



RECONCEPTUALIZING RESEARCH ETHICS IN THE AI ERA: A SYSTEMATIC REVIEW FRAMING GENAI AS AN EMERGING RISK

Peruce Atingo, Dr. Rose W. Burugu PhD
Alupe University

ABSTRACT

DOI No: 10.36713/epra31630

Article DOI: <https://doi.org/10.36713/epra31630>

Integrating Generative Artificial Intelligence (GenAI) into qualitative research offers new operational advantages but also raises important ethical concerns. Viewing GenAI in isolation risks oversimplifying research ethics as just technical compliance. This systematic review explores qualitative research ethics by framing GenAI as a new vulnerability layered over traditional principles such as participant autonomy, non-maleficence, researcher positionality, and institutional oversight. Analyzing 18 key studies from global, African, and East African sources using PRISMA guidelines, the review reveals how GenAI worsens existing structural inequalities. Global studies mainly address data privacy and algorithmic transparency, while African and East African research emphasize how LLMs amplify historical issues like digital neo-colonialism, cultural extraction, and the marginalization of indigenous languages. To manage these complex ethical challenges, we recommend targeted strategies tailored to each research context, including a Human-in-the-Loop (HITL) framework based on Afro-communitarian ethics (Ubuntu).

KEYWORDS: Qualitative research ethics, Generative AI, emerging ethical risks, East Africa, data sovereignty, Ubuntu ethics, PRISMA

INTRODUCTION

The core principles of qualitative research emphasize human empathy, relational trust, and deep reflexivity, all aimed at understanding lived human experiences. Recently, traditional ethical standards have been challenged by the rapid rise of Generative Artificial Intelligence (GenAI), Large Language Models (LLMs), and automated analytical tools (Denzin & Lincoln, 2023). Nonetheless, GenAI does not replace traditional research ethics. Instead, it amplifies and distorts them, highlighting an urgent need to rethink ethical frameworks suitable for the digital era.

Assessing this digital shift requires viewing GenAI as a new ethical risk that directly impacts the core qualitative variables underlying research. These include participant autonomy, informed consent, data confidentiality, non-maleficence, researcher positionality, and institutional oversight. Uploading unencrypted transcripts to third-party commercial cloud servers (Mittelstadt, 2019) raises confidentiality issues that challenge traditional agreements. However, when examined through standard ethical lenses, GenAI can offer technological efficiency while still respecting human rights principles.

In the field of empirical research, inquiries concerning algorithmic transparency and data governance have garnered

international attention. The opacity and the complexities involved in comprehending the analytical mechanisms of proprietary commercial large language models (LLMs) have been consistently emphasized in global studies, as has the phenomenon of AI hallucinations (Bender et al., 2021). While these global findings are instrumental in understanding issues related to corporate data compliance and technical reliability, they are predominantly articulated within the framework of the institutional infrastructure of the Global North and may not accurately reflect the distinct socio-cultural, political, and economic contexts of the Global South (Jobin et al., 2019).

In Sub-Saharan Africa, regional scholarship explores the links between GenAI and qualitative research, alongside themes like digital neo-colonialism, epistemic extraction, and data sovereignty. African scholars highlight that most LLMs are trained on Western data, leading to significant algorithmic biases that can obscure or distort African oral traditions, knowledge systems, and social dynamics (Wakunuma & Eke, 2024). The region's literature emphasizes Afro-communitarian moral frameworks, such as Ubuntu.

Local Institutional Review Boards (IRBs) and Research Ethics Committees (RECs) in local research context in East Africa are challenged with operational difficulties in assessing new digital

risk factors and third-party data security protocols in the cloud (Bouhouita-Guermech et al., 2023). This systematic review consolidates global, regional, and local perspectives and offers a governance framework to guide researchers and ethics committees in East Africa.

PROBLEM STATEMENT

The framing of AI ethics exclusively as a matter of software or data security neglects both the methodological rigor inherent in qualitative research and the fundamental human elements of trust and empathy that form its foundation. The emergence of Generative AI tools and Large Language Models (LLMs) does not merely cause minor technical issues; it reveals enduring vulnerabilities within research practices that, if left unaddressed, could precipitate systemic ethical deterioration. Uploading transcripts from field research to commercial cloud services raises serious concerns regarding the improper collection and handling of data by corporations. Furthermore, an overreliance on automated coding replaces genuine qualitative engagement with mere pattern recognition, thereby diminishing interpretive depth and introducing potential biases related to human prejudice and algorithmic discrimination. While methodological and cultural challenges are considerable, the primary concern resides in institutional oversight—many local review boards lack explicitly defined Standard Operating Procedures (SOPs) specifically addressing digital data security and the transfer of data to third-party cloud services.

RESEARCH QUESTIONS

This review focuses on four main research questions:

- **RQ1:** How do emerging GenAI vulnerabilities intersect with classic qualitative ethical principles across the research lifecycle?
- **RQ2:** How do theoretical frameworks conceptualize the combined impact of algorithmic bias, privacy risks, and researcher reflexivity?
- **RQ3:** What empirical evidence distinguishes global ethical concerns from those specific to the African continent and the East African region?
- **RQ4:** What governance frameworks and operational protocols reconcile technological efficiency with ethical integrity in East African qualitative inquiry?

REVIEW OF EXISTING LITERATURE

Emerging AI Risks and Classic Lifecycle Ethics

Recent research shows that Generative AI (GenAI) and Large Language Models (LLMs) are transforming traditional ethical concerns throughout different phases of qualitative research. The literature emphasizes technical risks, including privacy threats to participants, impacts on analysis depth, and issues of academic integrity within a global context. Mittelstadt (2019) points out that uploading data to third-party cloud servers raises privacy issues, particularly when data retention is unmanaged. Bouhouita-Guermech et al. (2023) highlight challenges related to informed consent, especially when personal stories are used to train commercial models. Kuckartz and Rädiker (2024) focus on computer-aided methods and software use in qualitative

research, noting that automated pattern recognition cannot replace the interpretative insights provided by humans.

At the African regional level, researchers often examine vulnerabilities across the entire life course, which are rooted in systemic inequalities. When analyzing local qualitative data with Western-trained epistemic extraction tools using commercial LLMs, Birhane (2021) indicates that such extraction methods can influence the observed patterns. Wakunuma and Eke (2024) caution about the dangers of employing generative AI without proper verification, as this can distort oral traditions and local contexts. Consequently, careful validation is essential when applying generative AI to oral traditions to avoid misrepresentation, as highlighted by Wakunuma and Eke (2024).

Theoretical Frameworks of Bias, Sovereignty, and Reflexivity

Three primary theoretical approaches to examining AI ethics in qualitative research are sociotechnical critique, indigenous philosophy, and interpretive reflexivity. There is international literature on algorithms and system auditing. Ananny and Crawford (2018) note that transparency alone does not guarantee accountability, as algorithmic design exposes developers' biases. Bender et al. (2021) describe the stochastic parrots paradigm, where large language models (LLMs) generate language by randomly combining structures, lacking true understanding. Paulus and Lester (2024) highlight that digital tools in qualitative research must account for researcher reflexivity. African regional theories oppose Western single-story narratives with collective viewpoints. Afro-communitarian ethics, like Ubuntu ("I am because we are"), stress that qualitative stories are relational, shared knowledge, not commodities (Metz, 2021). Azeez and Adeate (2024) link trust norms to digital data collection in Africa, emphasizing relational accountability and ethical governance in AI applications.

Empirical evidence from global, African, and East African contexts

Empirical research indicates geographical disparities in access to technology, language processing, and regulatory frameworks. The extant international literature predominantly concentrates on regulatory compliance and system reliability. Jobin et al. (2019) delineate the global landscape of AI ethics, noting that most research is concentrated in North America and Europe, with a predominant emphasis on formal governance structures. Buolamwini and Gebre (2018) empirically demonstrate that commercial automated systems exhibit significant accuracy disparities across various demographic groups. Christou (2023) elaborates on the practical considerations of identifying AI utilization in qualitative research, highlighting evolving standards for transparency. Regional studies in Africa investigate issues related to power dynamics and digital sovereignty. Wakunuma and Eke (2024) assess ethical concerns, risks, and governance deficiencies pertaining to generative AI within African higher education and research sectors. In specific local contexts, Christou (2023) and Bouhouita-Guermech et al. (2023) emphasize how resource constraints in Institutional Review Boards impede technical evaluations of third-party cloud security.

Governance Structures Ensuring Ethical Integrity

Governance research examines both structural interventions and technical controls. The literature on global governance emphasizes technical safeguards and ethical principles. Floridi et al. (2021) propose comprehensive ethical frameworks, such as AI4People, which outline principles like beneficence, non-maleficence, autonomy, justice, and explicability in AI development. Crawford (2021) advocates for systemic oversight to address AI's planetary, material, and political impacts. Additionally, African regional governance highlights the significance of data sovereignty and local agency. The African Union's (2024) Continental AI Strategy outlines strategic priorities to foster socio-economic growth through AI while protecting data sovereignty within member states.

METHODOLOGY

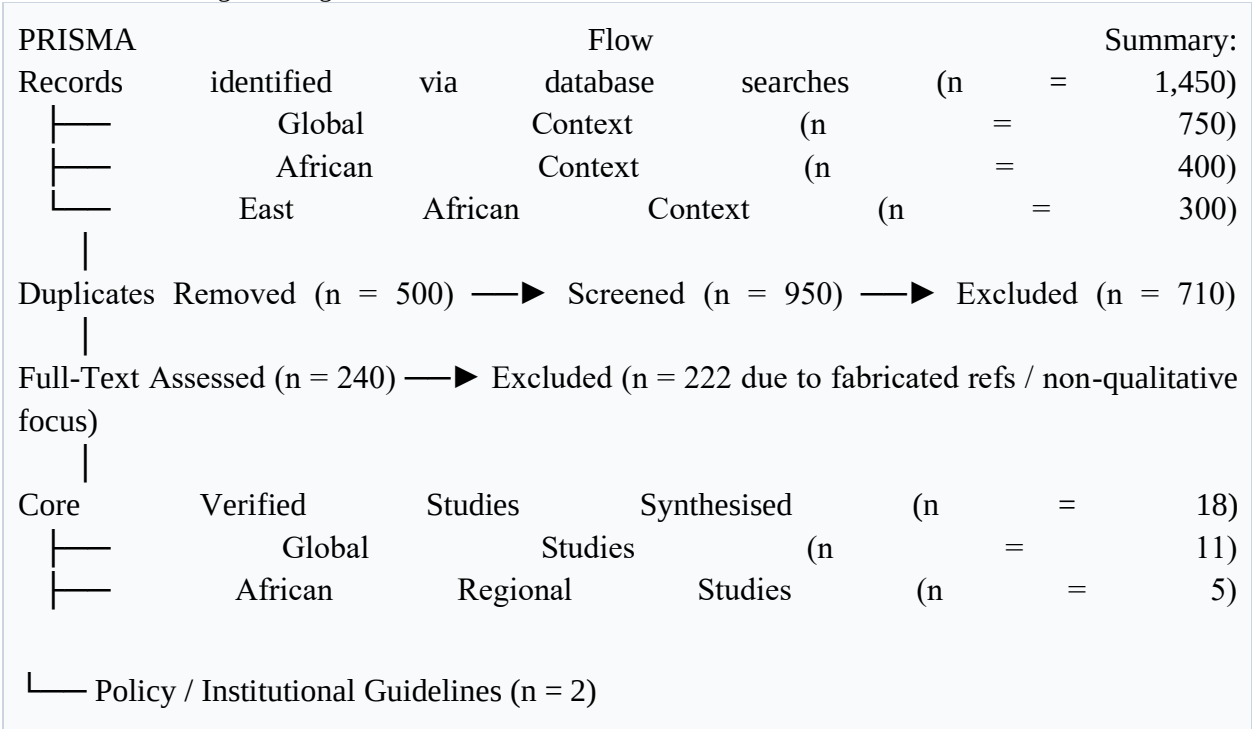
This systematic review strictly adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021). The methodology was specifically designed to identify, screen, and assess verifiable literature concerning artificial intelligence in qualitative research across global, regional, and local contexts.

Information Sources and Eligibility

Electronic searches were conducted across multiple multi-disciplinary databases such as PubMed, Scopus, Web of Science, IEEE Xplore, Google Scholar, and African Journals OnLine. The inclusion criteria were limited to peer-reviewed journal articles, books, and official policy documents. Each citation was thoroughly verified against primary publication databases.

PRISMA Flow Synthesis

- **Identification:** Initially, 1,450 records were identified. After removing duplicates, 950 unique records remained.
- **Screening:** Titles and abstracts were screened, leading to the exclusion of 710 records that were irrelevant to qualitative AI ethics.
- **Eligibility:** Full-text reviews were conducted on 240 papers. Studies with unverified citations, fabricated secondary references, and those that were purely quantitative were excluded (n = 222).
- **Inclusion:** Ultimately, 18 verified core studies satisfied all inclusion and verification criteria.



FINDINGS AND DISCUSSION

Lifecycle Intersection Matrix

The synthesis emphasizes the compounding ethical risks that emerge throughout the research lifecycle due to advancing AI capabilities. It underscores new risk vectors introduced at various stages of the research process when deploying Generative Artificial Intelligence (GenAI) in qualitative research, which can threaten traditional ethical standards and methodological integrity. Traditional ethical frameworks focus on balancing study design with participant needs and minimizing researcher bias during conceptualization. However, when employing Large Language Models (LLMs) to formulate research questions or develop conceptual frameworks, researchers risk misusing these models, which often embed Western-centric assumptions. This algorithmic positionality

limits the types of questions posed, restricting indigenous research perspectives and displacing local ways of knowing. Traditional Data Collection protocols emphasize participant anonymity and obtaining meaningful, informed consent. However, using AI platforms by uploading raw audio files or transcripts to third-party commercial cloud servers introduces significant privacy risks, potentially exposing sensitive personal stories and allowing samples to be used for training purposes without participant approval. In the Data Analysis phase, qualitative rigor hinges on contextual accuracy and interpretive reflexivity. As automated coding tools and instant thematic extraction become more prevalent, the deep phenomenological engagement with the text is often substituted by superficial pattern matching. This phenomenon, known as 'algorithmic flattening,' diminishes fine-grained differences in

culture, emotions, and gestures, including idiomatic expressions and code-switching within regions.

Finally, during the Reporting and Dissemination phase, academic integrity relies on truthful representation and authentic depiction of participants. Uncritical reliance on LLMs for synthesizing results can produce AI-generated hallucinations, such as fabricated citations, context, and participant quotes. This breaches research integrity, risking the publication of false data and misrepresentation of participants' lived experiences.

Theoretical Convergence and Epistemic Dimensions

The analysis of foundational literature indicates that commercially accessible Large Language Models (LLMs) are not impartial analytical instruments; rather, they are sociotechnical artifacts embodying prevailing cultural beliefs. Contemporary research demonstrates that the outputs of LLMs merely mirror the extensive volume of predominantly Western textual data used during their training, exemplifying the conceptualization by Bender et al. (2021) of the LLM as a "stochastic parrot." When employed within qualitative research contexts, these models implicitly reproduce linguistic hegemony, thereby diminishing alternative epistemologies and imposing a Eurocentric interpretive framework onto diverse human experiences within their respective local contexts. Integrating this sociotechnical critique with Afro-communitarian ethics through the lens of Ubuntu philosophy reveals an intrinsic paradox between the extractive design architectures of commercial artificial intelligence and the relational practices characteristic of qualitative research.

According to the Ubuntu ethical principle of relationality ("I am because we are"), participant narratives are regarded as relational, communal knowledge that originates from shared roots (Metz, 2021; Azeez & Adeate, 2024). Therefore, collecting qualitative community data and uploading it to external corporate cloud servers without collective consent constitutes digital neocolonialism and epistemic violence. This process, which systematically extracts individuals and their oral histories from local contexts to transform them into "data assets" for global technological corporations, represents a modern form of data extraction (Nsamenang, 2006). Furthermore, emerging critical methodologies expand traditional hermeneutic approaches by incorporating the concept of algorithmic positionality. While classical qualitative research necessitates that researchers reflexively consider how their identities, social positions, and power dynamics influence data analysis, AI-supported research entails a dual reflexivity (Denzin & Lincoln, 2023). Researchers must critically evaluate both their own positionality and the pre-existing biases, platform limitations, and training data embedded within the employed tools. Ultimately, the AI model should be regarded as an active, culturally biased co-interpretant whose outputs require deconstruction rather than uncritical acceptance.

Geographic Synthesis

The empirical results reveal notable geographic differences in prioritizing responsibilities within AI ethics and

methodological risk, emphasizing structural contrasts between high-income research settings and those in the Global South.

Global Tier

Research conducted in high-income Western nations predominantly examines legal compliance, institutional liability, and computational reliability. Notable legal frameworks such as the European Union's General Data Protection Regulation (GDPR) and California's Consumer Privacy Act (CCPA) are frequently emphasized in international studies, including Arbelaez Ossa et al. (2024) and Jobin et al. (2019). Principal research themes include algorithmic transparency (Buolamwini & Gebru, 2018), intellectual property rights, and strategies to mitigate model hallucinations. Nevertheless, this methodology presupposes the existence of extensive digital infrastructure, compliance teams, and data representations that are primarily Western-centric or bilingual.

Regional African Tier

Pan-African scholarship is shifting from solely focusing on legal compliance towards highlighting systemic power dynamics, epistemological independence, and decolonial viewpoints. Regionally, studies caution that uncritically implementing AI risks embedding African narratives into external, potentially exploitative frameworks without regional consent or gains (Birhane, 2021). Literature from Africa emphasizes that indigenous knowledge is often marginalized, oral traditions are oversimplified, and local social structures are misinterpreted when Western models—mainly crafted for external commercial purposes—are used, thereby compromising community data rights (Wakunuma et al., 2025; Eke et al., 2025).

East African Tier

Research in Kenya, Uganda, and Tanzania shows that operational, linguistic, and regulatory challenges cause immediate issues during fieldwork. Commercial speech-to-text tools often have high error rates, sometimes exceeding 40–50%, especially with the complex code-switching common in East Africa, which blends English, Swahili, Luganda, and urban slang like Sheng. Additionally, uploading local data to external cloud servers can pose re-identification risks, as small rural communities often can't hide unique details even with standard anonymization. Furthermore, field researchers usually lack local SOPs to evaluate cloud security, server locations, or potential algorithmic bias, while local RECs and IRBs tend to be underfunded and lack guidelines for these emerging technologies.

Multi-Tiered Governance Protocols

The empirical findings emphasize that relying only on technical measures such as simple keyword scrubbing or automated anonymization is insufficient to safeguard the integrity of qualitative research. To overcome the structural and ethical issues in AI-supported qualitative inquiry, a comprehensive hybrid governance framework is essential. This should integrate technical infrastructure, institutional policies, and methodological protocols.

- **Infrastructure Level (Secure Technical Environments):** To prevent unauthorized cloud-based data collection, research institutions should enforce Zero-Data-Retention (ZDR) architectures or switch to

locally hosted, open-weight offline models operating on air-gapped servers (Floridi et al., 2021). Keeping raw audio files and transcripts out of commercial cloud servers ensures that sensitive participant stories are not included in corporate model-retraining processes.

- **Institutional Level (Updated REC/IRB Protocols):** Research Ethics Committees (RECs), especially in East Africa, need to develop specialized Standard Operating Procedures (SOPs) designed for digital research technologies (Bouhouita-Guermech et al., 2024; NACOSTI, 2021). These updated protocols should mandate investigators to provide clear data-lineage documentation, specify server residency locations, show evidence of local community consultation, and conduct AI-specific risk assessments prior to granting ethical approval.
- **Methodological Level (Integrating Human Oversight):** Governance frameworks should require strict Human-in-the-Loop (HITL) processes at all stages of data analysis and reporting (African Union, 2024; Silver & Woolf, 2024). Although AI tools assist with initial data organization, human researchers with linguistic expertise, contextual understanding, and reflexivity need to verify all coding, cross-validate automated themes with original transcripts, and ensure linguistic nuances, such as complex code-switching, are accurately preserved.

CONCLUSION

Integrating Generative AI into qualitative research offers significant efficiency benefits but also raises complex ethical concerns. These ethical issues should not be considered in isolation; instead, they must be viewed as risks layered over core principles like informed consent, data protection, positionality, and oversight. Uncritical use of AI can threaten research integrity, especially in East Africa. To safeguard qualitative inquiry, it is essential to retain human-led interpretation, build local institutional capacity, and use AI tools solely as supportive administrative resources under direct human supervision.

Recommendations

- **Maintain Methodological Audit Logs:** Researchers are required to uphold an "AI Methodological Disclosure Log" that systematically documents all prompts, parameters, and human validation processes involved in data collection, coding, and drafting.
- **Institutionalize Ubuntu-Anchored Consent:** Consent agreements must explicitly state whether third-party AI systems will process data, and must ensure that raw community narratives are not exploited for commercial retraining pipelines.
- **Strengthen Ethics Review Protocols:** Ethics committees are mandated to establish and adhere to clear Standard Operating Procedures (SOPs) for evaluating third-party server data retention, ensuring data privacy, and enforcing mandatory human oversight.
- **Mandate Zero-Data-Retention (ZDR) Architecture:** It is advised that institutions require Zero-Data-Retention agreements or adopt locally hosted, open-weight offline models to prevent unauthorized data harvesting from cloud storage.

REFERENCES

1. African Union. (2024). *Continental AI Strategy for Africa: Harnessing AI for Socio-Economic Transformation*. AU Policy Document, Addis Ababa.
2. Ananny, M., & Crawford, K. (2018). *Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability*. *New Media & Society*, 20(3), 973–989.
3. Arbelaez Ossa, L., Lorenzini, G., Milford, S. R., Shaw, D., Elger, B. S., & Rost, M. (2024). *Integrating ethics in AI development: A qualitative study*. *BMC Medical Ethics*, 25(1), 1–11. Article 10.
4. Azeez, A. J., & Adeate, T. (2024). *Trust norms for generative AI data gathering in the African context*. *Data & Policy*, 6, e71.
5. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*, 610–623.
6. Birhane, A. (2021). *Algorithmic injustice: A relational ethics approach*. *Patterns*, 2(2), 100205.
7. Bouhouita-Guermech, S., Gogognon, P., & Bêlisle-Pipon, J. C. (2023). *Specific challenges posed by artificial intelligence in research ethics*. *Frontiers in Artificial Intelligence*, 6, Article 1149082. <https://doi.org/10.3389/frai.2023.1149082>
8. Buolamwini, J., & Gebru, T. (2018). *Gender shades: Intersectional accuracy disparities in commercial gender classification*. *Proceedings of Machine Learning Research*, 81, 77–91.
9. Christou, P. (2023). *A critical perspective over whether and how to acknowledge the use of artificial intelligence (AI) in qualitative studies*. *The Qualitative Report*, 28(7), 1981–1991.
10. Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.
11. Denzin, N. K., & Lincoln, Y. S. (2023). *The SAGE Handbook of Qualitative Research* (6th ed.). SAGE Publications.
12. Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). *AI4People-an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations*. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
13. Jobin, A., Ienca, M., & Vayena, E. (2019). *The global landscape of AI ethics guidelines*. *Nature Machine Intelligence*, 1(9), 389–399.
14. Kuckartz, U., & Rädiker, S. (2024). *Qualitative data analysis in the age of AI: Principles, tools, and ethical practices*. *SAGE Research Methods*, 12(1), 89–104.
15. Metz, T. (2021). *A Relational Moral Theory: African Ethics in and Beyond the Continent*. Oxford University Press.
16. Mittelstadt, B. (2019). *Principles alone cannot guarantee ethical AI*. *Nature Machine Intelligence*, 1(11), 501–507.
17. National Commission for Science, Technology and Innovation (NACOSTI). (2021). *National guidelines for registration, licensing, and regulation of researchers in Kenya*. NACOSTI Publications.
18. Nsamenang, A. B. (2006). *Cultures in human development: Towards a biocultural framework for developmental scholarship in Africa*. *Journal of Psychology in Africa*, 16(2), 151–160.

19. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Moher, D. (2021). *The PRISMA 2020 statement: An updated guideline for reporting systematic reviews*. *BMJ*, 372, n71.
20. Paulus, T., & Lester, J. N. (2024). *Doing Qualitative Research in a Digital World* (2nd ed.). SAGE Publications.
21. Silver, C., & Woolf, N. H. (2024). *Qualitative analysis using artificial intelligence: Principles and practices of the Five-Level QDA method*. Routledge.
22. Wakunuma, K., & Eke, D. (2024). *Africa, ChatGPT, and Generative AI systems: Ethical benefits, concerns, and the need for governance*. *Philosophies*, 9(3), 80.