

**Book Review of *Trustworthy AI: African Perspectives***

*Reframing Trustworthiness Through Context, Agency, and Technological Sovereignty*

**Sahaja Deverashetty**

*Undergraduate Researcher, Bachelor of Computer Science, Golden Gate University*

**Kshitij Taluja**

*Undergraduate Researcher, Bachelor of Computer Science, Golden Gate University*

**Dr. Smrite Goudhaman**

*Global Professor of Practice, Golden Gate University*

**Editors:** Damian Okaibedi Eke, Kutoma Wakunuma, Simisola Akintoye, and George Ogoh

**Title:** *Trustworthy AI: African Perspectives*

**Publisher:** Palgrave Macmillan

**Year:** 2025

**ISBN:** 978-3-031-75673-3

**Pages:** 281

**Abstract**

Trustworthy AI: African Perspectives, edited by Damian Okaibedi Eke, Kutoma Wakunuma, Simisola Akintoye, and George Ogoh (2025), examines how trustworthy artificial intelligence should be understood when social, cultural, institutional, historical, and material conditions are treated as part of trustworthiness itself. This review argues that the volume's principal contribution is not the proposal of a single African model, but its challenge to universal and predominantly technical definitions of trustworthy AI. Across discussions of agency, decoloniality, accessibility, affordability, data governance, and technological sovereignty, the volume shows why context matters to both the design and evaluation of AI systems. Its main limitation is more specific than a lack of empirical evidence: evidence about how these principles can be operationalized is uneven across the edited collection, and their application is likely to vary substantially among African countries, communities, sectors, and institutions. Read through an action research lens,

the book offers a useful basis for testing trustworthy AI through stakeholder participation, contextual evaluation, and iterative organizational change.

**Keywords:** trustworthy AI; African perspectives; decoloniality; AI governance; technological sovereignty

Trustworthy AI: African Perspectives starts from a deceptively simple premise: an AI system cannot be judged trustworthy independently of the environment in which it is developed and used. The edited volume brings together conceptual, normative, and empirical contributions on governance, healthcare, road safety, gender, data justice, technology studies, and AI design. What connects these varied chapters is an emphasis on context, agency, and power. The volume therefore expands the question of trustworthiness beyond technical performance and asks who participates in AI development, whose values shape the system, what resources are available to use it safely, and who retains meaningful control over data and deployment.

Importantly, the book does not simply reject established trustworthy-AI frameworks. The High-Level Expert Group on Artificial Intelligence (2019), for example, defines trustworthy AI through principles including human agency, privacy, transparency, fairness, safety, well-being, and accountability. Eke et al. (2025) do not dismiss these concerns. Instead, the volume asks whether they are sufficient when affordability, accessibility, local adaptability, infrastructure, history, and technological dependence are taken seriously. This is a more useful argument than a simple contrast between Western and African models because it treats existing frameworks as starting points that may require contextual reinterpretation rather than as approaches that should be discarded.

The book's discussion of decoloniality sharpens this argument by directing attention to ownership and decision-making. Questions about who provides data, who defines design priorities, who controls infrastructure, whose knowledge is represented, and who benefits from deployment connect trustworthy AI to wider debates about algorithmic colonization and decolonial AI (Birhane, 2020; Mohamed et al., 2020). The organizational implications are concrete. During procurement, an organization could evaluate not only price and technical performance but also data location, contractual control, local maintenance capacity, and opportunities for affected stakeholders to influence deployment. During data collection and system design, it could examine who is represented, whose categories are being imposed, and whether communities can contest or revise those choices. During evaluation, it could include accessibility, affordability, and local social effects alongside conventional accuracy measures. In this sense, decoloniality changes operational decisions rather than remaining only a theoretical position.

The volume is also stronger than a purely conceptual collection because several chapters examine particular sectors and institutions, including Nigerian digital-health startups and road-safety applications. These cases show why technical performance cannot be separated from institutional capacity, data practices, ethics oversight, or the conditions of use. The main evidentiary limitation should therefore be stated carefully. The book does not generally lack empirical evidence, and conceptual or normative chapters should not be judged solely by empirical criteria. Rather, the evidence for how proposed principles can be operationalized is uneven across the volume. Questions remain about whether the

same criteria can be implemented consistently across organizations with very different infrastructure, regulatory capacity, resources, and community expectations.

That limitation becomes especially important because 'African perspectives' cannot be treated as a single context. African countries and communities differ in language, legal systems, infrastructure, public-sector capacity, health systems, educational institutions, digital access, and relationships with international technology providers. A governance practice that is workable in one country or sector may be inappropriate in another.

Contextualizing trustworthy AI therefore requires attention not only to differences between Africa and other regions, but also to differences within Africa itself. Comparative and action-oriented research across countries, communities, sectors, and institutions is necessary if the book's principles are to become adaptable practices rather than broad regional claims.

The discussion of technological sovereignty adds another useful layer. The volume emphasizes greater African agency over AI development and use, but sovereignty should not be confused with technological self-sufficiency. Complete independence from international investment, cloud infrastructure, hardware, expertise, or collaboration may be unrealistic and is not required for meaningful control. A more practical understanding of sovereignty concerns who has authority over governance, data, procurement, infrastructure choices, system adaptation, and deployment. An organization may rely on international technologies while still negotiating stronger data protections, requiring local accountability, retaining meaningful procurement choices, and building the capacity to

evaluate or reject systems that do not fit local needs. This distinction makes the book's argument more actionable.

The clearest connection to action research appears when these principles are treated as hypotheses to test in practice. Consider an organization introducing an AI-supported healthcare system. In an initial cycle, clinicians, administrators, patients, and community representatives could define what trustworthiness means locally and establish baseline measures for accessibility, affordability, data governance, clinical usefulness, and user confidence. A limited deployment could then be introduced, with the effects observed across different groups rather than only through aggregate technical performance. The organization could identify intended and unintended consequences, revise procurement terms, workflow design, data practices, or training, and then test the revised arrangement in a second cycle. Repeating this process would make stakeholder participation and contextual trustworthiness observable rather than rhetorical.

This example also integrates the volume's connection to Sustainable Development Goals 9 and 10. Innovation and infrastructure development can expand access to useful technologies, but they can also deepen inequality when benefits depend on costly infrastructure, specialist expertise, proprietary data, or services that remain inaccessible to disadvantaged groups (United Nations, 2015). An action research evaluation should therefore ask two questions at the same time: does the intervention improve capacity or service quality, and who gains or loses access, voice, or control as a result? Evaluating both innovation and inequality outcomes keeps the SDG discussion connected to the book's core concerns with affordability, participation, agency, and institutional capacity.

Overall, *Trustworthy AI: African Perspectives* makes a valuable contribution by changing how the problem of trustworthiness is framed. Its strongest claim is not that there is one complete African model that can replace existing frameworks. It is that trustworthy AI must be developed and evaluated with the communities, institutions, histories, and material conditions in which systems will operate. The volume is less conclusive on how consistently its principles can be translated into practice across highly diverse settings, but this limitation also defines a productive research agenda. For organizations and action researchers, the next step is to test, compare, and refine these principles across different African contexts while preserving local agency, stakeholder participation, and meaningful control over technology.

## References

- Birhane, A. (2020). Algorithmic colonization of Africa. *SCRIPTed*, 17(2), 389-409.  
<https://doi.org/10.2966/scrip.170220.389>
- Eke, D. O., Wakunuma, K., Akintoye, S., & Ogoh, G. (Eds.). (2025). *Trustworthy AI: African perspectives*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-031-75674-0>
- High-Level Expert Group on Artificial Intelligence. (2019). *Ethics guidelines for trustworthy AI*. European Commission. <https://doi.org/10.2759/346720>
- Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33(4), 659-684. <https://doi.org/10.1007/s13347-020-00405-8>
- United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*. United Nations General Assembly. <https://sdgs.un.org/2030agenda>