: A review analysis', in *International Conference on Circuit Power and Computing Technologies (ICCPCT)*, IEEE, Kollam, July 06–08, 2023, pp. 1749–1754.

Mogalakwe, M., 2009, 'The documentary research method: Using documentary sources in social research', *Eastern Africa Social Science Research Review* 25(1), 43–58. https://doi.org/10.1353/eas.0.0006

Murdoch, B., 2021, 'Privacy and artificial intelligence: Challenges for protecting health information in a new era', *BMC Medical Ethics* 22(1), 1–5. https://doi.org/10.1186/s12910-021-00687-3

Nagy, S. & Hajdú, N., 2021, 'Consumer acceptance of the use of artificial intelligence in online shopping: Evidence from Hungary', *Amfiteatru Economic* 23(56), 155–173. https://doi.org/10.24818/EA/2021/56/155

Naidoo, V., 2021, 'POPI and its possible relation to AI or machine learning systems', *Without Prejudice* 21(4), 67–68.

Netshakhuma, N.S., 2020, 'Assessment of a South Africa national consultative workshop on the Protection of Personal Information Act (POPIA)', *Global Knowledge, Memory, and Communication* 69(1/2), 58–74. https://doi.org/10.1108/GKMC-02-2019-0026

Perrault, R., Yoav, S., Brynjolfsson, E., Jack, C., Etchmendy, J., Grosz, B. et al., 2019, *Artificial intelligence index report 2019*, HAI Centre Stanford University, viewed 11 November 2023, from https://wp.oecd.ai/app/uploads/2020/07/ai_index_2019_introduction.pdf.

Prem, E., 2019, 'Artificial intelligence for innovation in Austria', *Technology Innovation Management Review* 9(12), 5–15. https://doi.org/10.22215/timreview/1287

Rahman, A., Hossain, M.S., Muhammad, G., Kundu, D., Debnath, T., Rahman, M. et al., 2023, 'Federated learning-based AI approaches in smart healthcare: Concepts, taxonomies, challenges and open issues', *Cluster Computing* 26(4), 2271–2311. https://doi.org/10.1007/s10586-022-03658-4

Ramcharan, S., 2020, *How will the information regulator manage and control advancements in technology to minimise its adverse effects on the protection of personal information?*, Doctoral dissertation, University of KwaZulu Natal, viewed 20 November 2023, from https://ukzn-dspace.ukzn.ac.za/bitstream/handle/10413/19126/Ramcharan_Sarika_2020.pdf?sequence=1&isAllowed=y.

Sallam, M., 2023, 'ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns', *Healthcare* 11(6), 887. https://doi.org/10.3390/healthcare11060887

Shah, N., Engineer, S., Bhagat, N., Chauhan, H. & Shah, M., 2020, 'Research trends on the usage of machine learning and artificial intelligence in advertising', *Augmented Human Research* 5, 1–15. https://doi.org/10.1007/s41133-020-00038-8

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A. et al., 2018, 'A general reinforcement learning algorithm that masters chess, shogi and Go through self-play', *Science* 362(6419), 1140–1144. https://doi.org/10.1126/science.aar6404

Singh, T. & Mishra, J., 2021, 'Learning with artificial intelligence systems: Application, challenges, and opportunities', in S. Verma & P. Tomar (eds.), *Impact of AI technologies on teaching, learning, and research in higher education*, pp. 236–253, IGI Global, Pennsylvania.

Staunton, C., Adams, R., Anderson, D., Croxton, T., Kamuya, D., Munene, M. et al., 2020, 'Protection of Personal Information Act 2013 and data protection for health research in South Africa', *International Data Privacy Law* 10(2), 160–179. https://doi.org/10.1093/idpl/ipz024

Sun, L., Gupta, R.K. & Sharma, A., 2022, 'Review and potential for artificial intelligence in healthcare', *International Journal of System Assurance Engineering and Management* 13, 54–62. https://doi.org/10.1007/s13198-021-01221-9

Swales, L., 2021, 'The Protection of Personal Information Act and data de-identification', *South African Journal of Science* 117(7–8), 1–3. https://doi.org/10.17159/sajs.2021/10808

Thaldar, D. & Townsend, B., 2021, 'Exempting health research from the consent provisions of POPIA', *Potchefstroom Electronic Law Journal/Potchefstroomse Elektroniese Regsblad* 24(1), 1–32. https://doi.org/10.17159/1727-3781/2021/v24i0a10420

Townsend, B. & Botes, M., 2023, 'Bridging the regulatory gaps created by smart and connected technologies in South Africa', *South African Journal of Bioethics and Law* 16(2), 36–41. https://doi.org/10.7196/SAJBL.2023.v16i2.201

Thowfeek, M.H., Samsudeen, S.N. & Sanjeetha, M.B.F., 2020, 'Drivers of artificial intelligence in banking service sectors', *Solid State Technology* 63(5), 6400–6411.

Truby, J., Brown, R. & Dahdal, A., 2020, 'Banking on AI: Mandating a proactive approach to AI regulation in the financial sector', *Law and Financial Markets Review* 14(2), 110–120. https://doi.org/10.1080/17521440.2020.1760454

Wirtz, B.W., Weyerer, J.C. & Sturm, B.J., 2020, 'The dark sides of artificial intelligence: An integrated AI governance framework for public administration', *International Journal of Public Administration* 43(9), 818–829. https://doi.org/10.1080/01900692.2020.1749851

Yaiprasert, C. & Hidayanto, A.N. 2023, 'AI-driven ensemble three machine learning to enhance digital marketing strategies in the food delivery business', *Intelligent Systems with Applications* 18, 200235. https://doi.org/10.1016/j.iswa.2023.200235

Zheng, X. & Cai, Z., 2020, 'Privacy-preserved data sharing towards multiple parties in industrial iots', *IEEE Journal on Selected Areas in Communications* 38(5), 968–979. https://doi.org/10.1109/JSAC.2020.2980802