E-Journal of Humanities, Arts and Social Sciences (EHASS)
ISSN – Online 2720-7722 | Print 2821-8949
Volume 6 Issue 15 2025 pp 4369 - 4379
Available online at: https://noyam.org/journals/ehass/
DOI: https://doi.org/10.38159/ehass.202561512

Algorithmic Justice in South Africa: Safeguarding Human Rights in AI-Driven Legal Systems

Bulelani Thukuse [1] & Paul S. Masumbe [1]
[1] School of Law, Faculty of Law, Humanities and Social Sciences, Walter Sisulu University, South Africa.

ABSTRACT
This paper investigated the intersection between emerging algorithmic technologies and the enforcement of human rights within South Africa's legal system. As artificial intelligence (AI) becomes increasingly integrated into legal processes ranging from predictive policing to judicial decision-making, urgent questions arise regarding the compatibility of such technologies with constitutional protections, ethical standards, and democratic accountability. The study critically examined how AI-driven legal tools may inadvertently entrench existing biases, obscure accountability, or compromise the right to a fair trial, especially for historically marginalised groups. This study drew on South Africa's constitutional framework, global human rights principles, and a comparative analysis of AI regulation in the *EU and Canada* to assess how algorithmic systems can operate in a manner that respects and promotes justice, fairness, and transparency. The key focus areas of this study included the risk of data-driven discrimination, the opacity of algorithmic reasoning, and the adequacy of current regulatory safeguards. Through a detailed analysis of case law, journal articles, statutory developments, and technological trends, this study evaluates whether South Africa's legal and institutional frameworks are sufficiently equipped to manage the risks and opportunities presented by AI. The paper offers normative and policy-oriented recommendations to ensure that algorithmic tools deployed within the legal domain uphold constitutional values, enhance legal accountability, and foster trust in the justice system. The study contributes by proposing a rights-based framework to ensure that AI in South Africa's legal system upholds justice, accountability, and full regulation of AI.

Keywords: Algorithmic Justice, Artificial Intelligence, Human Rights, Legal Technology, South African Constitution, Legal Accountability, AI Regulation.

Correspondence
Paul S. Masumbe
Email:
pmasumbe@wsu.ac.za

Publication History
Received: 12th May, 2025
Accepted: 11th November, 2025
Published online:
30th December, 2025

To Cite this Article:
Thukuse, Bulelani, and Paul S. Masumbe. "Algorithmic Justice in South Africa: Safeguarding Human Rights in AI-Driven Legal Systems," *E-Journal of Humanities, Arts and Social Sciences* 6, no. 15 (2025): 4369 - 4379, https://doi.org/10.38159/ehass.202561512.

INTRODUCTION
In the digital age, the convergence of artificial intelligence (AI) and legal systems has emerged as both a promise and a peril for modern democracies.[1] Across the globe, AI is steadily reshaping the way legal systems function.[2] From streamlining routine administrative work to supporting legal research and even influencing decisions around bail and sentencing, AI is becoming an integral part of the legal toolkit. According to a 2025 report by the Secretariat and the Association of Certified E-Discovery Specialists

[1] Patricia Gomes Rêgo de Almeida, Carlos Denner dos Santos, and Josivania Silva Farias, "Artificial Intelligence Regulation: A Framework for Governance," *Ethics and Information Technology* 23, no. 3 (September 21, 2021): 505–25, https://doi.org/10.1007/s10676-021-09593-z.
[2] Secretariat, "AI Adoption Surges in the Legal Industry: Key Findings from the 2025 Secretariat and ACEDS Global Artificial Intelligence Report," 2025, https://secretariat-intl.com/insights/ai-adoption-surges-in-the-legal-industry/.

(ACEDS), nearly three out of four legal professionals, about 74%, anticipate incorporating AI-powered tools into their daily work within the next year.[3] This growing reliance on AI reflects its increasing value in enhancing efficiency, consistency, and access to justice across legal processes. South Africa, with its progressive Constitution and deep commitment to human rights, stands at a critical juncture where the adoption of AI-driven tools must be carefully scrutinised to safeguard the fundamental values that underpin its democratic society.[4] The country's history of systemic discrimination, entrenched inequalities, and struggle for social justice provides a unique backdrop for examining how algorithmic technologies may impact vulnerable and historically marginalised communities.

The integration of AI into legal decision-making introduces efficiency and predictive capabilities that can potentially enhance access to justice and reduce case backlogs.[5] However, it also raises pressing concerns about algorithmic bias, the opacity of machine learning processes, and the risk of eroding procedural fairness.[6] Tiwari further argued that artificial intelligence has the potential to significantly enhance judicial efficiency by streamlining legal research, assisting in the drafting of judgments, and employing predictive analytics to anticipate case outcomes.[7] These tools can help reduce delays and promote consistency in legal reasoning across similar cases. However, the integration of AI into judicial processes also introduces serious legal and constitutional challenges, most notably in relation to due process, judicial accountability, and the transparency of decision-making systems.[8] Chen on his analysis, which shows research that AI systems, when trained on biased datasets, may perpetuate or even amplify historical prejudices embedded in data, disproportionately affecting disadvantaged groups.[9]

Moreover, the "black box" nature of many AI models undermines the principle of accountability, making it difficult to ascertain how decisions are made or to challenge them effectively in court.[10] This opacity not only hampers accountability but also risks infringing Section 9 of the South African Constitution,[11] which guarantees equality and prohibits unfair discrimination.[12] Confronted with these tensions, South Africa's courts and policymakers must proceed with caution, ensuring AI systems uphold constitutional rights rather than erode them. While the potential benefits of AI cannot be ignored, its deployment in sensitive areas such as predictive policing, surveillance, and judicial sentencing must be carefully aligned with constitutional safeguards and international human rights norms.[13] Drawing on comparative models like the European Union's Artificial Intelligence Act, which establishes a robust regulatory framework for high-risk AI applications, especially those used in law enforcement and the judiciary, South Africa can develop a balanced approach that encourages innovation while proactively protecting against violations of rights and freedoms.[14]

However, Petse and Phindelo are of the view that using ChatGPT as a research tool in South African legal practice poses serious risks to the integrity of the courts and the broader administration of justice.[15] They argued that this AI tool is seen not just as flawed, but also as potentially deceptive, particularly when legal professionals rely on it without proper oversight. There have been instances where ChatGPT has generated inaccurate or entirely fictitious legal content, leading to the circulation of

---

[3] Secretariat, "AI Adoption Surges in the Legal Industry: Key Findings from the 2025 Secretariat and ACEDS Global Artificial Intelligence Report."

[4] Ryszard Lisinski, "Regulation of Artificial Intelligence in South Africa: Progress, Challenges and the Road Ahead," Fluxman, 2025, https://www.fluxmans.com/article/ai-regulation-south-africa.

[5] Tiwari Shashvat, " Legal Implications of AI in Judicial Decision-Making," *International Journal of Legal and Law Research*, 2024.

[6] Shashvat, " Legal Implications of AI in Judicial Decision-Making."

[7] Shashvat, " Legal Implications of AI in Judicial Decision-Making."

[8] Shashvat, " Legal Implications of AI in Judicial Decision-Making."

[9] Zhisheng Chen, "Ethics and Discrimination in Artificial Intelligence-Enabled Recruitment Practices," *Humanities and Social Sciences Communications* 10, no. 1 (2023): 1–12.

[10] Inioluwa Deborah Raji et al., "Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 2020, 33–44.

[11] South Africa, *The Constitution of the Republic of South Africa* (Juta Law Firm: Juta, 2009). Sec. 9.

[12] Editorial, "Legal and ethical principles governing the use of artificial intelligence," De Rebus (2022), noting that arbitrary or opaque AI systems would be contrary to Section 9 of the Constitution, which prohibits unfair discrimination.

[13] Kelly Stone, "Responsible Use of Artificial Intelligence for Crime Prevention in South Africa: A Topical Guide," *Department of Science & Innovation (South Africa)*, n.d.

[14] "AI and Policing: The Benefits and Challenges of Artificial Intelligence for Law Enforcement (Europol, 2024)," n.d.

[15] Isiphile Petse and Usenathi Phindelo, "The Influence of ChatGPT-Generated Data on the Administration of Justice in South Africa," *Turf Law Journal*, 2025.

misleading documents. When such unverified information finds its way into court proceedings, it not only compromises judicial outcomes but also places legal practitioners at risk.[16] They further argued that presenting false information knowingly or not could amount to perjury under common law, especially if the practitioner fails to fact-check the AI-generated content before submitting it in court.[17] This paper examines the rise of algorithmic decision-making in South Africa's Legal System. The study explored the opacity of AI reasoning and its implications for legal accountability. It examined how legal AI systems can be regulated to ensure ethical and rights-based governance. Furthermore, it provided a comparative analysis of how the EU, Canada, and the UK regulate high-risk AI in legal contexts. The study argues for a careful and values-based approach to using AI in the legal system, one that is guided by the Constitution's commitment to justice and equality and supported by strong legal safeguards. Hence, this study will attempt to investigate whether South Africa's evolving legal and institutional frameworks can govern the rise of AI in ways that respect, protect, and fulfil fundamental human rights, ensuring that technological innovation does not come at the cost of constitutional justice.

METHODOLOGY

This study used a qualitative, desktop-based approach to examine the legal and human rights implications of AI in South Africa's judicial system. It analysed constitutional provisions, legislation, case law, and academic literature, focusing on how AI impacts accountability, equality, and access to justice. The above research method is more ideal because it includes a thorough review of academic sources and no data collection. A comparative legal lens was employed, drawing on regulatory frameworks from the European Union and Canada, to analyse how these jurisdictions govern high-risk AI applications within their justice systems. AI is viewed not just as a tool for efficiency, but as a potential threat to fairness if not properly governed.

DISCUSSION

The Rise of Algorithmic Decision-Making in South Africa's Legal System

Artificial intelligence is steadily becoming a key component of South Africa's legal landscape, changing the way legal professionals conduct research, draft legal documents, and evaluate case strategies.[18] AI are at the forefront of this transformation, merging authoritative legal databases with sophisticated AI functions to improve accuracy and efficiency in legal work.[19] However, recent developments, most notably the case of *Mavundla v MEC Department of Co-Operative Government and Traditional Affairs and Others,*[20] have revealed serious risks tied to AI in legal practice. In that matter, AI-generated submissions included several fictitious case citations, sparking concern across the legal fraternity. The incorporation of artificial intelligence into legal proceedings in South Africa has been minimal, with few reported cases addressing its use directly.[21] However, an exception was found in *Parker v Forsyth NO and Others*:[22]

> "A case heard in the Regional Court for the Gauteng Division, Johannesburg. During oral argument, the plaintiff's attorney raised points concerning the legal standing of body corporates and their ability to initiate or be subject to defamation claims. Following the hearing, and upon the court's request, the plaintiff's legal representative submitted a list of legal authorities to support the argument. The defendant's attorneys, upon reviewing the list, were unable to trace any of the cited cases and subsequently sought clarification regarding their source. It was then revealed that the plaintiff's attorney had neither accessed nor reviewed the judgments in question. Instead, the references were generated by ChatGPT, an artificial intelligence chatbot. This incident highlights the risks associated with unverified AI-generated legal content and

[16] Petse and Phindelo, "The Influence of ChatGPT-Generated Data on the Administration of Justice in South Africa."
[17] Petse and Phindelo, "The Influence of ChatGPT-Generated Data on the Administration of Justice in South Africa."
[18] Lexis Nexis South Africa, "LexisNexis South Africa Highlights Responsible AI Integration in Law," January 2025, https://www.lexisnexis.com/blogs/za/b/legal/posts/lexisnexis-south-africa-highlights-responsible-ai-integration-in-law.
[19] Lexis Nexis South Africa, "LexisNexis South Africa Highlights Responsible AI Integration in Law."
[20] *Mavundla v MEC Department of Co-Operative Government and Traditional Affairs and Others* (7940 2024P) [2025] ZAKZPHC 2 (8 January 2025).
[21] *Mavundla v MEC Department of Co-Operative Government and Traditional Affairs and Others*, para 42.
[22] *Parker v Forsyth NO and others* (Regional Court, Johannesburg, Gauteng) unreported case no 1585/20 (29 June 2023) (Parker), and reported on Law Library South Africa as *Parker v Forsyth NO and others* [2023] ZAGPRD 1, available at https://lawlibrary.org.za/akn/zagp/judgment/zagprd/2023/1/eng@2023-06-29. (Accessed 06 July 2025).

> reinforces the importance of professional diligence and ethical responsibility when incorporating technology into legal practice."[23]

In this case, the court analysed section 28(2) of the CPA[24] and stated that this provision provide that if someone gives false information under oath in terms of section 21(1) or 25(1), leading to the issuing and execution of a search or arrest warrant, and that person is later convicted of perjury as a result, the court may award compensation to any person who suffered harm due to the unlawful entry, search, or seizure.[25] In a recent contribution to the legal literature, van Eck attempted to deal with the growing role of artificial intelligence in legal practice in his 2024 article.[26] The article offers an in-depth analysis of South Africa's current legal stance, as well as comparative insights from international jurisdictions, on the ethical and practical implications of using AI tools, particularly ChatGPT, in legal research and practice. Van Eck further acknowledged that:

> "While AI platforms like ChatGPT may offer the promise of increased efficiency within the legal sector, their reliability remains deeply questionable."[27]

As he noted, the responses generated by such tools frequently contain fabricated or inaccurate legal information, especially when prompted for legal citations or case law.[28] The abovementioned case of *Parker v Forsyth NO and Others*[29] serves as a key focus of Van Eck's analysis. He critically evaluated the court's response to the plaintiff attorney's submission of fictitious case law sourced from ChatGPT.[30] As South Africa continues to embrace legal innovation, the need to strike a balance between digital automation and adherence to constitutional principles, ethical standards, and professional diligence is becoming more urgent.

Confronting Algorithmic Bias in AI-Driven Justice Systems in South Africa

AI systems rely heavily on large datasets for training, but these datasets often include embedded biases that reflect historical and societal inequalities.[31] If such biases go unaddressed, AI tools risk producing outcomes that are not only unfair but also discriminatory, thereby reinforcing existing social disparities.[32] To prevent this, it is essential to implement strategies that can detect, reduce, and manage bias within AI models. Equally important is ensuring transparency throughout the development and deployment of AI technologies, as this fosters ethical accountability and allows stakeholders to better understand how decisions are made.[33] In integrating AI into South Africa's justice system, especially in areas like predictive policing, bail assessment, and sentencing, there is an urgent need to confront algorithmic bias.[34] These systems typically rely on historical data that echo the structural inequalities of apartheid-era law

---

[23] *Parker v Forsyth NO and others* 86-87.

[24] Section 28(2) of the Criminal Procedure Act.

[25] Section 28(2) of the CPA, see also *Parker v Forsyth NO and Others,* where the court further stated that this can be done on application by the affected person or by the prosecutor acting on their behalf. In such cases, the compensation process follows the provisions of section 300 of the CPA, applied with the necessary changes (mutatis mutandis).

[26] Michele van Eck, "Error 404 or an Error in Judgment? An Ethical Framework for the Use of ChatGPT in the Legal Profession," *Journal of South African Law/Tydskrif Vir Die Suid-Afrikaanse Reg* 2024, no. 3 (2024): 469–90.

[27] van Eck, "Error 404 or an Error in Judgment? An Ethical Framework for the Use of ChatGPT in the Legal Profession."

[28] van Eck, "Error 404 or an Error in Judgment? An Ethical Framework for the Use of ChatGPT in the Legal Profession."471.

[29] *Parker v Forsyth NO and others* (n22).

[30] Van Eck argued that although the incident raised serious concerns about ethical standards and professional diligence, the judgment failed to provide any meaningful guidance on the duties of legal practitioners when using AI-generated material. Instead, the court appeared to overlook the attorney's conduct entirely. According to Van Eck, this omission risks setting a dangerous precedent, as it raises the question of whether the attorney's use of unverified sources constituted mere negligence or something more serious in breach of ethical standards, See para 474-475.

[31] George Benneh Mensah, "Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems," *Preprint, November* 10, no. 1 (2023): 1.

[32] Mensah, "Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems."

[33] George Benneh Mensah, Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems, EGRC Africa Regulation Research Consortium (November 2023), DOI: 10.13140/RG.2.2.23381.19685/1.

[34] SD Law Cape Town Attorneys South African Police Service reliance risk, "Predictive Policing Employs Algorithms … Historical Crime Data—"AI in Law Enforcement," 2025, https://www.sdlaw.co.za/articles/ai-in-law-enforcement/?utm_source.

enforcement, and without correction, they risk embedding those injustices into automation.[35] Larsen and Rajab, in their analysis, asserted that:

> "The rapid evolution of artificial intelligence (AI) has transformed countless sectors by improving efficiency and decision-making. Nevertheless, the embedded biases in AI algorithms also raise profound ethical, legal, and social concerns that demand our attention."[36]

Zhou and others claimed that the rapid adoption of AI and machine learning (ML) algorithms has opened doors for innovation across various fields, yet it has also introduced significant concerns around fairness, bias, and accountability.[37] In their analysis, Zhou et al. offer a comprehensive overview of how these challenges emerge during the design and deployment of ML systems, emphasising the need for transparency and ethical safeguards in algorithmic decision-making.[38] While artificial intelligence represents a transformative advancement in technological innovation, its increasing application across sectors has raised profound legal and ethical concerns, particularly regarding the presence of bias in algorithmic systems.[39] Such bias, if left unchecked, threatens to undermine foundational principles of equality, due process, and non-discrimination enshrined in both domestic and international human rights law.[40] For instance, Pasipamire in his study posits how datasets drawn from crime records can disproportionately flag low-income or black neighbourhoods, perpetuating over-policing and wrongful profiling.[41]

In South Africa, where the legal order continues to confront historical inequalities, the introduction of AI tools in areas like policing and judicial decision-making makes addressing algorithmic bias a legal imperative. Sunita Menon argues that AI systems, often presented as neutral, can inadvertently replicate deep-rooted biases, creating what she terms "digital neocolonialism," where datasets and algorithms imported from global tech giants fail to reflect African lived realities.[42] Menon contends that the assumed objectivity of technology obscures embedded power imbalances, reinforcing existing marginalisation unless datasets and system design are reconfigured to reflect local contexts.[43] This critique highlights that fairness in legal AI cannot be reduced to technical tweaks; it must be grounded in decolonial and context-sensitive design principles that actively dismantle exclusionary structures.

Opacity in AI Reasoning and the Challenge of Legal Accountability

The rising use of artificial intelligence in legal decision-making poses serious concerns for transparency and accountability, especially when such systems function as opaque "black boxes." In the South African context, where constitutional principles emphasize fairness, openness, and justice, the lack of clarity in how AI reaches decisions may undermine essential rights such as due process and access to a fair trial.[44] AI is now deeply embedded in modern life, transforming multiple sectors and significantly improving operational efficiency.[45] Yet, as these technologies evolve, it becomes increasingly important to engage with the ethical challenges that accompany their development and use.

[35] South African Police Service reliance risk, "Predictive Policing Employs Algorithms … Historical Crime Data—"AI in Law Enforcement."

[36] Nancy Larsen and Husam Rajab, "Ensuring Fairness and Mitigating Bias in AI Algorithms," 2024.

[37] Nengfeng Zhou et al., "Bias, Fairness, and Accountability with AI and ML Algorithms," *ArXiv Preprint ArXiv:2105.06558*, 2021.

[38] Zhou et al., "Bias, Fairness, and Accountability with AI and ML Algorithms."

[39] Temitope Ajibola, "Navigating Bias and Fairness in AI Algorithms: Pioneering a Human-Centric Revolution," *Medium*, 2023.

[40] Ajibola, "Navigating Bias and Fairness in AI Algorithms: Pioneering a Human-Centric Revolution."

[41] Notice Pasipamire and Abton Muroyiwa, "Navigating Algorithm Bias in AI: Ensuring Fairness and Trust in Africa," *Frontiers in Research Metrics and Analytics* 9 (2024): 1486600.

[42] Sunita Menon, "Postcolonial Differentials in Algorithmic Bias: Challenging Digital Neo-Colonialism in Africa," *SCRIPTed* 20 (2023): 383.

[43] Menon, "Postcolonial Differentials in Algorithmic Bias: Challenging Digital Neo-Colonialism in Africa."

[44] Mensah, "Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems."

[45] Mensah, "Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems."

Technologies like deep learning models often provide little to no insight into their reasoning, which makes it difficult for both legal professionals and those affected by these decisions to question or understand the outcomes.[46] However, he further stated that:

> "The rapid advancement of large language models (LLMs) has significantly reshaped the field of natural language processing, allowing for the production of text that closely resembles human expression. However, as AI-generated content becomes more prevalent, it brings with it complex ethical dilemmas especially concerning the assignment of moral responsibility for the outputs these systems produce."[47]

In some cases, the lack of clarity becomes dangerous when there are no clear procedures to detect bias or trace how decisions were made, potentially causing harm that cannot be easily corrected.[48] For South Africa to maintain the integrity of its legal system, it is essential that regulations require AI processes to be transparent and that all decisions influenced by such technologies remain open to human evaluation and legal scrutiny.[49] Elizabeth Oluwagbade adds that accountability becomes even more difficult when AI tools are built and used across different legal systems, creating further challenges in legal oversight and access to justice.[50] Artificial intelligence (AI) is playing an increasingly central role in decision-making across vital sectors such as healthcare, finance, criminal justice, and national security. However, as these systems are developed, deployed, and managed across multiple jurisdictions, they complicate established frameworks of legal accountability and regulatory oversight.[51] For instance, in the Netherlands, Judge RJJ van Acht openly acknowledged turning to AI for guidance in resolving a neighbourhood dispute over a home extension that allegedly interfered with a neighbour's solar panel performance.[52] While tools like ChatGPT, DeepSeek and Google's Gemini are increasingly used in various fields, they've also faced ongoing criticism for being unreliable at times and prone to basic errors, raising concerns about their role in serious decision-making.[53] In addition, quoted in The Telegraph, Lord Justice Birss reflected on the practical value of large language models like ChatGPT, saying:

> "What I find most interesting is their ability to summarise information. It's genuinely helpful and yes, I've used it myself." "It gave me a paragraph. I already knew what the answer was, I was about to write it myself, but it produced the summary for me, and I included it in my judgment. It was spot on and very useful."[54]

He further shared that he once asked ChatGPT for a summary of a particular area of law. Lord Justice Birss is believed to be the first British judge to publicly acknowledge using AI assistance in writing a judgment.[55]

AI-Driven Data Governance and the Protection of Personal Information

According to Sagar Patel, the adoption of artificial intelligence within legal and administrative systems has raised heightened concerns about the protection and ethical handling of personal data, particularly in countries like South Africa, where constitutional safeguards for privacy and dignity are well established.[56] Tools such as Large Language Models (LLMs) routinely process highly sensitive information, including

[46] Parvaiz A Ganai and Irshad A Naikoo, "The Ethical Paradox of AI-Generated Texts: Investigating the Moral Responsibility in Generative Models," 2025.
[47] Ganai and Naikoo, "The Ethical Paradox of AI-Generated Texts: Investigating the Moral Responsibility in Generative Models."
[48] Mensah, "Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems."
[49] Mensah, "Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems."
[50] Elizabeth Oluwagbade, "Accountability Without Borders: Cross-Jurisdictional Challenges in Regulating AI-Driven Decisions," n.d.
[51] Oluwagbade, "Accountability Without Borders: Cross-Jurisdictional Challenges in Regulating AI-Driven Decisions."
[52] Perkin Amalaraj, "Decision by Judge to Use ChatGPT to Reach a Verdict Sparks Fury," 2025, https://www.llex.hk/news/other/ridiculous-judge-uses-chatgpt-to-help-reach-a-case-verdict/ar-AA1ogb0L.
[53] Amalaraj, "Decision by Judge to Use ChatGPT to Reach a Verdict Sparks Fury."A Dutch judge has come under fire for consulting ChatGPT. The judge twice went to the chatbot to retrieve facts for a case
He was slammed by AI experts, who said that ChatGPT is easily fallible. A Dutch judge has been slammed after he used ChatGPT to help reach a verdict in a case where thousands of euros were on the line.
[54] Amalaraj, "Decision by Judge to Use ChatGPT to Reach a Verdict Sparks Fury."55.
[55] Amalaraj, "Decision by Judge to Use ChatGPT to Reach a Verdict Sparks Fury." 56.
[56] Sagar Patel, "Safeguarding Sensitive Data in LLM Race – Awareness and Protection," 2025, https://doi.org/10.34218/IJAIRD_03_02_001.

health records, financial details, and personal identifiers, often without individuals fully grasping the potential risks.[57] Patel further emphasised that conventional data protection frameworks are no longer sufficient in AI-driven environments, which require adaptive and proactive safeguards designed for algorithmic technologies.[58] In recent years, artificial intelligence (AI) has quickly become embedded in various aspects of society, ranging from healthcare and the judiciary to public administration and other critical sectors.[59] Svetlana Vashurina emphasised that public trust in AI systems is strongly tied to transparency, data security, and robust regulatory mechanisms, especially when these technologies intersect with fundamental rights.[60] While South Africa's Protection of Personal Information Act (POPIA)[61] offers a foundational legal structure, many scholars caution that the legislation must evolve to keep pace with the specific challenges posed by AI-powered data processing.[62] To truly protect individuals in the digital era, a rights-based approach to AI governance must incorporate algorithmic audits, meaningful user consent protocols, and clearly enforceable accountability systems.[63]

Regulating Legal AI Systems for Ethical and Rights-Based Governance

Regulating artificial intelligence within South Africa's legal system must be firmly anchored in both ethical standards and the foundational values of the Constitution to prevent misuse and safeguard human rights.[64] Legal scholars have raised concerns that in the absence of comprehensive legal frameworks, AI technologies may function without adequate transparency or accountability, particularly in high-stakes areas such as judicial decision-making.[65] The National AI Policy Framework (2024) echoes these concerns, emphasising the development of human-centered AI guided by principles of fairness, inclusivity, and respect for fundamental rights. Oluwagbade argued that effective ethical oversight of AI must include public accountability mechanisms, independent algorithmic audits, and clear procedures for individuals to challenge decisions made by automated systems.[66] Similarly, Benneh Mensah further noted that the need to align South Africa's AI regulation with international human rights norms, advocating for digital systems that prioritise justice, equity, and accessibility.[67] The cross-border nature of AI systems makes it significantly more difficult to enforce transparency, uphold due process, and provide effective remedies when these technologies cause harm.[68] Drawing lessons from global regulatory efforts such as the European Union's AI Act, South Africa is well-positioned to adopt a rights-based approach that fosters innovation while embedding strong legal protections against harm.[69]

A Comparative Look at How the EU, Canada, and the UK Regulate High-Risk AI in Legal Settings

European Union

The EU AI Act introduces a risk-based framework for regulating artificial intelligence. AI systems are classified into different risk categories, with high-risk systems such as those used in critical infrastructure, law enforcement, or health facing stringent requirements.[70] These include obligations around

[57] Sagar Patel, "Safeguarding Sensitive Data in LLM Race – Awareness and Protection."58.

[58] Sagar Patel, "Safeguarding Sensitive Data in LLM Race – Awareness and Protection." 59.

[59] Vashurina Svetlana, "Trust in Artificial Intelligence: Regulatory Challenges and Prospects," *Legal Issues in the Digital Age*, no. 2 (2025): 69–86.

[60] Svetlana, "Trust in Artificial Intelligence: Regulatory Challenges and Prospects."

[61] K. Calitz, "Combating Fake News through Data Protection: The Potential Role of POPIA," *South African Law Journal* 140, no. 2 (2023): 259–89.Sec 4: This section outlines the foundational principles for processing personal information, including fairness, transparency, and accountability.

[62] Svetlana, "Trust in Artificial Intelligence: Regulatory Challenges and Prospects."

[63] Svetlana, "Trust in Artificial Intelligence: Regulatory Challenges and Prospects."

[64] Mensah, "Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems."

[65] Mensah, "Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems."

[66] Oluwagbade, "Accountability Without Borders: Cross-Jurisdictional Challenges in Regulating AI-Driven Decisions."

[67] Mensah, "Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems."

[68] Oluwagbade, "Accountability Without Borders: Cross-Jurisdictional Challenges in Regulating AI-Driven Decisions."

[69] Oluwagbade, "Accountability Without Borders: Cross-Jurisdictional Challenges in Regulating AI-Driven Decisions." 70.

[70] EU Artificial Intelligence Act: The EU AI Act classifies artificial intelligence systems based on the level of risk they pose; a) Unacceptable risk AI, such as social scoring or systems that manipulate human behaviour is outright banned; b) High-risk AI makes up the bulk of the regulation. These systems, often used in sensitive areas like healthcare, education, or public safety, are subject to strict

transparency, data management, and human oversight to ensure safety and accountability. To support innovation while maintaining public trust, the Act also includes a "sandbox" approach.[71] This allows developers to test AI systems in a controlled environment under the watch of regulators, helping ensure compliance with the law without stifling technological progress. For instance, in *Breyer v Germany*,[72] this case concerned the German government's practice of retaining and processing users' dynamic IP addresses when they visited official websites, where:

> "The applicant raised concerns that such data collection, particularly in the context of AI and automated systems, posed risks to privacy and could amount to mass surveillance. As part of their advocacy, they organised public demonstrations and published articles criticising state surveillance practices. Notably, the first applicant also served as a member of the Schleswig-Holstein State Parliament. In this case, the European Court of Human Rights (ECHR) found that dynamic IP addresses can be classified as personal data, especially when combined with other identifying information. The Court held that such retention must comply with strict data protection standards under EU law. The judgment underscored the importance of transparency, legal safeguards, and accountability when AI systems are used to handle personal data."[73]

While this constituted an interference with the right to privacy,[74] the data collected was limited in scope. Article 8 of the European Convention on Human Rights provides the right to respect private life.[75] The Court further held that the retrieval of such data by authorities was subject to sufficient legal safeguards and oversight.[76] As a result, the interference was found to be proportionate to the legitimate objectives of safeguarding national security and combating crime. The following section deals with artificial intelligence in Canada.

Canada

Moreover, Canada is currently going through a significant period of political upheaval, and amid this turbulence, the proposed Artificial Intelligence and Data Act (AIDA) has effectively stalled following the prorogation of Parliament.[77] Originally introduced in June 2022 in the House of Commons, AIDA aimed to create a broad and comprehensive regulatory framework to govern AI systems nationwide. However, from the outset, the Act faced considerable controversy and debate throughout its time in Parliament.[78] Canada has been working on the Artificial Intelligence and Data Act (AIDA), which sets out rules for high-impact AI systems, such as requiring risk assessments, measures to reduce those risks, and transparency through public reporting. [79] Introduced as part of Bill C-27, the Digital Charter Implementation Act of 2022, AIDA's goal is to promote the safe, transparent, and value-driven development and use of AI across the country.[80] That said, by early 2025, AIDA had encountered notable setbacks. It did not pass into law due to parliamentary prorogation and faced criticism for its unclear scope and the absence of an independent regulatory body to oversee compliance.[81] Nevertheless, the ideas and principles behind AIDA remain an important part of ongoing conversations about responsible AI governance in Canada.

---

rules on transparency, data quality, human oversight, and accountability; c) Limited risk AI is regulated more lightly. Developers and deployers must ensure users are informed when they're interacting with an AI system for example, when using chatbots or encountering deepfakes; d) Minimal risk AI, which includes most current consumer applications like AI in video games or spam filters (as of 2021), is largely unregulated. However, this may change as generative AI technologies become more widespread and integrated into everyday tools.

71 Jon Truby et al., "A Sandbox Approach to Regulating High-Risk Artificial Intelligence Applications," *European Journal of Risk Regulation* 13, no. 2 (2022): 270–94.

72 *Breyer v. Germany*, 2020 (§ 88).

73 *Breyer v. Germany* (n 74). Para 15-16.

74 *Breyer v. Germany* (n 75).

75 European Convention on Human Rights, art 8.

76 *Breyer v. Germany* (n 76). Para 15-16.

77 Blair Attard-Frost, "The Death of Canada's Artificial Intelligence and Data Act: What Happened, and What's next for AI Regulation in Canada," *Montreal AI Ethics Institute*, 2025.

78 Attard-Frost, "The Death of Canada's Artificial Intelligence and Data Act: What Happened, and What's next for AI Regulation in Canada." 79.

79 Teresa Scassa, "Regulating AI in Canada: A Critical Look at the Proposed Artificial Intelligence and Data Act," *Can. B. Rev.* 101 (2023): 1.

80 Scassa, "Regulating AI in Canada: A Critical Look at the Proposed Artificial Intelligence and Data Act." 80.

81 Scassa, "Regulating AI in Canada: A Critical Look at the Proposed Artificial Intelligence and Data Act."82.

Discussion Summary

This paper has explored how artificial intelligence is increasingly shaping South Africa's legal system and the challenges this poses for human rights and justice. It highlights concerns about AI reinforcing bias, reducing transparency, and undermining fair trial rights, particularly for vulnerable groups. By examining South African constitutional principles and comparing AI regulations in the EU and Canada, the study has assessed the readiness of South Africa's legal system to manage these technologies responsibly. The study has further identified gaps in current oversight and risks linked to opaque algorithms and discriminatory data.

RECOMMENDATIONS

To address the issues above, this study proposes the following recommendations:

- To address these concerns, South Africa must establish a robust legal and ethical framework that governs the use of AI in legal contexts. This includes mandatory transparency and explainability standards, independent oversight mechanisms, and regular audits of AI systems.
- Human rights impact assessments should be required before deploying any legal AI tool. Legal practitioners must also be equipped to question and interpret algorithmic outputs.
- An AI regulation tool is required in South Africa. This will prevent the unethical use of AI.
- Moreover, public engagement and civil society input should inform regulatory development.
- By adopting the above measures, South Africa can ensure AI technologies strengthen, rather than undermine, constitutional values and access to justice.

CONCLUSION

The integration of AI into South Africa's legal system presents both opportunities and serious risks. Without clear regulation, algorithmic tools may reinforce systemic bias, obscure accountability, and threaten the right to a fair trial, particularly for historically disadvantaged communities. This comparative study demonstrated the need for South Africa to draw lessons from international approaches while adapting AI regulation to its own constitutional and social realities. It showed that AI could enhance justice delivery when guided by fairness, transparency, and accountability. However, it also revealed that, without adequate oversight, AI risked reinforcing both global and local inequalities. The study also highlighted how the EU and Canada addressed human rights concerns in their digital governance. It emphasised that South Africa needed to establish its own ethical and legal standards for AI within the justice system.

BIBLIOGRAPHY

South African Legislation

Constitution of the Republic of South Africa, 1996.
Criminal Procedure Act 51 of 1977.
Protection of Personal Information Act 4 of 2013.

South African Case Law

Mavundla v MEC Department of Co-Operative Government and Traditional Affairs and Others (7940/2024P) [2025] ZAKZPHC 2 (8 January 2025).
Parker v Forsyth NO and Others (Regional Court, Johannesburg, Gauteng) unreported case no 1585/20 (29 June 2023), reported as Parker v Forsyth NO and Others [2023] ZAGPRD 1.

International Case Law

Breyer v Germany (2020).

International Sources

European Convention on Human Rights, 1950.
European Union Artificial Intelligence Act (EU AI Act), 2024.

Journal Articles and Websites

"AI and Policing: The Benefits and Challenges of Artificial Intelligence for Law Enforcement (Europol, 2024)," n.d.

Ajibola, Temitope. "Navigating Bias and Fairness in AI Algorithms: Pioneering a Human-Centric Revolution." *Medium*, 2023.

Almeida, Patricia Gomes Rêgo de, Carlos Denner dos Santos, and Josivania Silva Farias. "Artificial Intelligence Regulation: A Framework for Governance." *Ethics and Information Technology* 23, no. 3 (September 21, 2021): 505–25. https://doi.org/10.1007/s10676-021-09593-z.

Amalaraj, Perkin. "Decision by Judge to Use ChatGPT to Reach a Verdict Sparks Fury," 2025. https://www.llex.hk/news/other/ridiculous-judge-uses-chatgpt-to-help-reach-a-case-verdict/ar-AA1ogb0L.

Attard-Frost, Blair. "The Death of Canada's Artificial Intelligence and Data Act: What Happened, and What's next for AI Regulation in Canada." *Montreal AI Ethics Institute*, 2025.

Calitz, K. "Combating Fake News through Data Protection: The Potential Role of POPIA." *South African Law Journal* 140, no. 2 (2023): 259–89.

Chen, Zhisheng. "Ethics and Discrimination in Artificial Intelligence-Enabled Recruitment Practices." *Humanities and Social Sciences Communications* 10, no. 1 (2023): 1–12.

Eck, Michele van. "Error 404 or an Error in Judgment? An Ethical Framework for the Use of ChatGPT in the Legal Profession." *Journal of South African Law/Tydskrif Vir Die Suid-Afrikaanse Reg* 2024, no. 3 (2024): 469–90.

Ganai, Parvaiz A, and Irshad A Naikoo. "The Ethical Paradox of AI-Generated Texts: Investigating the Moral Responsibility in Generative Models," 2025.

Larsen, Nancy, and Husam Rajab. "Ensuring Fairness and Mitigating Bias in AI Algorithms," 2024.

Lexis Nexis South Africa. "LexisNexis South Africa Highlights Responsible AI Integration in Law," January 2025. https://www.lexisnexis.com/blogs/za/b/legal/posts/lexisnexis-south-africa-highlights-responsible-ai-integration-in-law.

Lisinski, Ryszard. "Regulation of Artificial Intelligence in South Africa: Progress, Challenges and the Road Ahead." Fluxman, 2025. https://www.fluxmans.com/article/ai-regulation-south-africa.

Menon, Sunita. "Postcolonial Differentials in Algorithmic Bias: Challenging Digital Neo-Colonialism in Africa." *SCRIPTed* 20 (2023): 383.

Mensah, George Benneh. "Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, Transparency, and Accountability in AI Systems." *Preprint, November* 10, no. 1 (2023): 1.

Oluwagbade, Elizabeth. "Accountability Without Borders: Cross-Jurisdictional Challenges in Regulating AI-Driven Decisions," n.d.

Pasipamire, Notice, and Abton Muroyiwa. "Navigating Algorithm Bias in AI: Ensuring Fairness and Trust in Africa." *Frontiers in Research Metrics and Analytics* 9 (2024): 1486600.

Petse, Isiphile, and Usenathi Phindelo. "The Influence of ChatGPT-Generated Data on the Administration of Justice in South Africa." *Turf Law Journal*, 2025.

Raji, Inioluwa Deborah, Andrew Smart, Rebecca N White, Margaret Mitchell, Timnit Gebru, Ben Hutchinson, Jamila Smith-Loud, Daniel Theron, and Parker Barnes. "Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 33–44, 2020.

Sagar Patel. "Safeguarding Sensitive Data in LLM Race – Awareness and Protection," 2025. https://doi.org/10.34218/IJAIRD_03_02_001.

Scassa, Teresa. "Regulating AI in Canada: A Critical Look at the Proposed Artificial Intelligence and Data Act." *Can. B. Rev.* 101 (2023): 1.

Secretariat. "AI Adoption Surges in the Legal Industry: Key Findings from the 2025 Secretariat and ACEDS Global Artificial Intelligence Report," 2025. https://secretariat-intl.com/insights/ai-adoption-surges-in-the-legal-industry/.

Shashvat, Tiwari. " Legal Implications of AI in Judicial Decision-Making." *International Journal of Legal and Law Research*, 2024.

South Africa. *The Constitution of the Republic of South Africa*. Juta Law Firm: Juta, 2009.

South African Police Service reliance risk, SD Law Cape Town Attorneys. "Predictive Policing Employs Algorithms … Historical Crime Data—"AI in Law Enforcement," 2025. https://www.sdlaw.co.za/articles/ai-in-law-enforcement/?utm_source.

Stone, Kelly. "Responsible Use of Artificial Intelligence for Crime Prevention in South Africa: A Topical Guide." *Department of Science & Innovation (South Africa)*, n.d.

Svetlana, Vashurina. "Trust in Artificial Intelligence: Regulatory Challenges and Prospects." *Legal Issues in the Digital Age*, no. 2 (2025): 69–86.

Truby, Jon, Rafael Dean Brown, Imad Antoine Ibrahim, and Oriol Caudevilla Parellada. "A Sandbox Approach to Regulating High-Risk Artificial Intelligence Applications." *European Journal of Risk Regulation* 13, no. 2 (2022): 270–94.

Zhou, Nengfeng, Zach Zhang, Vijayan N Nair, Harsh Singhal, Jie Chen, and Agus Sudjianto. "Bias, Fairness, and Accountability with AI and ML Algorithms." *ArXiv Preprint ArXiv:2105.06558*, 2021.

ABOUT AUTHORS

Bulelani Thukuse (LLB) (LLM Candidate) School of Law, Faculty of Law, Humanities & Social Sciences, Walter Sisulu University, South Africa. He also serves as a reviewer for the South African Journal of Agricultural Extension (SAJAE). His research interests cut across International Criminal Law, Human Rights Law, and Labour Law, particularly issues relating to minimum wages as well as selected aspects of Private Law.

Professor Paul S. Masumbe (LLD) is Associate Professor and Head of School of Law, Walter Sisulu University. He is external examiner for masters and doctoral degrees in law at various universities, reviewer and editorial board member of local and internal journals. His research interests include International Criminal Law, Human Rights Law, Labour Law with emphasis on dismissal, and selected areas of Private Law.