JURNAL ADABIYA
E-ISSN: 2549-1776 P-ISSN: 2549-3124
Vol. 28, No. 2, August 2026, pp. 106~124
DOI: 10.22373/adabiya.v28i2.36354

Developing Locally-Grounded Ethical Artificial Intelligence Frameworks for South African Academic Libraries

Dickson Mdhlalose
[1] Department of Information and Communication Technology, National Electronic Media Institute of South Africa, Johannesburg, South Africa.
Email Correspondence: dsskosana@gmail.com[1]

Article Info

Article history:

Received 07 / 18, 2026
Revised 08 / 18, 2026
Accepted 08 / 22, 2026

Keywords:
Artificial Intelligence; South African Academic Libraries; Epistemic Justice; Algorithmic bias; Artificial intelligence governance.

ABSTRACT

This paper explores the real-world ethical challenges and governance gaps that South African academic libraries face as they adopt Artificial Intelligence (AI). Through conversations with library directors, focus groups with frontline librarians, and analysis of information technology (IT) policy documents from six universities, the study uncovers four main problem areas: absence of local, relevant AI governance; algorithms often overlook African and Global South scholarship; staff don't always have the training or direction they need; and librarians themselves are often sidelined, expected to use new technologies rather than shape how they are used. To address these issues, the study offers a practical Ethical AI Integration Framework (EAIF), built on five key pillars: epistemic justice, algorithmic transparency, data sovereignty, multilingual equity, and participatory stewardship. Ultimately, the research argues that South African academic libraries need to develop their own ethical approaches rooted in local values, languages, and a commitment to fair access to knowledge rather than relying on imported frameworks that may not fit their contexts. This study provides, for the first time an evidence-based, co-designed governance model for AI adoption in African libraries by shifting the narrative beyond mere description of adoption towards providing the tools for responding to the epistemic exclusion, multilingual silencing and de-skilling in an interdependent yet distinct, complex cluster of ethical, infrastructural and agentic issues neglected by existing work which has tended to foreground access and equipment over ethical implementation.

Kata Kunci:
Kecerdasan buatan; Perpustakaan akademik Afrika Selatan; Keadilan epistemik; Bias algoritmik; Tata kelola kecerdasan buatan.

ABSTRAK

Makalah ini mengeksplorasi tantangan etis dunia nyata dan kesenjangan tata kelola yang dihadapi oleh perpustakaan akademik di Afrika Selatan dalam mengadopsi Kecerdasan Buatan (Artificial Intelligence/AI). Melalui wawancara dengan para direktur perpustakaan, diskusi kelompok terpumpun (focus groups) bersama pustakawan garda depan, serta analisis dokumen kebijakan teknologi informasi (TI) dari enam universitas, studi ini mengungkap empat bidang masalah utama: ketiadaan tata kelola AI lokal yang relevan; algoritma yang kerap mengabaikan karya ilmiah dari Afrika dan kawasan Selatan Global (Global South); staf yang tidak selalu memiliki pelatihan atau arahan yang memadai; serta pustakawan itu sendiri yang sering dikesampingkan dan hanya diharapkan untuk menggunakan teknologi baru alih-alih ikut membentuk cara pemanfaatannya. Untuk mengatasi permasalahan tersebut, studi ini menawarkan Ethical AI Integration Framework (EAIF) praktis yang dibangun di atas lima pilar utama: keadilan epistemik, transparansi algoritmik, kedaulatan data, kesetaraan multibahasa, dan pengelolaan partisipatif (participatory stewardship). Pada akhirnya, penelitian ini menegaskan bahwa perpustakaan akademik di Afrika Selatan perlu mengembangkan pendekatan etis mereka sendiri yang berakar pada nilai-nilai lokal, bahasa daerah, serta komitmen terhadap akses pengetahuan yang adil, alih-alih bergantung pada kerangka kerja impor yang mungkin tidak sesuai dengan konteks mereka. Studi ini menyajikan, untuk pertama kalinya, sebuah model tata kelola berbasis bukti yang dirancang bersama (co-designed) untuk adopsi AI di perpustakaan-perpustakaan Afrika. Model ini menggeser narasi dari sekadar deskripsi adopsi menuju penyediaan instrumen untuk merespons eksklusi epistemik, pembungkaman multibahasa, dan

penurunan keterampilan (de-skilling) dalam sebuah gugus isu etis, infrastruktural, dan agensial yang kompleks, saling bergantung namun berbeda, yang selama ini terabaikan oleh kajian terdahulu yang cenderung lebih mengedepankan aspek akses dan peralatan ketimbang implementasi yang etis.

This is an open access article under the CC BY-SA license.

1. INTRODUCTION

The rise of artificial intelligence (AI) in academic library services is changing the profession. Libraries worldwide now use AI-powered chatbots, machine-learning tools for cataloguing, and data analytics for collection decisions. These technologies help libraries do more and reach people who might otherwise be left out, but they raise complicated ethical questions that require thoughtful, locally informed oversight.

In South Africa, these challenges are especially urgent. The country's 26 public universities serve about 1.1 million students, many of them the first in their families to attend university. South Africa's constitution guarantees access to information, equality, and dignity, values that should be central to how AI is used in libraries. The country aims to build a knowledge-driven economy yet remains one of the most unequal societies in the world, a mix of high hopes and deep inequalities that shape the ethical issues its libraries face in adopting AI.

Research on AI in libraries, though growing, has been dominated by studies from the United States, Europe, and Australia (Gorichanaz, 2023; Raber & Budd, 2022; Tammaro et al., 2021), while African contexts, particularly South Africa, remain sparse and largely descriptive, focused on access and infrastructure rather than ethical or professional dimensions (Okonkwo & Ezema, 2022; Molepo & Fourie, 2020; Chukwuemeka & Ugwuanyi, 2023; Dlamini & Ngoepe, 2021). This imbalance is not merely academic; it reflects and reproduces the epistemic hierarchies that AI systems risk entrenching (Couldry & Mejias, 2021; Mohamed et al., 2020). Meanwhile, South African libraries face a flood of commercial AI tools, discovery systems, chatbots, recommendation engines, developed and tested largely in Europe and North America, often assuming reliable connectivity and shared language and academic traditions that do not reflect South African realities. Adopting such tools without clear local policy risks worsening inequality, marginalising African scholarship, and compromising the privacy of already vulnerable users.

Molepo and Fourie (2020) found that most South African academic libraries lacked formal AI adoption strategies, with fewer than 30% of staff trained in AI literacy. More recent data from the Library and Information Association of South Africa (LIASA, 2023) show that while 67% of directors believe AI will transform library services within five years, only 14% report having an institutional AI governance policy, a strategy-practice gap that represents a critical vulnerability.

A review of the literature reveals four gaps this study addresses: (1) the absence of localised ethical governance frameworks accounting for South Africa's 11 official languages, dual economy, and constitutional commitments (Andrade & Alves, 2022; Sewchurran & Gregoriou, 2022); (2) despite growing attention to algorithmic bias generally (Obermeyer et al., 2022; Zou & Schiebinger, 2021), little research examines how library discovery and recommendation systems disadvantage Global South scholarship specifically (Brock, 2020; Helgesson & Eriksson, 2021); (3) the well-documented digital divide in South African higher education (Brown & Czerniewicz, 2020; Czerniewicz et al., 2021) has not been systematically connected to AI adoption in libraries; and (4) the profession has not adequately theorised the librarian as an active ethical agent and co-creator of AI governance, rather than a passive recipient of externally determined change (Gorichanaz, 2023; Pawley, 2022).

AI's arrival in library services is the latest phase in a longer history of automation, from integrated library systems in the 1970s to discovery layers and linked data in the 2000s, and now to user-facing applications such as chatbots, automated metadata enrichment, and predictive collection-development analytics (Sadeh, 2022; Tammaro et al., 2021; Choi et al., 2022; Online Computer Library Center [OCLC], 2023). As with earlier waves, these tools bring efficiency gains alongside unintended consequences, including opaque decision-making and the risk that automated systems reproduce biases present in historical library data (Bawden & Robinson, 2022; Brock, 2020; Olteanu et al., 2020). Breed (2023) argues the profession has been insufficiently critical in adopting commercial AI, often prioritising vendor efficiency narratives over community-centred values, a critique that resonates with wider Global South debates on the absence of local agency in technology design (Couldry & Mejias, 2021; Ndovu & Mutula, 2021).

Global ethical frameworks, including the United Nations Educational, Scientific and Cultural Organization's (UNESCO) Recommendation on the Ethics of AI (2021), the International Federation of Library Associations and Institutions' (IFLA) statements on algorithmic transparency (2022), and the American Library Association's principles (2021), provide an important normative foundation but have been critiqued for universalising assumptions that ignore Global South structural conditions (Mohamed et al., 2020; Munn, 2022). A decolonial reading exposes three silences: they presuppose infrastructural conditions, such as stable electricity, connectivity, and cloud access, uncommon across much of sub-Saharan Africa (Dlamini & Ngoepe, 2021; Molepo & Fourie, 2020); they operate within an Anglophone tradition that renders non-English scholarship and knowledge systems less visible (Couldry & Mejias, 2021; Noble, 2021); and they reflect liberal individualist notions of privacy that sit uneasily with African relational ontologies and histories of state surveillance (Benjamin, 2020; Tladi, 2021).

Miranda Fricker's concept of epistemic injustice, elaborated in library science by Pawley (2022) and Gorichanaz (2023), offers a useful lens: testimonial injustice arises when a speaker's credibility is unduly discounted because of identity, while hermeneutical injustice arises when collective interpretive resources fail groups. AI discovery systems trained

predominantly on Global North citation data risk exactly this, algorithmically ranking African scholarship as less relevant or discoverable (Helgesson & Eriksson, 2021; Sewchurran & Gregoriou, 2022); one 2022 audit found African-authored articles received 34% fewer algorithmic recommendations than equivalent European or North American articles, even controlling for citation counts (Brock, 2020).

South Africa has formally embraced the Fourth Industrial Revolution (4IR) as a national development framework, with higher education identified as a key site for innovation (Republic of South Africa, 2020; Department of Higher Education and Training [DHET], 2022), yet translation into library-specific strategy remains slow. This is compounded by a well-documented digital divide spanning access, use, and quality dimensions (Brown & Czerniewicz, 2020): roughly 43% of South African households lack home internet access, rising to 64% in rural areas (South African Institute of Race Relations [SAIRR], 2023), and 38% of students rely primarily on mobile data (Czerniewicz et al., 2021), a connectivity mode poorly suited to data-intensive AI platforms.

South Africa's 11 official languages represent not only communicative but epistemological diversity, each carrying distinct scholarly vocabularies (Ndlovu-Gatsheni, 2021; Okonkwo & Ezema, 2022). Because large language models are trained overwhelmingly on English data, one 2023 benchmark found leading commercial models performed 67% worse on isiZulu comprehension tasks than on English (Adelani et al., 2023), an equity issue, not merely a technical one, for libraries serving multilingual student populations.

A recurring risk in this literature is technological determinism, portraying AI as an autonomous force that renders librarians' professional agency invisible (Pawley, 2022; Raber & Budd, 2022) and positions them as passive consumers rather than critical interlocutors (Gorichanaz, 2023; Munn, 2022). This study instead adopts the emerging framing of the 'librarian-arbiter': a professional who combines domain knowledge, ethical commitment, and technological literacy to actively shape AI systems as data ethicist, pedagogical guide, and custodian of epistemic justice (Anderson & Rainie, 2022; Chukwuemeka & Ugwuanyi, 2023; Buckingham, 2020; Helgesson & Eriksson, 2021). Realising this vision methodologically, the study draws on participatory action research (PAR), premised on the principle that those most affected by a problem are best placed to analyse and address it, positioning frontline librarians as co-researchers rather than mere informants (Andrade & Alves, 2022; Munn, 2022; Participatory Research in Asia [PRIA], 2021).

The novelty of this study lies in its combination of empirical breadth and conceptual contribution. Where prior South African and African scholarship has remained largely descriptive, cataloguing levels of technology access without probing governance or ethics (Okonkwo & Ezema, 2022; Molepo & Fourie, 2020), this study offers the first evidence-based, co-designed governance model for AI adoption in African academic libraries. By combining director interviews, frontline focus groups, and policy analysis across six universities, it moves beyond describing adoption towards providing an actionable framework for addressing epistemic exclusion, multilingual silencing, and professional de-

skilling as an interdependent cluster of ethical, infrastructural, and agentic concerns, an integration that existing work, with its tendency to foreground access and equipment over ethical implementation, has yet to achieve.

This study aims to develop a contextualised, actionable ethical framework to guide the strategic and responsible integration of AI in South African academic libraries. The specific objectives are to: (a) document the current state of AI adoption and governance in South African academic libraries; (b) identify the ethical risks, biases, and equity challenges specific to the South African library context; (c) examine librarians' perceptions of their professional agency in relation to AI; and (d) co-create a prototype Ethical AI Integration Framework (EAIF) grounded in local values, constitutional principles, and the principle of epistemic justice.

2. METHOD

This study employed a qualitative, multi-site case study design informed by the principles of participatory action research. The selection of a qualitative approach reflects the epistemological conviction that the ethical dimensions of AI adoption in South African academic libraries are fundamentally contextual, relational, and value-laden, and therefore not fully accessible through quantitative methods alone (Creswell & Poth, 2022). Six South African public universities were purposively selected to ensure diversity across institutional types (historically advantaged and historically disadvantaged), geographic locations (urban, peri-urban, and rural), and student population size and demographics.

Data collection involved three complementary methods. Semi-structured interviews were conducted with 18 library directors and senior managers (three per institution), focusing on institutional strategy, AI policy, and resource allocation. Forty-two focus groups were conducted with frontline librarians across the six sites, with each group comprising six to eight participants. Groups explored perceptions of AI's impact on professional roles, experiences of AI tools in practice, and views on ethical governance. At the close of each focus group, participants were invited to complete a supplementary structured survey capturing their perceptions of AI discovery-tool representativeness and their self-rated AI competencies. Of the 294 frontline librarians invited across the 42 groups (an average of seven participants per group), 276 completed the survey, a response rate of 93.9%, yielding the librarian-level quantitative dataset (n = 276) reported throughout the Results and Discussion section. A systematic document analysis was conducted of institutional IT policies, data governance frameworks, and strategic planning documents available across the six institutions. In a participatory workshop phase, librarians at each site were invited to engage in structured co-design sessions to prototype components of a local ethical AI charter.

Data was analysed using reflexive thematic analysis (Braun & Clarke, 2022), with coding conducted iteratively across three rounds. Member checking was employed to validate emerging themes with participant groups, and intersectional analysis was applied to examine how themes varied across institutional types, geographic contexts, and staff

demographic characteristics. All research was conducted with ethical clearance from the lead institution's Research Ethics Committee (Protocol REC-2024-0847), and written informed consent was obtained from all participants.

To ensure trustworthiness, the study drew on the criteria proposed by Lincoln and Guba (1985) for naturalistic inquiry. Credibility was strengthened through data triangulation across interviews, focus groups, document analysis, and the survey instrument, together with member checking of emerging themes with participant groups. Dependability and confirmability were supported by an audit trail documenting coding decisions across the three analytic rounds and by peer debriefing among the research team. Transferability was addressed through thick description of institutional contexts and purposive selection of sites for maximum variation, allowing readers to judge the applicability of the findings to their own settings.

3. RESULTS AND DISCUSSION

The analysis yielded four overarching thematic findings, each of which is reported and discussed below. Quantitative data from the document analysis and a supplementary survey instrument administered to all participant librarians (n = 276) are presented in tabular and figural form where relevant.

Theme 1: The Policy Vacuum, Strategic Absence and Institutional Fragmentation

The most striking finding from the document analysis was the near-total absence of AI-specific governance policies across the six participating institutions. Table 1 presents a comparative summary of AI policy documentation across the six sites, classified by whether each is a Historically Disadvantaged Institution (HDI) or a Historically Advantaged Institution (HAI). Only one institution (University F, a historically advantaged, research-intensive university in Gauteng) had any AI-specific policy document, and it was a draft framework still under consultation. The remaining five institutions relied on generic IT acceptable use policies, the most recent of which dated to 2019 and made no reference to machine learning, automated decision-making, or data ethics in the context of library services.

Table 1: AI Governance Policy Status Across Six South African University Libraries (2024)

Institution	Type	AI Policy Status	Data Gov. Policy	AI Training Programme
University A	HDI – Rural	None	None	None
University B	HDI – Urban	None	Generic IT only	Informal only
University C	HAI – Urban	None	Partial (2021)	None
University D	HDI – Peri-urban	None	None	None

University E	Comprehensive	None	Generic IT only	Pilot (2024)
University F	HAI – Research	Draft (2023)	Comprehensive	Formal programme

Note. HDI = Historically Disadvantaged Institution; HAI = Historically Advantaged Institution. Data sourced from institutional document analysis and director interviews (2024).

These findings are consistent with and extend the earlier survey data from LIASA (2023), which found that only 14% of South African academic library directors reported having an institutional AI governance policy. The qualitative interview data reveal the mechanisms underlying this gap. Library directors at all five institutions without policies attributed the vacuum to three interrelated factors: inadequate resourcing for strategic planning ('We are still fighting for the basics band width and database subscriptions. AI policy feels like a luxury we cannot afford': Director, University A); a lack of national-level guidance from government or library associations ('Nobody has told us what the rules are. We are waiting for someone to lead': Director, University D); and a tendency to defer governance decisions to central IT departments that lack library-specific expertise ('IT sees AI as a technical problem. We see it as an ethical one. Those are two very different conversations': Director, University C).

These findings align with Dlamini and Ngoepe's (2021) analysis of digital governance gaps in South African public institutions, and with Ndovu and Mutula's (2021) broader argument that the lack of AI regulatory frameworks in sub-Saharan Africa creates conditions of epistemic and ethical vulnerability. They also resonate with Sewchurran and Gregoriou's (2022) finding that South African higher education institutions tend to adopt technology reactively, in response to vendor pressure or peer-institution practices, rather than proactively, in accordance with locally grounded value frameworks.

Theme 2: Algorithmic Bias and the Marginalisation of African Scholarship

Focus group discussions at all six institutions surfaced deep concern about how AI-powered discovery platforms systematically disadvantage African and Global South scholarship. Figure 1 presents data from a structured survey item administered to all participant librarians (n = 276), asking them to assess the degree to which AI discovery tools at their institution adequately represented African scholarship.

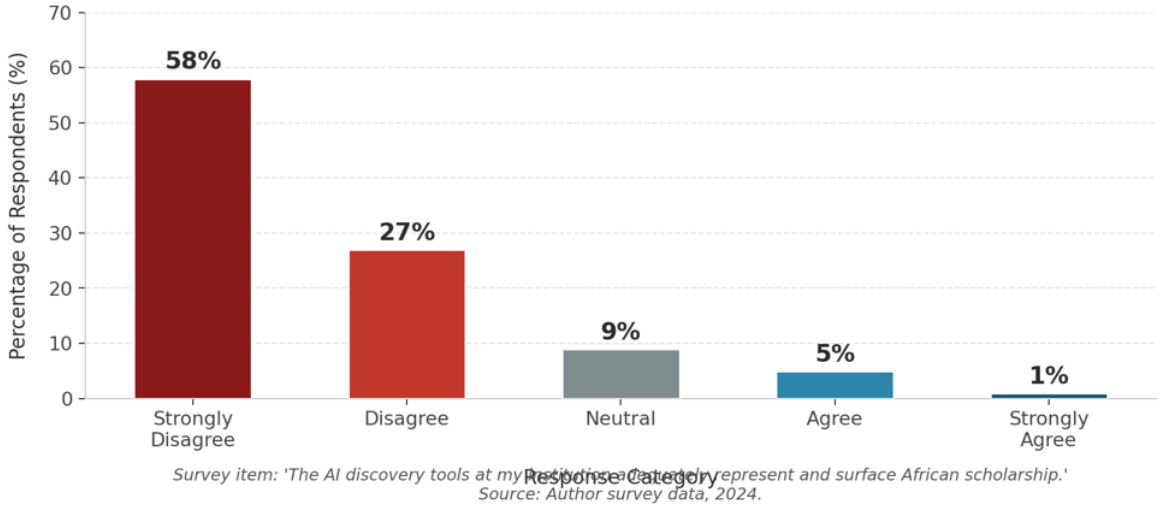


Figure 1: Librarian Perceptions of African Scholarship Representation in AI Discovery Tools

The quantitative data are striking: 85% of participant librarians disagreed or strongly disagreed that AI discovery tools at their institutions adequately surfaced African scholarship, while only 6% agreed. Qualitative focus group data provides a rich contextual explanation for these perceptions. Librarians consistently identified several mechanisms of algorithmic marginalisation. First, citation-based ranking algorithms privilege works that are highly cited within global index databases, which are themselves dominated by output from Global North publishers; African journals and institutional repositories receive comparatively few citations within these databases, regardless of their scholarly quality or local relevance (Helgesson & Eriksson, 2021; Okonkwo & Ezema, 2022).

Second, the metadata vocabularies and subject classification schemes embedded in AI-driven cataloguing tools, primarily derived from the Library of Congress Subject Headings, contain cultural biases and terminological gaps that render African intellectual traditions less findable (Brock, 2020; Noble, 2021). Third, conversational AI reference tools trained on global library corpora are ill-equipped to assist students seeking African-authored scholarship, often returning recommendations dominated by Northern sources even for explicitly African research topics.

These findings constitute a form of institutional epistemic injustice as theorised by Fricker and applied to the library context by Pawley (2022) and Gorichanaz (2023). The practical implications are serious: when students researching African political philosophy, indigenous knowledge systems, or sub-Saharan public health are systematically directed away from African scholarship by AI recommendation engines, the library, despite its professional commitments, becomes an instrument of epistemic marginalisation rather than epistemic empowerment. This finding reinforces the arguments of Couldry and Mejias (2021) regarding data colonialism and the empirical findings of Brock (2020) and Mohamed et al. (2020) regarding the racialised dimensions of algorithmic discrimination.

Theme 3: The Capacity Deficit: Staff Training, Digital Literacy, and Professional Identity

Another major issue is the gap between the AI skills that librarians know they need and what they have. For example, when asked to rate themselves across eight important AI skills, most librarians said they fell short in key areas that matter for ethical library work.

Table 2: Librarian Self-Rated AI Competency Versus Perceived Importance (n = 276)

AI Literacy Domain	Mean Self-Rating (1–5)	Mean Importance Rating (1–5)	Competency Gap
Understanding machine learning basics	2.1	4.3	2.2
Evaluating algorithmic bias	1.8	4.7	2.9
Data privacy and protection law	2.6	4.8	2.2
Algorithmic transparency advocacy	1.6	4.5	2.9
Multilingual AI tool evaluation	1.4	4.6	3.2
Instructing users in AI literacy	2.3	4.4	2.1
Epistemic justice and AI	1.3	4.5	3.2
Participatory AI design processes	1.2	4.2	3.0

Note. Self-rating scale: 1 = No competency, 5 = Expert. Importance scale: 1 = Not important, 5 = Essential. Competency gap = Importance rating minus self-rating.

The data reveals a pervasive and severe competency gap across all eight domains, with the largest gaps recorded for multilingual AI tool evaluation (3.2), epistemic justice and AI (3.2), and algorithmic transparency advocacy (2.9). Particularly concerning is the domain of epistemic justice and AI, where the mean self-rating of 1.3 indicates that most participant librarians have received no preparation in this conceptual area, despite rating it as highly important for ethical practice. These figures are consistent with Molepo and Fourie's (2020) finding that fewer than 30% of South African library staff had received AI-related training, and with the more recent LIASA (2023) survey data indicating that this proportion had risen to only 37% by 2023, a modest improvement given the pace of technological change.

Librarians in the study described this skills gap as more than just a lack of training; it sometimes made them question their place in the profession. Many said they felt sidelined by new technology, especially when students trusted the chatbot's answer over their expert advice. This was even more common at under-resourced universities, where staff had fewer opportunities for AI training. As one librarian put it, 'I have a master's degree and decades of experience, but when the chatbot gives a wrong answer, and the student believes it anyway, I feel invisible.' Importantly, focus group discussions also revealed significant

reserves of professional agency and critical insight among librarians. Participants consistently demonstrated sophisticated analyses of AI's limitations and risks, even in the absence of formal AI literacy training. This aligns with Pawley's (2022) argument that librarians possess underutilised ethical expertise that, if properly supported and resourced, could position them as powerful advocates for equitable AI governance. The challenge, in this reading, is not a deficit of professional wisdom but a deficit of institutional recognition and investment.

Theme 4: Towards the Librarian-Arbiter Agency, Co-Creation, and Ethical Stewardship

When librarians were invited to co-create solutions, the results were inspiring. In workshops across all six universities, librarians worked together to draft their own AI ethics charters. Despite differences between the institutions, the core values they chose were strikingly similar and deeply rooted in the needs and realities of their own communities.

Table 3: Principles Articulated in Prototype AI Ethics Charters Across Six Institutions

Governance Principle	**Frequency Endorsed (of 6 sites)**	**Representative Articulation**
Epistemic justice and African scholarship inclusion	6/6	AI must actively surface African voices, not bury them
Algorithmic transparency and explainability	6/6	Users have the right to know how recommendations are generated
Data sovereignty and user privacy	6/6	Student data is not a commercial resource
Multilingual equity and accessibility	5/6	AI tools must work in our students' home languages
Participatory governance and librarian agency	6/6	Librarians must be co-designers, not passive users
Decolonial metadata and classification	4/6	Subject headings must reflect African knowledge systems
Ongoing algorithmic auditing	5/6	Bias is not a one-time problem; we must keep checking
Community-centred accountability	6/6	Technology must serve our communities, not replace them

Note. Data derived from participatory co-design workshop outputs across six university library sites (2024). Articulations are representative quotes from workshop documentation.

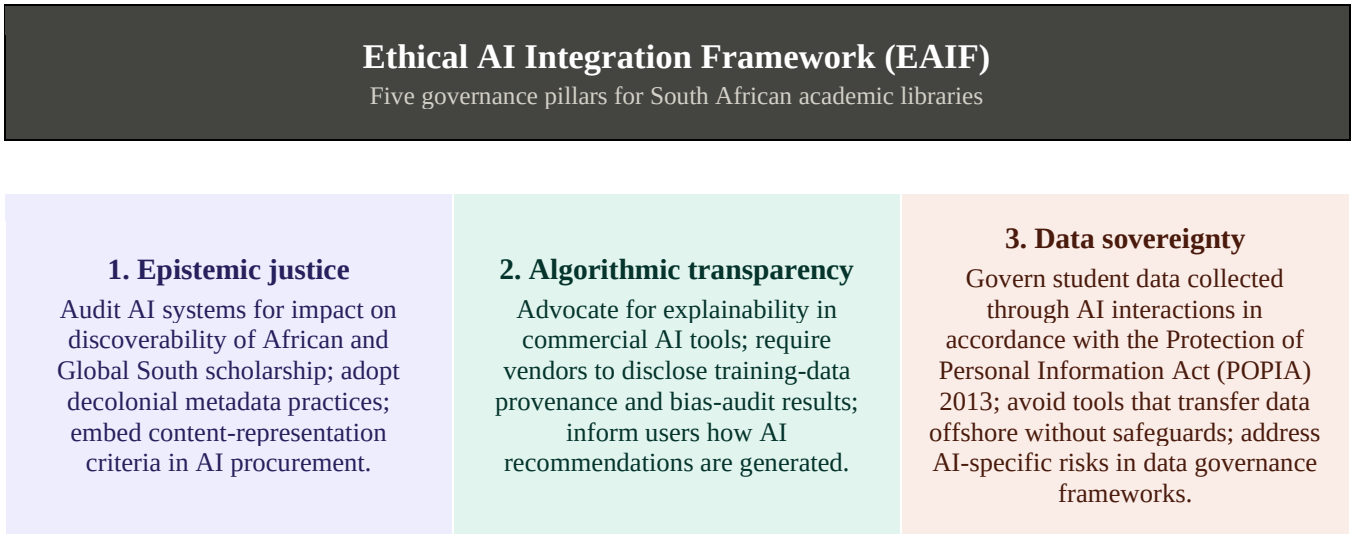
The unanimity of endorsement for epistemic justice, algorithmic transparency, data sovereignty, participatory governance, and community-centred accountability across all six sites' institutions, which differ significantly in size, resources, historical positioning, and geographic context, is a finding of considerable significance. It suggests that there is a latent,

shared ethical vocabulary among South African academic librarians awaiting institutional recognition and formal codification. The fact that these principles were generated not from expert elicitation or top-down policy development but through facilitated practitioner deliberation provides strong evidence for the methodological value of participatory action research in the co-creation of professional governance frameworks (Andrade & Alves, 2022; Participatory Research in Asia [PRIA], 2021). These workshops also showed how librarians see themselves not as people who will be replaced by AI, but as essential partners to technology. They want to work alongside AI, using their own judgment and local knowledge to ensure library services remain fair and relevant. As one participant said, 'The algorithm can search. Only I can care about whether what it finds is just.'

A Proposed Ethical AI Integration Framework

Synthesising the findings from all four phases of data collection and from the co-design workshops, this study proposes an Ethical AI Integration Framework (EAIF) for South African academic libraries, see Figure 2. EAIF is structured around five governance pillars, each of which encompasses a set of operational principles and practical governance actions. The framework is intended not as a prescriptive template but as a flexible scaffold that individual institutions can adapt to their specific community contexts and institutional values.

The five pillars of the EAIF are: (1) Epistemic Justice, which requires that AI systems be audited for their impact on the discoverability of African and Global South scholarship, that decolonial metadata practices be adopted, and that library AI procurement decisions include explicit criteria for content representation; (2) Algorithmic Transparency, which mandates that librarians advocate for explainability in commercial AI tools, that institutions require vendors to disclose training data provenance and bias audit results, and that users be informed in accessible terms about how AI-driven recommendations are generated;



4. Multilingual equity

Systematically evaluate AI tool performance across all 11 official languages; invest in developing or procuring multilingual AI capabilities; advocate for greater African-language AI investment.

5. Participatory stewardship

Establish librarians as co-designers of AI systems; mandate ongoing professional development in AI literacy and ethics; create institutional mechanisms for community consultation.

Figure 2: Ethical AI Integration Framework

(3) Data Sovereignty, which insists that student data collected through AI interactions be governed in accordance with South Africa's Protection of Personal Information Act (POPIA) 2013, that libraries avoid AI tools that transfer user data to offshore servers without appropriate safeguards, and that institutional data governance frameworks explicitly address AI-specific risks; (4) Multilingual Equity, which requires systematic evaluation of AI tool performance across all 11 official languages, investment in the development or procurement of multilingual AI capabilities, and advocacy with vendors and national bodies for greater investment in African language AI; and (5) Participatory Stewardship, which establishes librarians as co-designers of AI systems, mandates ongoing professional development in AI literacy and ethics, and creates institutional mechanisms for community consultation on AI adoption decisions.

To provide examples of how these were applied: under the Data Sovereignty principles, clauses were added to a draft policy at University F requiring that AI vendor contracts prohibit student interaction logs from being sent to non-South African servers, and that chatbot conversation data should be deleted after 90 days, unless a student affirmatively consents to retaining it-which, in effect, ensures that POPIA is implemented as opposed to simply mentioned. Under the Multilingual Equity principle, the proposal included piloting isiZulu question processing on library chatbots at two of the six co-design workshops in collaboration with local natural language processing (NLP) work (such as Masakhane), and requiring vendors' tool procurement to be benchmarked against isiZulu comprehension weaknesses indicated in Adelani et al. (2023), rather than relying on vendor claims of multilingual support.

Here are three ways the EAIF is unique from prior AI ethics frameworks: First, while previous frameworks such as the UNESCO Recommendation (2021), the IFLA statement (2022) and the ALA principles (2021) express lofty, global, idealist aspirations, the EAIF was inductively co-designed with librarians working on the front lines of six institutions and led by its library directors, with all pillars anchored in real harms backed up by the actual testimony of local stakeholders rather than global ideals. Second, the EAIF gives data sovereignty and multilingual equity explicit status as standalone operational pillars rather than treating these either as implicit categories under "fairness" or "privacy," due to challenges with issues related to the failure of algorithms in isiZulu and POPIA's offshore data transfer threat, which global guidelines fail to mention in any meaningful way. Third, rather than treating librarians merely as end-users and implementers of work crafted elsewhere, the EAIF positions them as co-authors, collaborators, and guardians of

governance: as participatory Stewards, alongside epistemic justice, algorithmic fairness and data sovereignty, rather than simply another locally designed variation on a globally adopted paradigm.

As the EAIF was designed with and for the South African reality, its reasoning is probably valid far beyond the South African context. A multilingual population poorly represented in the English-centric AI training set and other factors underlying the design-such as a colonial tradition of the epistemic hierarchy perpetuating the discrediting of local research, inconsistent digital infrastructures within both urban and rural institutions, and library staff cut from AI acquisition decisions-are to a varying extent to be found throughout the library systems of the Global South, whether in Nigeria, Kenya, India, regions of Latin America, and Southeastern Asia. Five dimensions the EAIF are purposefully built as, in effect, domains amenable to governance rather than fixed procedures: institutions in other contexts can insert their respective languages to enforce multilingual equity, national data protection law to assert data sovereignty, or ignored scholarly tradition to attain the recognition of their respective epistemologies, respectively. Further, comparative research, explored at greater length in Recommendations, needs to be undertaken to test and adapt this model in other contexts; however, the process-to empower, through co-construction, rather than simply receive-may offer useful precedent for other library and information service systems in the South.

4. CONCLUSION

This study identified the ethical issues and governance void in South African academic libraries in the era of AI on the evidence of six university libraries. The study found three conditions occurring together across these sites: an almost complete absence of AI-specific policy governance; AI discovery tools with an inordinate bias against African scholarship; and library staff with the relevant expertise being rarely asked to participate in the adoption process. The study then proposes the EAIF model as a potential starting point, though not the panacea for the ethical challenges, community led model for AI governance in SA libraries, based upon 5 pillars; epistemic justice, algorithmic transparency, data sovereignty, multilingual equity, and participatory stewardship that were code signed with librarians and framed according to SA constitutional values that position the librarian role centre stage in AI adoption at institutions.

Apart from offering a practical approach, this study advances theories within Library and Information Science (LIS) in two main ways. Firstly, by expanding on Fricker's (2007) view of epistemic injustice as an individual/inter-personal process to an institutional and systemic approach, articulating the way how discovery systems, by using algorithms systematically, re-establish testimonial and hermeneutical injustices within the professional information environment of the library. Secondly, by advancing the concept of librarianship as a more “active ethical agent,” a " librarian-arbiter” and not a “passive receiver” of vendor systems, which aims to challenge technologically deterministic discourse within LIS and offers LIS scholars a fact-supported counter-alternative to vendor-driven technological

determinist arguments concerning AI use. This study adds to an expanding corpus of scholarship that argues that ethics of AI should be contextually relevant rather than internationally imposed wholesale. South librarians have capabilities and dedication needed to steer this movement but require formal institutional legitimation and funding that comes with institutional space and resources.

5. STUDY LIMITATIONS

Several limitations of this study should be acknowledged. First, the sample of six institutions, though purposively selected for diversity, cannot be considered representative of the full population of 26 South African public universities, and the findings should be interpreted as illuminative rather than generalisable in a statistical sense. Second, the participatory action research design, while generating rich and locally grounded data, introduces the risk of facilitation bias; the co-design workshops were led by researchers with established commitments to epistemic justice frameworks, which may have influenced the principles that librarian participants prioritised.

Third, the study was conducted during 2024, a period of rapid change in the AI landscape; some of the specific tools and capabilities described may have evolved significantly by the time of publication, potentially affecting the continuing relevance of findings. Fourth, the study did not include the perspectives of library users, students, and academic staff, whose experiences with AI tools in library contexts would provide an important complementary data source. Fifth, the EAIF proposed in this study has not yet been implemented and evaluated in practice; its efficacy as a governance instrument remains to be tested through longitudinal implementation research.

6. RECOMMENDATIONS

Based on the study's findings, the following recommendations are directed to key stakeholders in the South African academic library ecosystem.

a. Library directors and senior management

1) Make AI governance an issue of strategic importance; convene ad-hoc working groups to build institutional policy on AI ethics based on EAIF principles.
2) Supplement insufficient, one-size-fits-all IT acceptable use policies with policies crafted at the library level in collaboration with our front-line librarians and our users.
3) Harness purchasing power via consortia like the Committee of Higher Education Libraries of South Africa (CHELSA) and engage in contract negotiation to include data sovereignty guarantees, mandatory bias transparency, and multinational performance benchmarking requirements for AI tools (Anderson & Rainie, 2022; Breed, 2023).

b. Library and Information Association of South Africa (LIASA)

1) Create a national AI ethics policy for the library profession following in the vein of the IFLA (2022) statement but based primarily in terms of South African

constitutional law principles, the provisions of the POPIA Act and its unique trilingual national status.

2) Drive awareness of AI literacy and data ethics as a skill set central to LIS curriculum development standards, which will enable fresh librarians with the capability to perform as AI ethical guardians (Chukwuemeka & Ugwuanyi, 2023; Molepo & Fourie, 2020).

c. Department of Higher Education and Training (DHET) and Department of Science and Innovation (DSI)

1) Create specific budget allocations to address the AI literacy training needs for the staff of academic libraries, just as has already happened to address the training of staff at public higher learning institutions in digital literacy, which serves all staff at the institution.
2) Require that biased audits are performed on all AI tools to be utilised by public-funded educational institutions, and to account for the epistemic-injustice risk as part of the national AI policy (DSI, 2022; Republic of South Africa, 2020)

d. Library and information science researchers

1) Perform longitudinal EAIF implementation studies to ascertain practical viability as a governance instrument.
2) Commission user studies regarding how students of other linguistic backgrounds interact in their experience of library services with the help of the EAIF.
3) Work towards a cross-regional expansion into other African nation contexts grappling with colonial epistemic legacy, tech deficit, capacity shortage (Okinkwo & Ezema, 2022; Ndovu & Mutula, 2021).

e. AI vendors and technology developers

1) Invest in multilingual, culturally appropriate development rather than assuming Global North defaults transfer directly to African contexts.
2) Submit to independent bias audits and engage library professionals as governance partners rather than passive customers (Adelani et al., 2023; Noble, 2021).
3) Support multilingual development, which respects Africa's cultural landscapes, instead of assuming Global North's design features auto-translate to Africa.
4) Accept third-party bias auditing, where Library professionals are active players as government partners, as in Adelani et al. (2023) and Noble (2021.

ACKNOWLEDGEMENTS

The author would like to thank all the employees working in the libraries who have participated in this study.

FUNDING STATEMENT

This study was not supported by any grants from funding bodies in the public, private, or non-profit sectors.

AVAILABILITY OF DATA AND MATERIALS

The data supporting this study's findings are available on request from the corresponding author.

ETHICS STATEMENT

Before the research began, ethical approval was acquired for this study.

REFERENCES

Adelani, D. I., Alabi, J., Alkhamissi, B., Ojo, A., Emezue, C., Nasution, R., ... & Ruder, S. (2023). MasakhaNER 2.0: Africa-centric transfer learning for named entity recognition. *Transactions of the Association for Computational Linguistics, 11*, 21–37. https://doi.org/10.1162/tacl_a_00536

American Library Association. (2021). *ALA statement on artificial intelligence*. https://www.ala.org/tools/ethics/artificial-intelligence

Anderson, J., & Rainie, L. (2022). *AI and the future of humans: Expert perspectives on governance*. Pew Research Centre.

Andrade, M., & Alves, P. (2022). Participatory action research in library settings: A methodological review. *Journal of Librarianship and Information Science, 54*(3), 412–429. https://doi.org/10.1177/09610006211015022

Bawden, D., & Robinson, L. (2022). *Introduction to information science* (2nd ed.). Facet Publishing.

Benjamin, R. (2020). *Race after technology: Abolitionist tools for the new Jim Code*. Polity Press.

Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE Publications.

Breed, M. (2023). Library systems report 2023: Advancing AI integration. *American Libraries, 54*(5), 22–35.

Brock, A. (2020). Critical technocultural discourse analysis. *New Media & Society, 22*(6), 1012–1030. https://doi.org/10.1177/1461444820914736

Brown, C., & Czerniewicz, L. (2020). Debunking the 'digital native': Beyond digital apartheid, towards digital democracy. *Journal of Computer Assisted Learning, 36*(5), 583–594. https://doi.org/10.1111/jcal.12476

Buckingham, D. (2020). A critical engagement with digital literacy. *Journal of Media Literacy Education, 12*(2), 114–119.

Choi, W., An, S., & Han, G. (2022). AI-powered reference services: User satisfaction and interaction quality. *College & Research Libraries, 83*(4), 512–530. https://doi.org/10.5860/crl.83.4.512

Chukwuemeka, A., & Ugwuanyi, C. (2023). Artificial intelligence and the changing roles of academic librarians in Africa. *African Journal of Library, Archives and Information Science, 33*(1), 45–61.

Couldry, N., & Mejias, U. A. (2021). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media, 20*(4), 336–349. https://doi.org/10.1177/1527476418796632

Creswell, J. W., & Poth, C. N. (2022). *Qualitative inquiry and research design: Choosing among five approaches* (5th ed.). SAGE Publications.

Czerniewicz, L., Agherdien, N., Badenhorst, J., Bhatt, D., Chili, M., De Villiers, M., ... & Wiens, K. (2021). A wake-up call: COVID-19 and its impact on the core business of higher education institutions in South Africa. *Studies in Higher Education, 46*(1), 1–16. https://doi.org/10.1080/03075079.2020.1813351

Department of Higher Education and Training. (2022). *Digital transformation in higher education: A strategic framework 2022–2027*. Government Printer.

Department of Science and Innovation. (2022). *South Africa's national artificial intelligence policy framework*. Government Printer.

Dlamini, N., & Ngoepe, M. (2021). Digital transformation in South African public institutions: Governance challenges and opportunities. *South African Journal of Information Management, 23*(1), 1–11. https://doi.org/10.4102/sajim.v23i1.1312

Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.

Gorichanaz, T. (2023). The librarian as ethical AI steward: Professional identity in the age of algorithmic systems. *Library Quarterly, 93*(2), 101–118. https://doi.org/10.1086/724021

Helgesson, G., & Eriksson, S. (2021). Plagiarism in research and the inequitable treatment of scholars from the Global South. *Research Ethics, 17*(1), 5–18. https://doi.org/10.1177/1747016120954412

International Federation of Library Associations and Institutions. (2022). *IFLA statement on libraries and artificial intelligence*. https://www.ifla.org/publications/node/93554

Islam, M. A., & Tsuji, K. (2021). AI-driven reference services: A review of implementations and outcomes. *Journal of the American Society for Information Science and Technology, 72*(8), 1093–1107. https://doi.org/10.1002/asi.24453

Library and Information Association of South Africa. (2023). *State of academic libraries in South Africa: Annual survey report 2023*. LIASA.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.

Luthmann, A., & Park, J. (2022). Automated metadata enrichment using natural language processing: Challenges for multilingual collections. *Cataloguing & Classification Quarterly, 60*(5), 348–366. https://doi.org/10.1080/01639374.2022.2056128

Masakhane. (n.d.). *Masakhane: African languages in AI*. https://www.masakhane.io/

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

Molepo, M., & Fourie, I. (2020). Digital divide, digital literacy and academic libraries in South Africa. *South African Journal of Libraries and Information Science, 86*(2), 1–14. https://doi.org/10.7553/86-2-1810

Munn, L. (2022). The uselessness of AI ethics. *AI and Society, 37*(4), 1–6. https://doi.org/10.1007/s00146-021-01180-8

Ndlovu-Gatsheni, S. J. (2021). *Decolonisation, development and knowledge in Africa: Turning over a new leaf*. Routledge.

Ndovu, T., & Mutula, S. (2021). AI governance in sub-Saharan African public institutions: A systematic review. *Information Development, 37*(4), 617–635. https://doi.org/10.1177/02666669211020428

Noble, S. U. (2021). *Algorithms of oppression: How search engines reinforce racism* (2nd ed.). New York University Press.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2022). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

OCLC. (2023). *Artificial intelligence in libraries: A global benchmark study*. OCLC Research. https://www.oclc.org/research/publications/2023/oclc-ai-libraries-global.html

Okonkwo, C., & Ezema, I. J. (2022). Challenges and prospects of AI adoption in African academic libraries. *Library Hi Tech, 40*(3), 718–735. https://doi.org/10.1108/LHT-05-2021-0176

Olteanu, A., Castillo, C., Diaz, F., & Kıcıman, E. (2020). Social data: Biases, methodological pitfalls, and ethical boundaries. *Frontiers in Big Data, 2*, 13. https://doi.org/10.3389/fdata.2019.00013

Park, J., & Lu, C. (2021). Machine learning for subject cataloguing: A review of approaches and implications for metadata professionals. *Library Resources & Technical Services, 65*(1), 34–52. https://doi.org/10.5860/lrts.65n1.34

Participatory Research in Asia. (2021). *Participatory action research: Principles and practice* (3rd ed.). PRIA.

Pawley, A. L. (2022). Unequal legacies: Race and multiculturalism in the LIS curriculum. *Library Quarterly, 92*(3), 269–291. https://doi.org/10.1086/719978

Raber, D., & Budd, J. M. (2022). Information as sign: Semiotics and information science. *Journal of Documentation, 78*(1), 1–18. https://doi.org/10.1108/JD-04-2020-0054

Republic of South Africa. (2020). *Presidential Commission on the Fourth Industrial Revolution: Final report*. The Presidency.

Sadeh, T. (2022). The evolution of library discovery systems. *Information Technology and Libraries, 41*(2), 1–22. https://doi.org/10.6017/ital.v41i2.14219

Sewchurran, K., & Gregoriou, P. (2022). Ethical technology adoption in South African higher education: A critical discourse analysis. *South African Journal of Higher Education, 36*(3), 212–231. https://doi.org/10.20853/36-3-4852

South African Institute of Race Relations. (2023). *South Africa survey 2023*. SAIRR.

Tammaro, A. M., Dick, A., & Salarelli, A. (2021). Artificial intelligence and libraries: Toward a future agenda for research and practice. *JLIS.it, 12*(3), 1–12. https://doi.org/10.4403/jlis.it-12740

Tladi, L. (2021). Privacy rights in the digital age: A South African perspective. *Potchefstroom Electronic Law Journal, 24*(1), 1–27. https://doi.org/10.17159/1727-3781/2021/v24i0a8784

UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organisation. https://doi.org/10.54676/GPPC3851

Universities South Africa. (2024). *Higher education sector statistics 2023*. USAf.

Vakkari, P. (2023). Searching and information seeking in digital environments: An overview. *Information Processing & Management, 60*(1), 103153. https://doi.org/10.1016/j.ipm.2022.103153

Vieira, D. V., & Garcez, P. M. (2021). Machine learning for collection development in academic libraries: A predictive model for acquisition decisions. *Library & Information Science Research, 43*(4), 101121. https://doi.org/10.1016/j.lisr.2021.101121

World Bank. (2024). *Poverty and inequality platform: South Africa country profile 2023*. World Bank Group. https://pip.worldbank.org/country-profiles/ZAF

Zhang, Y., & Arlitsch, K. (2022). Predictive analytics for library collection management: Methodologies and outcomes. *Library Collections, Acquisitions, & Technical Services, 45* (1), 1–14. https://doi.org/10.1080/14649055.2021.1989762

Zou, J., & Schiebinger, L. (2021). AI can be sexist and racist—it's time to make it fair. *Nature, 559*(7714), 324–326. https://doi.org/10.1038/d41586-018-05707-8