Artificial Intelligence Safety in Development Work: A Responsibility We Cannot Ignore

by **Abisoye Ajose** *[July 2026]*



AI Safety image [Source: iStock]

Introduction

According to the International Organisation for Standardisation, Artificial Intelligence (AI) refers to computer systems capable of performing tasks that typically require human intelligence, such as reasoning, learning, perception and language understanding.[1]

Artificial Intelligence is estimated to yield a revenue of $1.2 trillion to the African economy, which would represent a 5.6% boost to the continent's total economic output by the end of the decade.[2]

[1] International Organization for Standardization (ISO) (n.d.) *Artificial intelligence: What it is, how it works and why it matters.* (Accessed: 23 July 2026).

[2] United Nations Economic Commission for Africa (UNECA) (2024) Artificial intelligence and Africa. Africa Renewal.

Artificial intelligence is no longer a futuristic concept confined to research labs and tech giants. Artificial Intelligence tools are transforming how Development Organisations across Africa collect, manage, analyse data, and prepare reports. From accelerating data collection to uncovering insights that drive impactful programmes, AI is enabling positive change and innovation at an unprecedented pace in the Development sector in Africa.

Yet, with such transformative potential comes the need for caution. There is a vital question that the Development sector has been slow to answer: Are we using AI safely, ethically, and responsibly? The stakes in development work are too high to get this wrong.

How AI Is Gaining Popular Use in Development Work

The appeal of AI in development work is easy to understand. Generative AI tools such as Claude AI and Notebook LM, along with predictive analytics platforms and automated data systems, can achieve in minutes what once took weeks.

Development practitioners are using AI to synthesize research, construct outcome variables, and extract data.[3]

In Uganda, the United Nations is applying natural language processing to radio broadcasts as a way of tracking public sentiment and evaluating how well its programmes are landing. Radio recordings gathered in Luganda and Acholi were first transcribed into text using automatic speech recognition.[4] The team then scanned the resulting transcripts for keywords and phrases tied to three areas of interest: attitudes toward refugees, the effects of disasters on livelihoods and health, and how effective radio-based awareness campaigns had been.[5]

Another case in point is Planet Labs, which uses satellite imagery to track agricultural development[6] in Kenya, flooding in Sri Lanka, and the expansion of informal settlements[7] in Dar es Salaam. In Tanzania specifically, the company trained an AI model to detect areas where the built-up footprint was expanding, then paired this with 3D reconstructions of the satellite images to estimate building heights.[8] Together, these methods allowed them to track population shifts and feed that insight into urban planning and policy decisions.

Donor pressure to demonstrate efficiency and shrinking funding space are accelerating this trend,[9] and for many development organisations operating in under-resourced contexts, AI feels like a force multiplier, a way to extend limited capacity and reach more people.[10]

Yet adoption is often outpacing understanding of the potential risks of AI tools. Development practitioners are integrating AI tools with minimal training, limited institutional guidance, and little awareness of the risks embedded in these systems

Recommendations to Foster AI Safety in Development Work

Responsible AI use in Development work requires deliberate action, not good intentions alone.

Here are four strategies that organisations should embed into their practice.

Understand the Foundations of AI Tools: Before deploying any AI system, Development organisations must ask: What data was this model trained on? Does it reflect the communities we serve, or does it encode biases that could distort our analysis and decision-making?

A model trained predominantly on data from the Global North may produce outputs that are irrelevant or harmful when applied in different cultural and geographic contexts.

Integrate Human Oversight in AI-Driven Decisions: Artificial Intelligence should augment professional judgment, not replace it. High-stakes decisions such as targeting beneficiaries, assessing vulnerability, and allocating resources must remain under meaningful human oversight. Organisations

[3] 3ie (2024) *Making AI work: Practical applications in impact evaluation.* 3ie Blog. (Accessed: 23 July 2026).
[4] Pulse Lab Kampala (2017) *Using machine learning to accelerate sustainable development solutions in Uganda.* UN Global Pulse.(Accessed 23 July 2026)
[5] Fu, G. (2022) *Toward ethical artificial intelligence in international development.* Center for Digital Acceleration Insights.(Accessed: 23 July 2026).
[6] Vilaca, N. (2016) *Financing smallholder farms with FarmDrive.* Planet.(Accessed: 23 July 2026).
[7] O'Shea, T. (2018) *New project from World Bank & Planet uses satellite imagery & machine learning to drive sustainable urbanization in Tanzania.* Planet.(Accessed: 23 July 2026).
[8] Fu, G. (2022) *Toward ethical artificial intelligence in international development.* Center for Digital Acceleration Insights.(Accessed: 23 July 2026).
[9] Singh, M. (2026) *AI for nonprofits: the smarter way to scale fundraising impact, not costs.* Flexsin Technologies.(Accessed: 24 July 2026)
[10] Van Osch, M. (2026) *Empowering nonprofits with AI: the complete 2026 guide.* Zenphi. Available at: (Accessed: 24 July 2026).

should establish clear protocols for when AI outputs require human review, and ensure that frontline staff have the authority and capacity to question or override automated recommendations.

Promote Data Protection Practices:
Development work involves some of the most sensitive data imaginable: health records, displacement histories, financial profiles of people in crisis. Using AI tools, particularly cloud-based platforms without robust data governance frameworks, puts beneficiaries at risk. Organizations must conduct data protection impact assessments before deploying AI, and ensure compliance with relevant frameworks such as the GDPR or national data protection laws.

Build Institutional Literacy, Not Just Individual Skills:
AI safety cannot rest on the shoulders of a few tech-savvy staff members. It requires organization-wide literacy from leadership setting policy to field staff applying tools. Investing in structured AI training, developing internal usage guidelines, and creating feedback mechanisms for staff to flag concerns are all essential components of a safety-conscious culture for AI usage.

Engage Regulators Jointly:

Organisations should work collectively with regulators in their respective countries such as Kenya's Office of the Data Protection Commissioner and South Africa's Information Regulator to develop sector-specific guidance and codes of conduct on the use of AI with vulnerable populations.

Strengthening Collective Accountability in AI Governance

Artificial Intelligence offers genuine promise for a sector grappling with complex problems and constrained resources. A promise without prudence is a liability.

Development Organisations in Africa have a duty of care to their beneficiaries, partners, and communities to do more than just adopt AI. They should make accountability a priority, ask key questions before using these tools, and set clear policies to guide staff in using AI ethically.

Development organisations must move beyond organisation-specific responsible AI policies towards a coordinated, sector-wide approach to AI governance. This requires addressing existing power imbalances by ensuring that local communities, civil society organisations and governments, particularly in the Global South, have a meaningful voice in decisions about how AI is designed and deployed.

Responsible AI frameworks must prioritise the rights and interests of affected communities rather than primarily serving organisational or commercial interests. Development actors, donors and technology providers should collectively establish standards for transparency, data protection, bias mitigation and accountability, supported by independent audits and clear mechanisms for reporting AI-related harms. Ultimately, AI safety must become a shared sectoral responsibility, ensuring that technological advancement does not deepen existing inequalities or undermine the communities development work seeks to serve.

Conclusion

Artificial Intelligence offers genuine promise for a sector grappling with complex problems and constrained resources. A promise without prudence is a liability.

Development Organisations in Africa have a duty of care to their beneficiaries, partners, and communities to do more than just adopt AI. They should make accountability a priority, ask key questions before using these tools, and set clear policies to guide staff in using AI ethically.