Artificial Intelligence and Transparency: Addressing Risks, Inclusivity, Human Rights and Algorithmic Bias through Ethical Impact Assessment (EIA)

Kariuki Muigua

Table of Contents

Artificial Intelligence and Transparency: Addressing Risks, Inclusivity, Human Rights and Algorithmic Bias through Ethical Impact Assessment (EIA)

Kariuki Muigua*

Abstract

This paper critically examines how Ethical Impact Assessment (EIA) can be utilised to address risks and challenges in AI. The paper defines EIA. It observes that EIA provides a practical and valuable approach towards managing some of the key risks and challenges in AI. The paper identifies the major ethical risks and challenges associated with the use of AI. It also discusses how the use of EIA can address these risks including transparency, inclusivity, human rights and algorithmic bias.

1.0 Introduction

Artificial Intelligence (AI) is a having a transformative impact on nations, societies, businesses and individuals all over the world. It has been observed that harnessing and integrating AI into decision-making processes leads to more data-driven decision-making, thereby enhancing efficiency, strategic planning and execution[1]. AI tools and systems provide numerous benefits for society and the economy, including improving efficiency of processes in both public and private spheres, fostering innovation, assisting in risky tasks, and addressing complex global challenges[2].

The United Nations acknowledges that AI is a powerful technology that can accelerate progress towards most of the Sustainable Development Goals (SDGs)[3]. For example, AI can promote inclusivity, foster innovation and reduce inequalities within and among nations thus providing opportunities to foster Sustainable Development[4]. In particular, it has been observed that AI is improving healthcare, finance, education, agriculture, and environmental governance including climate action thus fostering progress towards the SDGs[5].

Despite its role in driving Sustainable Development, it has been observed that harnessing the benefits of AI depend on managing risks and concerns including skewed data in AI systems that can cause harmful decisions, lack of transparency and accountability in AI tools, and the likelihood to widen digital divides

* PhD in Law (Nrb),SC, FCIArb (Chartered Arbitrator),OGW, LL. B (Hons) Nrb, LL.M (Environmental Law) Nrb; Dip. In Law (KSL); FCPS (K); Dip. in Arbitration (UK); MKIM; Mediator; Consultant: Lead expert EIA/EA NEMA; BSI ISO/IEC 27001:2005 ISMS Lead Auditor/ Implementer; ESG Consultant; Advocate of the High Court of Kenya; Professor of Environmental Law and Conflict Management at the University of Nairobi, Faculty of Law; Member of the Permanent Court of Arbitration (PCA) [June, 2026].

[1] Gabelaia. I., & Hendieh. J., 'The transformative power of AI and its impact on business strategy, financial operations, and marketing decision-making: a case study method' *International Journal of Innovation Science*., Volume 18, Issue 3 (2026)

[2] Benefits and Risks of AI., Available at https://www.consilium.europa.eu/en/policies/benefits-and-risks-of-ai/ (Accessed on 29/06/2026)

[3] United Nations., 'Artificial Intelligence (AI)' Available at https://www.un.org/en/global-issues/artificial-intelligence (Accessed on 29/06/2026)

[4] Ibid

[5] Ibid

and propagate errors[6]. The United Nations points out that without adequate safeguards, AI could further exacerbate inequalities and digital divides, undermine human rights and disproportionately affect the most vulnerable[7]. Harnessing the transformative power of AI therefore depends on how the underlying risks and challenges are addressed.

This paper critically examines how Ethical Impact Assessment (EIA) can be utilised to address risks and challenges in AI. The paper defines EIA. It observes that EIA provides a practical and valuable approach towards managing some of the key risks and challenges in AI. The paper identifies the major ethical risks and challenges associated with the use of AI. It also discusses how the use of EIA can address these risks including transparency, inclusivity, human rights and algorithmic bias.

2.0 Risks and Ethical Concerns in Artificial Intelligence

Despite its transformative potential, the rapid growth and deployment of AI raises several risks and ethical concerns. For instance, transparency is a fundamental ethical risk and concern in AI. It has been observed that algorithmic decision-making through AI makes it difficult to understand how decisions are made which undermines transparency, accountability and due process requirements under the rule of law[8]. In particular, the integration of AI and other technologies into the public sphere raises concerns over transparency and accountability due to the secretive nature of algorithmic decision-making[9]. Due to challenges in explaining, managing and regulating AI outcomes, transparency is pertinent concern while harnessing this transformative technology[10]. Consequently, it has been observed that improving people's understanding about how AI models are trained, the data they are trained on and how they determine outcomes builds trust in AI decisions towards transparency and accountability[11].

Inclusivity is another major ethical concern in AI. In particular, it has been observed that the digital divide undermines the ability of developing countries to harness AI and other technologies leading to AI and technology gaps that do not take into account inclusivity[12]. Digital divide refers to inequality in access, affordability and quality of modern technologies and digital tools including AI[13]. It has been observed that billions of people all over the world lack access to robust, affordable, sustainable and inclusive internet connectivity and modern technologies and digital tools with many more lacking basic digital literacy skills[14].

---

[6] Organisation for Economic Co-operation and Development., 'Governing with Artificial Intelligence' Available at https://www.oecd.org/en/publications/2025/06/governing-with-artificial-intelligence_398fa287.html (Accessed on 29/06/2026)

[7] United Nations., 'Artificial Intelligence (AI)' Op Cit

[8] Burgess. P., 'AI and the Rule of Law' Available at https://www.bloomsbury.com/us/ai-and-the-rule-of-law-9781509963171/ (Accessed on 29/06/2026)

[9] Key Human Rights Challenges of AI., Available at https://ennhri.org/ai-resource/key-human-rights-challenges/ (Accessed on 29/06/2026)

[10] What is AI Transparency?., Available at https://www.ibm.com/think/topics/ai-transparency (Accessed on 29/06/2026)

[11] Ibid

[12] United Nations., 'Artificial Intelligence (AI)' Op Cit

[13] Muller. C., 'What is the Digital Divide?' Available at https://www.internetsociety.org/blog/2022/03/what-is-the-digital-divide/#:~:text=The%20digital%20divide%20is%20the%20gap%20between,**Health%20disparities**%20*%20**Social%20determinants%20of%20health (Accessed on 29/06/2026)

[14] United Nations Habitat., 'Addressing the Digital Divide: Taking Action towards Digital Inclusion' Available at https://unhabitat.org/sites/default/files/2021/11/addressing_the_digital_divide.pdf (Accessed on 29/06/2026)

In particular, Africa and the Global South face challenges such as underdeveloped digital infrastructure, lack of accessible, affordable and quality internet connectivity, a wide digital gender gap, limited digital skills among individuals and organisations, and inadequate regulatory and policy environments causing inequalities and injustices in the digital age[15]. The digital divide is evident in the development and adoption of AI with most models and policies being shaped by the Global North[16]. This challenge undermines inclusivity and the ability of the Global South to shape its own technology future.

The widespread adoption and use of AI can also potentially undermine human rights. For example, it has been observed that unregulated development and misuse of AI can potentially impact human dignity including through perpetuating biases, subjecting individuals and communities to surveillance and worsening inequalities and discrimination[17]. In addition, it has been pointed out that algorithmic decision-making in key areas such as finance, healthcare, hiring and justice systems reduces human beings to mere data points undermining the very nature of human dignity[18]. AI and automation is also disrupting the global job market and significantly impacting the right to fair and decent work[19]. In addition, since most AI models require extensive and unrestricted access to data, it may be challenging to determine what personal data is collected, how it is used, and for what specific purposes such data was used leading to data privacy and security concerns[20]. It has been observed that if the data used in AI models is not properly secured, this data could be vulnerable to breaches or unauthorized access undermining the right to privacy[21].

Algorithmic bias is another fundamental risk and ethical concern in AI that can undermine human rights and dignity. It has been correctly observed that AI models are only as good as the data they are trained on and therefore, if the data is biased, incomplete, or inaccurate, the AI model's predictions and decisions will also be biased, incomplete, or inaccurate[22]. It has been observed that due to historically discriminatory

[15] World Bank Group., 'Digital Transformation Drives Development in Africa' Available at https://www.worldbank.org/en/results/2024/01/18/digital-transformation-drives-development-in-afe-afw-africa (Accessed on 29/06/2026)

[16] Png. M-T., 'The Critical Roles of Global South Stakeholders in AI Governance' Available at https://academic.oup.com/edited-volume/41989/chapter-abstract/377785114?redirectedFrom=fulltext (Accessed on 29/06/2026)

[17] Randazzo. M.S., & Hill. G., 'Human dignity in the age of Artificial Intelligence: an overview of legal issues and regulatory regimes' Available at https://www.tandfonline.com/doi/full/10.1080/1323238X.2025.2483822#:~:text=Contributing%20to%20the%20emergent%20scholarship,fundamental%20rights%20and%20democratic%20values. (Accessed on 29/06/2026)

[18] Ibid

[19] World Economic Forum., 'How are Today's Biggest Tech Trends Affecting our Human Rights?' Available at https://www.weforum.org/agenda/2017/12/how-are-today-s-biggest-tech-trends-affecting-human-rights/ (Accessed on 29/06/2026)

[20] European Data Protection Supervisor., 'Agentic AI' Available at https://www.edps.europa.eu/data-protection/technology-monitoring/techsonar/agentic-ai_en (Accessed on 29/06/2026)

[21] Hoppin. B., 'Ethics in Online Dispute Resolution: Balancing Technology and Human Judgment' Available at https://odr.com/ethics-in-online-dispute-resolution-balancing-technology-and-human-judgment/#:~:text=3.,refine%20AI%20models%20without%20consent (Accessed on 29/06/2026)

[22] Vinciullo. A., 'The Risk of Law Firms Relying on Chat GPT to Perform Contract analysis and Legal Research.' Available at https://www.linkedin.com/pulse/risk-law-firms-relying-chat-gpt-perform-contract-legal-vinciullo/ (Accessed on 29/06/2026)

and biased datasets, AI is fuelling racial, gender and language among other baises[23]. For instance, it has been observed that people have been excluded from loan eligibility and job opportunities on basis of their race and gender, with mass surveillance through AI tools also contributing to racial profiling[24]. In particular, it has been observed that since most AI models are designed in the Global North, they may perpetuate bias, discrimination and prejudices in the Global South unless they are tuned to fit with local circumstances[25]. The use of imported AI models in Africa and the Global South is causing biases and discrimination including through exclusion of indigenous data sets, inadequate gender, racial and language considerations, racial profiling in criminal justice and law enforcement and mass surveillance among other forms of algorithmic bias[26].

In light of the foregoing risks and concerns, it is imperative to ensure safe, ethical and appropriate adoption of AI including through fostering transparency, upholding human rights, ensuring inclusivity and tackling algorithmic biases.

3.0 Utilising Ethical Impact Assessment to Foster Transparency in Artificial Intelligence

Ethical Impact Assessment (EIA) refers to a structured and systematic approach used to evaluate the ethical and moral implications of technology and innovation in order to ensure the safe, appropriate, responsive and purposeful adoption and use of modern technologies[27]. It has been observed that EIA involves identifying and evaluating both current and future ethical risks and concerns, particularly in relation to emerging technologies, in order to strengthen governance and decision-making processes[28]. Through EIA, it is possible to identify, understand, and address the potential ethical consequences of modern technologies before and during deployment[29].

EIA is particularly important in the current era of AI. For example, it has been observed that EIA examines and evaluates how AI systems and models may affect individuals, communities and nations, especially regarding fairness, transparency, autonomy, discrimination, bias and power imbalance[30]. In particular, EIA enables decision-makers to examine both positive and negative impacts of an AI system, identify and engage with relevant stakeholders, and develop mitigation strategies in order to address adverse ethical outcomes[31]. It has been observed that EIA can be utilised in both public and private sectors to ensure the

[23] Olurunju. N., 'African algorithmic governance: Benefit of a community-based approach' Available at https://researchictafrica.net/2022/04/03/african-algorithmic-governance-benefit-of-a-community-based-approach/ (Accessed on 29/06/2026)
[24] Ibid
[25] Orero. L.O., & Kaaniru. J., 'Automated Decision-Making Policies in Africa' Available at https://cipit.strathmore.edu/wp-content/uploads/2023/08/Policy-Brief-Design-Automated-Decision-Making.pdf (Accessed on 29/06/2026)
[26] Olurunju. N., 'African algorithmic governance: Benefit of a community-based approach' Op Cit
[27] Ethical Impact Assessment: Theoretical Foundations, Methodology, and Applications., Available at https://link.springer.com/article/10.1007/s11948-025-00574-9 (Accessed on 29/06/2026)
[28] Ibid
[29] Ethics Impact Assessment for AI., Available at https://verifywise.ai/lexicon/ethics-impact-assessments (Accessed on 29/06/2026)
[30] Ibid
[31] United Nations Educational, Scientific and Cultural Organization., 'Ethical Impact Assessment' Available at https://www.unesco.org/ethics-ai/en/eia (Accessed on 29/06/2026)

ethical design, development, and deployment of an AI system[32]. EIA is a tool that can be utilised by various stakeholders including governments, AI developers, businesses, researchers and academics in order to ensure the safe, ethical and appropriate development, deployment and use of AI in both public and private spheres[33].

It has been observed that EIA is a necessary tool since AI models and systems are not inherently neutral[34]. AI tools usually reflect the values and interests of those who design and deploy them[35]. For instance, it has been correctly pointed out that since most AI models are designed in the Global North, they may perpetuate bias, discrimination and prejudices in the Global South unless they are tuned to fit with local circumstances[36]. The use of imported AI models in Africa and the Global South is causing biases and discrimination including through exclusion of indigenous languages and data sets, inadequate gender and racial considerations, racial profiling in criminal justice and law enforcement and mass surveillance[37]. EIA therefore provides a practical approach towards the safe, ethical, appropriate and inclusive adoption of AI, especially in the Global South, in line with the promotion and protection of human rights and human dignity, environmental sustainability, fairness, justice, inclusion and gender equality[38].

In light of the current risks and challenges in AI including transparency and accountability concerns, the potential for human rights violations, lack of inclusivity and algorithmic biases, there is need to embrace EIA in order to ensure that this technology is harnessed in a safe and ethical manner. It has been observed that EIA considers the ethical risks throughout the entire AI lifecycle.[39] EIA therefore covers both *ex-ante* (prior) and *ex-post* (later) requirements[40]. At the *ex-ante/initial stage*, before the development and deployment of AI systems, EIA focuses on the importance of ensuring quality and representativeness data in order to avoid algorithmic biases, the diversity of the teams developing AI systems in order ensure inclusivity, the robustness and transparency of the algorithms in line with human rights and their auditability in order to identify and address risks and challenges including human rights violations and algorithmic biases[41]. In addition, at the *ex-post stage*, EIA involves regular monitoring and human review of AI tools and systems to identify and address new and emerging risks[42].

By effectively embracing EIA, it is possible to address the ethical risks and challenges in AI. For instance, through EIA, it is possible to address algorithmic biases and ensure fairness by ensuring that AI systems

[32] Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence., Available at https://my.ai.se/resources/ethical-impact-assessment-a-tool-of-the-recommendation-on-the-ethics-of-artificial-intelligence-1 (Accessed on 29/06/2026)
[33] United Nations Educational, Scientific and Cultural Organization., 'Ethical Impact Assessment' Op Cit
[34] Ibid
[35] Ibid
[36] Orero. L.O., & Kaaniru. J., 'Automated Decision-Making Policies in Africa' Op Cit
[37] Olurunju. N., 'African algorithmic governance: Benefit of a community-based approach' Op Cit
[38] United Nations Educational, Scientific and Cultural Organization., 'Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence' Available at https://www.unesco.org/en/articles/ethical-impact-assessment-tool-recommendation-ethics-artificial-intelligence (Accessed on 29/06/2026)
[39] United Nations Educational, Scientific and Cultural Organization., 'Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence' Op Cit
[40] Ibid
[41] Ibid
[42] Ibid

are trained on accurate, inclusive, fair and just data sets that take into account all sectors of the society[43]. Further, through EIA, it is possible to foster transparency and explainability by ensuring that those affected can understand the rationale of algorithmic decision-making[44]. In addition, it has been observed that EIA ensures accountability including through determining who is responsible if harm or human rights violations occur[45].

EIA is therefore a key feature of sound AI governance. AI governance aims to promote the safe development and adoption of AI that is universal, adapted to cultural diversities all over the world, free from biases and discrimination, and respectful of democratic values and fundamental rights and freedoms for every person[46]. It has been observed that through appropriate AI governance, it is possible to direct AI research, development and adoption towards ensuring safety, ethics, fairness and respect for human rights[47]. EIA can bolster AI governance by inculcating ethical requirements and standards throughout the AI lifecycle including research, design, development, deployment, use, maintenance, operation, monitoring, and eventual termination[48].

Adopting EIA is therefore necessary towards bolstering AI governance.

4.0 Conclusion

In light of the widespread development and adoption of AI, it is imperative to strengthen its governance in order to harness its benefits while mitigating risks and challenges. In particular, EIA provides a valuable and practical approach towards examining both positive and negative impacts of AI systems in order to ensure the safe, ethical, appropriate and purposeful adoption of this transformative technology[49]. It is therefore imperative to widely embrace and utilise EIA towards strengthen governance of AI by fostering transparency, addressing risks, ensuring inclusivity, protecting human rights and avoiding algorithmic bias.

[43] Ethics Impact Assessment for AI., Op Cit
[44] Ibid
[45] Ibid
[46] French Institute of International Relations., 'Artificial Promises or Real Regulation? Inventing Global AI Governance' Available at https://www.ifri.org/en/studies/artificial-promises-or-real-regulation-inventing-global-ai-governance (Accessed on 29/06/2026)
[47] What is AI Governance?., Available at https://www.ibm.com/think/topics/ai-governance (Accessed on 29/06/2026)
[48] United Nations Educational, Scientific and Cultural Organization., 'Ethical Impact Assessment' Op Cit
[49] Ibid

References

Benefits and Risks of AI., Available at https://www.consilium.europa.eu/en/policies/benefits-and-risks-of-ai/

Burgess. P., 'AI and the Rule of Law' Available at https://www.bloomsbury.com/us/ai-and-the-rule-of-law-9781509963171/

Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence., Available at https://my.ai.se/resources/ethical-impact-assessment-a-tool-of-the-recommendation-on-the-ethics-of-artificial-intelligence-1

Ethical Impact Assessment: Theoretical Foundations, Methodology, and Applications., Available at https://link.springer.com/article/10.1007/s11948-025-00574-9

Ethics Impact Assessment for AI., Available at https://verifywise.ai/lexicon/ethics-impact-assessments

European Data Protection Supervisor., 'Agentic AI' Available at https://www.edps.europa.eu/data-protection/technology-monitoring/techsonar/agentic-ai_en

French Institute of International Relations., 'Artificial Promises or Real Regulation? Inventing Global AI Governance' Available at https://www.ifri.org/en/studies/artificial-promises-or-real-regulation-inventing-global-ai-governance

Gabelaia. I., & Hendieh. J., 'The transformative power of AI and its impact on business strategy, financial operations, and marketing decision-making: a case study method' *International Journal of Innovation Science*., Volume 18, Issue 3 (2026)

Hoppin. B., 'Ethics in Online Dispute Resolution: Balancing Technology and Human Judgment' Available at https://odr.com/ethics-in-online-dispute-resolution-balancing-technology-and-human-judgment/#:~:text=3.,refine%20AI%20models%20without%20consent

Key Human Rights Challenges of AI., Available at https://ennhri.org/ai-resource/key-human-rights-challenges/

Muller. C., 'What is the Digital Divide?' Available at https://www.internetsociety.org/blog/2022/03/what-is-the-digital-divide/#:~:text=The%20digital%20divide%20is%20the%20gap%20between,**Health%20disparities**%20*%20**Social%20determinants%20of%20health

Olurunju. N., 'African algorithmic governance: Benefit of a community-based approach' Available at https://researchictafrica.net/2022/04/03/african-algorithmic-governance-benefit-of-a-community-based-approach/

Orero. L.O., & Kaaniru. J., 'Automated Decision-Making Policies in Africa' Available at https://cipit.strathmore.edu/wp-content/uploads/2023/08/Policy-Brief-Design-Automated-Decision-Making.pdf

Organisation for Economic Co-operation and Development., 'Governing with Artificial Intelligence' Available at https://www.oecd.org/en/publications/2025/06/governing-with-artificial-intelligence_398fa287.html

Png. M-T., 'The Critical Roles of Global South Stakeholders in AI Governance' Available at https://academic.oup.com/edited-volume/41989/chapter-abstract/377785114?redirectedFrom=fulltext

Randazzo. M.S., & Hill. G., 'Human dignity in the age of Artificial Intelligence: an overview of legal issues and regulatory regimes' Available at https://www.tandfonline.com/doi/full/10.1080/1323238X.2025.2483822#:~:text=Contributing%20to%20the%20emergent%20scholarship,fundamental%20rights%20and%20democratic%20values

United Nations Educational, Scientific and Cultural Organization., 'Ethical Impact Assessment' Available at https://www.unesco.org/ethics-ai/en/eia

United Nations Educational, Scientific and Cultural Organization., 'Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence' Available at https://www.unesco.org/en/articles/ethical-impact-assessment-tool-recommendation-ethics-artificial-intelligence

United Nations Habitat., 'Addressing the Digital Divide: Taking Action towards Digital Inclusion' Available at https://unhabitat.org/sites/default/files/2021/11/addressing_the_digital_divide.pdf

United Nations., 'Artificial Intelligence (AI)' Available at https://www.un.org/en/global-issues/artificial-intelligence

Vinciullo. A., 'The Risk of Law Firms Relying on Chat GPT to Perform Contract analysis and Legal Research.' Available at https://www.linkedin.com/pulse/risk-law-firms-relying-chat-gpt-perform-contract-legal-vinciullo/

What is AI Governance?., Available at https://www.ibm.com/think/topics/ai-governance

What is AI Transparency?., Available at https://www.ibm.com/think/topics/ai-transparency

World Bank Group., 'Digital Transformation Drives Development in Africa' Available at https://www.worldbank.org/en/results/2024/01/18/digital-transformation-drives-development-in-afe-afw-africa

World Economic Forum., 'How are Today's Biggest Tech Trends Affecting our Human Rights?' Available at https://www.weforum.org/agenda/2017/12/how-are-today-s-biggest-tech-trends-affecting-human-rights/