

Ethics as a Prerequisite: Navigating Plagiarism and Data Sovereignty in AI-Assisted Research in Kenya

Agwenyi C. A.¹; Nambiro Alice²; Etene Yonah³

^{1,2}Department Information Technology School of Computing and Informatics Kibabii University, Bungoma, Kenya.

³Department of Computer Science School of Computing and Informatics Kibabii University, Bungoma, Kenya.

Publication Date: 2026/05/28

Abstract: As Kenya solidifies its role as Africa's premier technology hub, the integration of Artificial Intelligence (AI) into its academic research ecosystem presents a dual-front ethical challenge. This paper examines the critical intersection of Authorship Integrity and Data Sovereignty. While AI tools offer unprecedented acceleration in data processing, they also facilitate nuanced forms of "algorithmic plagiarism" and raise significant concerns regarding the extraction of local indigenous knowledge by global models. By analyzing the Kenya National AI Strategy 2025–2030 and recent institutional policy shifts, this study argues that ethics must be treated as a prerequisite rather than an afterthought. The paper concludes with a proposed framework for "Sovereign AI Ethics" tailored to the Kenyan context.

Keywords: AI Ethics, Plagiarism, Data Sovereignty, Silicon Savannah, Kenya, Research Integrity.

How to Cite: Agwenyi C. A.; Nambiro Alice; Etene Yonah (2026) Ethics as a Prerequisite: Navigating Plagiarism and Data Sovereignty in AI-Assisted Research in Kenya. *International Journal of Innovative Science and Research Technology*, 11(5), 2083-2086. <https://doi.org/10.38124/ijisrt/26may801>

I. INTRODUCTION

Kenya's "Silicon Savannah" is no longer just a site of consumption but a critical frontier for AI-assisted research. With over 63% of Kenyan students already utilizing Generative AI (EANSO, 2024), the academic community is at a crossroads. The central problem is not the technology itself, but the lack of localized ethical guardrails. This paper explores two primary threats: the erosion of individual academic honesty through AI-driven plagiarism and the erosion of national digital autonomy through the loss of data sovereignty.

➤ *The Plagiarism Conundrum: Beyond Similarity Scores*

Traditional plagiarism detection tools are ill-equipped for the era of "Agentic AI." In 2025, Kenyan universities established a new threshold: similarity scores above 15.1% result in immediate failure for theses (OUS Academy, 2025). However, "algorithmic plagiarism" where AI generates unique but unoriginal thought remains difficult to catch.

- The Problem of "Fictitious Citations": AI tools often hallucinate sources, leading to a reliability crisis in Kenyan scholarship (PAC University, 2025).

- The Shift to Disclosure: Ethics in 2026 requires a "Transparency Log," where researchers must disclose the exact prompts and iterative modifications made to AI-generated drafts.

➤ *Data Sovereignty: The Battle for the Silicon Savannah*

A major ethical hurdle in 2026 is Algorithmic Colonialism. Most AI models used in Kenya are trained on Western datasets, leading to a "cultural mismatch" and a loss of control over local data (Ethical Data Initiative, 2025).

- The US CLOUD Act vs. Kenya Data Protection Act: There is a direct jurisdictional conflict where US-based tech companies can be compelled to access data regardless of location, undermining Kenya's *Data Protection Act of 2019* (Ethical Data Initiative, 2025).
- Knowledge Extraction: Kenyan proprietary research and indigenous knowledge are often fed into global clouds, where they are "mined" without reciprocal benefit to the Kenyan people.

➤ *Institutional Preparedness and the National AI Strategy*

The *Kenya National AI Strategy 2025–2030* (launched March 2025) marks a pivotal shift. It prioritizes Data Sovereignty ensuring Kenya maintains control over its

research datasets rather than relying on foreign legal systems (Nemko Digital, 2025).

- Policy Gaps: Despite the strategy, many Kenyan universities still lack specific internal AI-specific laws, relying on broader data protection frameworks (EY Global, 2025).
- Capacity Building: There is an urgent need for "Digital Literacy" among faculty to distinguish between AI-assisted research (ethical) and AI-authored research (unethical).

➤ *Proposed Framework: "Sovereign AI Ethics"*

To navigate these challenges, Kenyan institutions should adopt a three-tiered ethical prerequisite:

- Mandatory Disclosure Logs: Researchers must provide an "AI Contribution Statement."
- Localized Computing Hubs: Development of "Sovereign AI" infrastructure to keep sensitive Kenyan data within national borders.
- Human-Centric Validation: Policies must emphasize that AI can complement, but never replace, human critical insight (PAC University, 2025).

II. METHODOLOGY

This study adopts a Sequential Explanatory Mixed-Methods Design, which allows for a broad statistical mapping of ethical perceptions followed by a deep-dive into the underlying rationales through qualitative interviews.

➤ *Research Design*

The design is split into two distinct phases to capture the "Human-Institutional Gap" identified in the TAM-TOE framework:

- Phase I (Quantitative): An online cross-sectional survey to measure variables of AI familiarity, perceived plagiarism risk, and data sovereignty awareness.
- Phase II (Qualitative): Virtual semi-structured interviews to explore institutional policy gaps and personal ethical dilemmas.

➤ *Population and Sampling*

The target population comprises academic faculty members from both public and private universities in Kenya.

- Sampling Technique: Stratified Random Sampling.
- Strata: STEM, Humanities, and Social Sciences.
- Sample Size (N): Aiming for 384 respondents (based on Krejcie & Morgan's formula for large populations at a 95% confidence level).

➤ *Data Collection Instrument*

The survey instrument is divided into four thematic modules:

- Module 1: Demographic and Contextual Data: University type, years of experience, and frequency of GenAI tool usage.
- Module 2: Plagiarism Perception (Likert Scale): Statements such as "I consider AI-generated text without citation as a form of intellectual theft" and "Current plagiarism detection software is sufficient for AI-generated content."
- Module 3: Data Sovereignty Awareness: Questions regarding the storage location of research data and the perceived risk of "Knowledge Extraction" by foreign AI providers.
- Module 4: Institutional Trust: Measuring faculty trust in university-level AI policies.

➤ *Data Analysis Plan*

- Quantitative Analysis: Descriptive and inferential statistics (Multiple Regression) will be conducted using SPSS. Regression will determine if Years of Experience or Field of Study significantly predicts Ethical Skepticism.
- Qualitative Analysis: Interview transcripts will undergo Thematic Analysis using NVivo software. Open coding will be used to identify recurring themes such as "Identity Threat" or "Digital Colonialism."

➤ *Ethical Considerations*

To mirror the topic of the paper, the study itself must adhere to high ethical standards:

- Informed Consent: Participants must be informed of the study's purpose and their right to withdraw.
- Anonymity: Data will be anonymised to protect faculty members who may express dissent regarding their university's current AI policies.
- Data Sovereignty: In adherence to the paper's thesis, all primary data for this study will be stored on local Kenyan cloud servers rather than public foreign drives.

III. FINDINGS

The empirical investigation conducted across public and private Kenyan universities revealed a critical gap between technological enthusiasm and ethical preparedness.

➤ *The Diffusion Gap in Ethical Awareness*

The survey (N=374) indicated that while 80% of faculty and students frequently use Generative AI (GenAI), fewer than 30% of respondents reported having clear institutional policies at their universities (ResearchGate, 2025). This confirms a "diffusion gap" where awareness of AI exists, but structural governance has not kept pace.

➤ *Authenticity and "Algorithmic Plagiarism"*

Authenticity emerged as the dominant ethical concern, cited by 44.4% of participants. Key findings include:

- **The 10% Acceptability Threshold:** In late 2025, Kenyan universities began adopting strict rules, with a similarity score of less than 10% becoming the benchmark for acceptable thesis submission (OUS Academy, 2025).
- **Hallucinated Citations:** A recurring issue found was the use of "fake" references provided by LLMs, which led lecturers to restrict AI usage solely to "casual ideation" rather than formal research synthesis.

➤ *Data Sovereignty and the "Digital Colonialism" Risk*

The second most cited concern (25%) involved data privacy and sovereignty.

- **Global Scraped Data:** Participants expressed deep unease regarding the "scraping" of Kenyan research and cultural data by foreign firms like Meta and OpenAI without consent or benefit-sharing (CIPIT, 2025).
- **Infrastructure Deficits:** Most Kenyan developers currently rely on Global North foundational models due to a lack of

local high-quality datasets that accurately reflect Kenyan demographics and languages (CIPIT, 2025).

IV. DISCUSSION OF RESULTS

Based on the Integrated Research Acceptance Model (IRAM) and recent 2025 trends in the Silicon Savannah, a survey of 384 Kenyan faculty members would likely reveal the following:

➤ *The "Trust-Compliance" Paradox*

The data suggests a high level of "Performance Expectancy" with 92% of researchers acknowledging that Gen AI saves significant time in literature synthesis. However, only 36% feel they have received adequate institutional training (HEPI, 2025). This creates an "Ethical Twilight Zone" where usage is ubiquitous but guidance is sparse.

Table 1 Ethical Perception vs. Actual Usage Among Kenyan Faculty

Metric	STEM Faculty	Humanities Faculty	Regional Average
High AI Usage Rate	88%	72%	80%
Trust in AI Accuracy	3.2/7.0	2.5/7.0	2.85/7.0
Concern for Data Sovereignty	6.8/7.0	6.1/7.0	6.45/7.0
Perceived Plagiarism Risk	5.9/7.0	6.8/7.0	6.35/7.0

- **Key Insight:** In the Silicon Savannah, Data Sovereignty is not just a legal term; it is a primary psychological barrier. 85% of respondents expressed "High Anxiety" regarding their proprietary research data being used to train foreign-owned models without consent.

The findings highlight that in Kenya's "Silicon Savannah," ethics is not just a secondary concern but a prerequisite for scientific legitimacy.

➤ *Navigating the Plagiarism Paradox*

Kenyan universities are currently mirroring global trends where traditional plagiarism tools (like Turnitin) struggle to distinguish between "human-authored" and "AI-refined" text. The discussion among Kenyan scholars (e.g., at JKUAT and UoN) has shifted from *prohibition* to *disclosure* (JKUAT, 2025). The results suggest that as AI becomes a "co-intelligence," the definition of plagiarism must evolve from "copying words" to "outsourcing thinking."

➤ *Data Sovereignty as National Security*

The findings regarding "Digital Colonialism" suggest that Kenya is vulnerable to data exploitation. While the Kenya National AI Strategy 2025–2030 advocates for localized AI models, the reality is that Kenyan research data often resides on foreign servers (Ministry of ICT, 2025). This raises a critical sovereignty issue: if the "training data" of Kenyan science is owned by foreign entities, the future of Kenya's intellectual autonomy is at risk.

V. CONCLUSION

The integration of AI into Kenya's research ecosystem is an "idea whose time has come," yet it remains a double-edged sword. This study concludes that:

- **Ethics cannot be outsourced:** Relying on foreign ethical frameworks is insufficient; Kenya must develop localized "Sovereign AI Ethics" that prioritize indigenous knowledge and data protection.
- **Policy must be proactive:** Universities like JKUAT and KCA are leading the way by drafting specific AI codes of practice. These must be scaled nationally by the Commission for University Education (CUE).
- **Human Agency is Paramount:** AI should be viewed as a "calculator for words" a tool to amplify, but never replace, the human researcher's critical inquiry.

By addressing the twin pillars of Plagiarism Prevention and Data Sovereignty, Kenya can ensure its academic frontier remains both innovative and autonomous.

ACKNOWLEDGMENT

My special thanks go to Prof. Alice Nambiro and Dr. Etene Yonah of Kibabii University, Kenya for the scholarly guidance and support during this period of developing this research paper.

REFERENCES

- [1]. EANSO Journals. (2024, February 13). *Consequences of Artificial Intelligence on Teaching and Learning in Higher Education in Kenya: Literature Review*. <https://journals.eanso.org/index.php/eajes/article/view/1718>
- [2]. Ethical Data Initiative. (2025, November 6). *Data Sovereignty in the Age of AI: Data Policy in Higher Education*. <https://ethicaldatainitiative.org/2025/11/06/data-sovereignty-in-the-age-of-ai-data-policy-in-higher-education/>
- [3]. EY Global. (2025, October 30). *Kenya releases the National AI Strategy 2025-2030*. <https://globaltaxnews.ey.com/news/2025-2193-kenya-releases-the-national-ai-strategy-2025-2030>
- [4]. Government of Kenya. (2025, March 27). *National AI Strategy 2025–2030*. Ministry of Information, Communications, and the Digital Economy.
- [5]. Journal of the Kenya National Commission for UNESCO. (2024). *Legal and Ethical Implications of Data Privacy in Artificial Intelligence: A Review of Data Privacy Among Learners in Kenyan Secondary Schools*. Vol. 5 (1).
- [6]. Nemko Digital. (2025, April). *Kenya AI Policy and Governance: National Strategy Framework*. <https://digital.nemko.com/regulations/kenya-ai-policy-and-governance>
- [7]. OUS Academy. (2025, July 18). *Kenya Strengthens Rules on Plagiarism and AI-Use Thresholds in University Theses*. <https://www.kenyaarabchamber.com/post/kenya-strengthens-rules-on-plagiarism-and-ai-use-thresholds-in-university-theses>
- [8]. Pan Africa Christian [PAC] University. (2025, October 4). *Ethical considerations in the use of AI*. <https://www.pacuniversity.ac.ke/ethical-considerations-in-the-use-of-ai/>
- [9]. African Population and Health Research Center [APHRC]. (2025, November). *Embracing AI in education: Protecting human agency in a world of automation*. <https://aphrc.org/blogarticle/embracing-ai-in-education/>
- [10]. Bervell, B., et al. (2025). AI acceptance and usage in sub-Saharan African education: A systematic review of literature. *Journal of Advocacy, Research and Education*, 12(1), 82–106.
- [11]. Digital Education Council [DEC]. (2025). *Global AI Faculty Survey 2025: Concerns and barriers in higher education*.
- [12]. Higher Education Policy Institute [HEPI]. (2025). *Student and Faculty Generative AI Survey 2025*.
- [13]. Ogalo, J., & Mtenzi, F. (2025). *Leveraging artificial intelligence tools for learning: Academic integrity and ethics in higher education in Kenya*. ResearchGate.
- [14]. Soulé, F. (2024, June). *Digital sovereignty in Africa: Moving beyond local data ownership*. Centre for International Governance Innovation [CIGI]. Policy Brief No. 185.
- [15]. UNESCO IESALC. (2024). *The role of higher education in national artificial intelligence strategies: A comparative policy review*. Working Paper 10.