

Community-Governed AI for Endangered African Languages and Scripts^{*}

-A Multicultural Education Framework for Digital Language Justice-

Ghebru Bethel

(Busan University of Foreign Studies)

1. Introduction
2. Multicultural Education, Linguistic Minoritization, and Digital Language Justice
3. A Function-Based Typology of AI Applications for Endangered African Linguistic Heritage
4. Illustrative Preservation Scenarios and a Cross-Cutting Script-Infrastructure Layer
5. A Seven-Phase Roadmap for Community-Governed AI Preservation
6. Risks, Governance Architecture, Institutional Partnerships, and Policy Implications
7. Conclusion

* 이 논문은 2025년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임
(NRF-2025S1A6B5A01003729)

Abstract

This article examines how artificial intelligence (AI) can support the preservation and revitalization of endangered African languages and scripts through the lens of multicultural education and digital language justice. Bringing Banks's theory of multicultural education into conversation with May's analysis of linguistic minoritization, together with culturally sustaining pedagogy and community data sovereignty, it develops four analytical dimensions spanning recognition, rights, community agency, and intergenerational learning. It proposes a function-based typology of AI applications, examines Ongota, Yaakunte/Yaaku, Nluu, and African Ajami manuscript traditions as illustrative preservation scenarios, and treats African script technology as a cross-cutting infrastructural layer rather than an equivalent case. The article further develops a seven-phase implementation roadmap and identifies five interrelated risk domains: extractive data governance, exposure of restricted knowledge, AI error and epistemic distortion, infrastructural unsustainability, and unequal institutional power. It argues that AI-supported preservation should be evaluated not only through technical performance or data volume, but through its contributions to linguistic recognition, educational access, community authority, intergenerational learning, and cultural continuity.

[Key words] artificial intelligence, endangered African languages and scripts, multicultural education, linguistic minoritization, community governance, digital language justice, language revitalization

1. Introduction: Community-Governed AI, Linguistic Minoritization, and Multicultural Education

Artificial intelligence (AI) is increasingly presented as a potentially important resource for the documentation, revitalization,

and digital inclusion of endangered and low-resource languages, particularly in Africa, where exceptional linguistic diversity coexists with profound inequalities in educational provision, digital representation, and computational support. Many African languages remain central to family life, oral tradition, and community identity yet are marginalized in education, administration, media, and digital platforms, and severely endangered languages often lack even basic resources—recorded corpora, validated orthographies, dictionaries, fonts, keyboards, speech datasets, OCR, and locally relevant learning materials. This reflects a wider pattern of linguistic exclusion in natural language processing, in which a small number of globally dominant languages receive a disproportionate share of datasets, model development, and commercial investment (Joshi et al. 2020).

The problem, however, is not simply that some languages have fewer technological resources than others: linguistic endangerment is connected to historically produced relations of power. In many postcolonial African states, former colonial languages and a small number of nationally dominant African languages enjoy greater institutional authority, while other languages become minoritized not because their speakers are numerically insignificant but because their languages are assigned limited educational, political, and technological value. The erosion of an endangered language may therefore weaken access to oral histories, ecological knowledge, manuscript traditions, and community-specific ways of interpreting experience, making preservation an issue of educational equality, cultural recognition, and collective participation rather than a linguistic or technological concern alone.

This article approaches these concerns through a multicultural-education perspective informed by both educational transformation and linguistic rights. Banks's multidimensional account of multicultural education is valuable because it emphasizes the inclusion of marginalized knowledge, critical examination of how knowledge is constructed, equitable pedagogy, and institutional transformation. Yet the African linguistic contexts addressed here cannot be adequately understood only through a framework developed primarily in relation to racial and ethnic diversity in the United States. May's work on language, minority rights, and linguistic minoritization offers a complementary perspective by drawing attention to the institutional hierarchies through which dominant languages acquire legitimacy while minoritized languages are restricted to private, symbolic, or informal domains (Banks 2009; May 2012). Together these perspectives frame AI-supported preservation through four interconnected educational concerns: linguistic recognition and representation, equitable access and language rights, community agency and knowledge authority, and intergenerational, culturally sustaining learning.

From this perspective, multicultural education concerns the conditions under which historically marginalized communities can participate in the production, validation, teaching, and transmission of knowledge, not merely the representation of diverse languages in textbooks or occasional cultural celebration: an endangered language may be documented without becoming educationally accessible, a manuscript digitized without being returned to its community, and a learning application may display a minority language without allowing speakers, elders, teachers, or young people to

determine its content and use. The significance of AI-supported preservation therefore depends on whether it expands recognition, participation, learning, and community control. Recent policy developments provide relevant context: the African Union's Continental Artificial Intelligence Strategy identifies AI capacities for African languages as part of the continent's technological agenda (African Union 2024), and UNESCO has emphasized the potential of digital technologies for documenting and revitalizing endangered languages (UNESCO 2025), connecting preservation to wider debates about AI governance and technological sovereignty.

Policy recognition does not, however, resolve the practical and ethical challenges of implementation. AI may assist in cleaning and segmenting recordings, proposing transcriptions, recognizing printed or handwritten scripts, and generating draft learning materials, but many endangered African languages have very limited data, few fluent speakers, unstable or contested orthographies, and limited institutional support (Besacier et al. 2014; Joshi et al. 2020). Manuscript collections may be dispersed among families, religious institutions, and overseas archives (Mumin & Versteegh 2014; Ngom 2016), and some formally encoded African scripts remain difficult to use because of inadequate fonts, keyboards, and publishing tools. The usefulness of AI therefore depends on careful adaptation to each community's linguistic ecology, resources, cultural norms, and institutional capacities. These conditions also generate ethical and educational risks, including extractive data relations, exposure of restricted knowledge, and unverified machine-generated language, analyzed with their governance requirements, in Section 6.

The central argument of this article is that AI can contribute to the preservation and revitalization of endangered African languages and scripts only when technological tools are embedded within community-governed preservation ecosystems: the objective should not be the production of datasets, models, or archives as ends in themselves, but the strengthening of conditions under which communities can record, access, interpret, teach, and use their linguistic heritage. This requires moving beyond technological solutionism, since a language is not sustained by transcription and storage but when people continue to hear, speak, read, write, and teach it in meaningful domains, a point long emphasized in revitalization research (Fishman 1991; Grenoble & Whaley 2006; Hinton et al. 2018). Digital technologies contribute when they support community learning and everyday practice (Galla 2016) and when technical outputs are returned in usable, locally maintainable forms. AI-supported preservation should therefore be assessed not only through technical measures such as accuracy or dataset size, but through its contribution to recognition, educational access, community agency, and cultural continuity.

This article addresses a gap between rapidly expanding public discussion about AI for African languages and the more differentiated requirements of endangered-language preservation. Existing attention has focused largely on machine translation, large language models, and the inclusion of relatively widely spoken African languages in global platforms, but these developments do not fully address the needs of severely endangered oral languages, small revitalization movements, manuscript traditions, or script communities, whose specific requirements are examined in Sections 3 and 4.

The article therefore conceptualizes endangered African languages and scripts as distinct but interconnected preservation ecologies. Rather than treating AI as a universal solution, it asks how particular technologies may be aligned with specific linguistic, educational, cultural, and institutional conditions. The analysis is organized around four dimensions of multicultural education—linguistic recognition and representation, linguistic rights and equitable access, community agency and knowledge authority, and intergenerational and culturally sustaining learning—which are developed in detail in Section 2.

1.1. Research Design, Analytical Questions, and Case Selection

Methodologically, this article adopts a conceptual, applied, and policy-oriented approach; it is neither a systematic literature review nor an empirical multiple-case study. It synthesizes scholarship from language documentation and revitalization, low-resource natural language processing (NLP), digital heritage, multicultural and multilingual education, AI ethics, Indigenous and community data governance, and African script and manuscript studies, integrating the recurring preservation functions, educational objectives, and governance requirements identified there into an author-developed typology, illustrative scenarios, a seven-phase roadmap, and a risk-and-governance framework.

The article is guided by three questions: what functions AI can realistically perform in preserving and revitalizing endangered African languages and scripts; how AI applications should be adapted to different oral, written, manuscript, script,

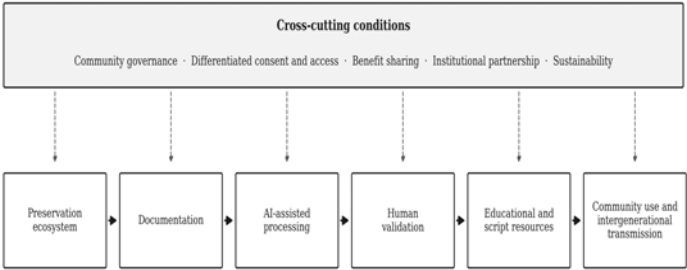
and educational ecologies; and what multicultural-education, community-governance, and institutional conditions are required for AI-supported preservation to advance linguistic recognition, equitable access, community agency, and long-term language use.

The illustrative cases are not presented as complete empirical case studies, nor as statistically representative of Africa's linguistic diversity. They are selected because they represent contrasting preservation ecologies: Ongota in Ethiopia illustrates the urgency of oral documentation where fluent speakers are extremely few; Yaakunte/Yaaku in Kenya, community revitalization connected to ecological memory and youth participation; Nluu in South Africa, pronunciation, orthographic development, and reparative recovery centred on elder authority; and African Ajami manuscript traditions, OCR, transliteration, archival recovery, and culturally governed access. A comparative design is deliberately retained because the article's analytical object is precisely this variation among preservation ecologies, which a single-country study could not demonstrate; an in-depth national implementation study—for example, focused on Ethiopia—remains an important direction for future research.

The discussion of African script technologies—N’Ko, Adlam, Bamum, Tifinagh, Vai, Ethiopic, and others—occupies a different analytical level from these four scenarios. They are examined not as a fifth equivalent case but as a cross-cutting infrastructural layer of fonts, keyboards, OCR, and publishing tools that may serve multiple script communities, since formal encoding alone does not guarantee everyday digital usability.

The article aims to contribute in three ways: it develops a function-based typology of AI applications for endangered

African linguistic heritage; it uses the four illustrative scenarios and the cross-cutting script-infrastructure layer to show why AI tools must be adapted to distinct preservation ecologies rather than transferred as uniform packages; and it proposes a seven-phase roadmap linking technical and archival activities to multicultural-education outcomes.



Note. Community governance, consent, benefit sharing, institutional partnership, and sustainability operate across the full preservation cycle. Source: Developed by the author.

〈Figure 1〉 Conceptual framework: From documentation to community-governed educational use

Section 2 develops the theoretical framework; Section 3 presents the typology; Section 4 examines the scenarios and the script-technology layer; Section 5 develops the roadmap; Section 6 analyses risks, governance, and policy implications; and Section 7 concludes.

2. Multicultural Education, Linguistic Minoritization, and Digital Language Justice

The preservation of endangered African languages and

scripts is often framed as a problem of cultural heritage or technological innovation. This section develops a different framing: preservation is also an educational project concerning whose knowledge is recognized, who has access to learning, and who exercises authority over linguistic resources in increasingly digital environments.

The article accordingly understands multicultural education as a process of institutional and pedagogical transformation rather than the symbolic celebration of diversity, developing this understanding through Banks's framework, May's account of linguistic minoritization, culturally sustaining pedagogy, and community data governance, integrated into the four analytical dimensions used throughout the article.

2.1. From the Recognition of Diversity to Educational Transformation

Banks's theory provides an important starting point because it presents multicultural education as a multidimensional project rather than a matter of curricular add-ons (Banks 2009): content integration asks whether marginalized languages and knowledge systems are represented in curricula; knowledge construction asks why some languages become associated with modernity while others are confined to folklore; equity pedagogy asks whether learners can study through and with their linguistic heritage; and institutional transformation requires schools, universities, and archives to change their own practices rather than expecting minoritized communities to adapt unilaterally.

Applied to AI-supported preservation, these dimensions

generate questions that are simultaneously educational and technological: not only whether AI can generate a transcription or recognize a script, but whether the resulting resource enters educational practice, whose categories structure it, and whether community members have authority to correct or reject its contents. A technically successful system may still reproduce educational inequality if it extracts community knowledge or produces materials inaccessible to local teachers and learners.

Banks's framework must nevertheless be used with contextual caution: it developed largely in relation to racial and ethnic inequalities in the United States, and the language politics of postcolonial African states involve different historical conditions. Colonial languages often remain dominant in schooling and publishing, and nationally influential African languages may receive more state support than smaller ones, producing hierarchies among African languages themselves. A framework developed in one national setting should therefore not be transferred without examining how colonialism, state formation, and language policy shape African multilingualism.

For this reason, the article treats Banks as providing a broad theory of educational transformation but not a complete explanation of linguistic marginalization; his framework is supplemented by May's work on language rights and minoritization, which more directly addresses the political processes through which languages acquire or lose institutional legitimacy.

2.2. Linguistic Minoritization and Language Rights in Postcolonial African Contexts

May's work is particularly relevant because it distinguishes numerical minority status from the political process of minoritization, through which languages are institutionally subordinated regardless of speaker numbers (May 2012): a language becomes vulnerable when denied intergenerational institutional support—schooling, media, technology—while a dominant language occupies those functions.

This distinction is essential for understanding endangered African languages, whose vulnerability may result from overlapping processes: colonial-language dominance, standardization of selected national languages, urbanization, displacement, and absent digital infrastructure. Endangerment is therefore not a natural decline in speaker numbers but is often produced by institutional arrangements that make transmission and public use increasingly difficult.

A language-rights perspective expands preservation beyond storing recordings: it involves communities' ability to access their heritage, participate in decisions about its use, and create educational opportunities for future speakers—as well as the right not to have restricted materials exposed or incorporated into AI systems without authorization.

This perspective also challenges celebratory approaches to multilingualism: the presence of many languages within a country can coexist with sharply unequal access to educational resources, technological support, and public recognition, so a multicultural-education approach must examine the hierarchy among languages rather than simply count them.

Combining Banks and May connects educational inclusion with structural linguistic inequality: Banks explains why education must change to recognize marginalized knowledge, and May explains why marginalized languages need institutional support if recognition is to become substantive. Together they establish that AI-supported preservation should be evaluated as an educational and political intervention, not only a technical one.

2.3. Culturally Sustaining and Intergenerational Learning

Language revitalization requires more than institutional recognition. Documentation alone cannot reverse language shift: grammars, dictionaries, recordings, and archives are valuable, but languages are sustained through intergenerational transmission and continuing use within families, schools, community institutions, and contemporary communication (Fishman 1991; Grenoble & Whaley 2006; Hinton et al. 2018).

Culturally sustaining pedagogy provides a useful orientation for this task: rather than treating the languages and cultural practices of marginalized learners as transitional resources to be replaced by dominant norms, it supports their continuing development and adaptation (Paris 2012; Paris & Alim 2017), so learners encounter an endangered language not only as an artefact of the past but as a medium for expressing contemporary experience.

This has direct consequences for AI-supported learning resources: an archive of historical recordings alone offers limited support for active use, a dictionary without cultural context fails to convey how words function socially, and an unvalidated chatbot may introduce false forms and undermine speaker

authority. Educational value depends on the relationship between the digital resource and lived learning practices.

Young people occupy an especially important position in this process. Often described only as recipients of endangered-language knowledge, they may also act as translators between generations, digital-content creators, and community researchers, opening new domains of language use through audio storytelling and messaging platforms. Intergenerational learning is correspondingly reciprocal: elders contribute linguistic and ecological knowledge, while younger participants contribute digital skills and diaspora connections. Community-governed AI projects can create settings in which this expertise interacts, an exchange central to multicultural education because it redistributes educational authority and recognizes communities as producers rather than merely sources of knowledge.

2.4. Community Governance, Data Sovereignty, and Decolonial AI

AI introduces governance questions that extend beyond conventional multicultural-education theory, because endangered-language data are inseparable from persons, relationships, and collective identities: a voice recording may contain knowledge belonging to a family, clan, or community, a manuscript may be individually owned yet carry collective spiritual significance, and a place name may reveal restricted ecological or ceremonial knowledge.

Conventional research consent may therefore be insufficient: consent to record does not imply consent to publish, train commercial AI, or generate synthetic voices, so preservation

requires differentiated, continuing governance across documentation, education, model training, and commercial reuse.

Research on data colonialism and decolonial AI warns that digital infrastructures can convert cultural knowledge into data controlled by institutions with greater technical, legal, and financial power (Birhane 2020; Couldry & Mejias 2019; Mohamed et al. 2020); for endangered-language communities the danger is not only extraction but that external systems define the categories of representation, determine what counts as an error, and retain control over the resulting models and platforms.

Community data sovereignty provides an alternative orientation, emphasizing meaningful community authority over the collection, access, reuse, and withdrawal of data connected to cultural heritage (Kukutai & Taylor 2016); the CARE Principles similarly emphasize collective benefit, authority to control, and ethics (Carroll et al. 2020). Such frameworks cannot be transferred mechanically across all African contexts, but they offer principles adaptable through community-specific processes.

Community governance should not, however, treat communities as internally homogeneous: decisions may involve differences among elders and youth, fluent and partial speakers, or supporters of competing orthographies, so responsible governance requires procedures for representation and disagreement, ensuring that invoking “the community” does not allow a small group to claim permanent authority.

Human-in-the-loop validation belongs to this governance structure: beyond improving accuracy, it recognizes that linguistic correctness cannot always be separated from social meaning, a transcription may be phonetically plausible but contextually inappropriate, so human review protects both

technical quality and epistemic authority.

Community governance also requires benefit sharing and capacity development. Communities should not provide recordings and knowledge while external researchers retain the equipment, datasets, technical skills, publications, and institutional recognition; benefits may include locally accessible archives, paid community-researcher roles, training, teacher resources, equipment, shared intellectual credit, and long-term technical support, shifting preservation from extractive research toward accountable partnership.

2.5. An Integrated Framework for Analyzing Community-Governed AI Preservation

The preceding discussion yields four dimensions through which AI-supported preservation can be examined as a multicultural education project. These dimensions are analytically distinct but closely interconnected.

2.5.1. *Linguistic recognition and representation*

This dimension concerns whether marginalized languages, scripts, histories, voices, and knowledge systems become visible in educational and digital environments—not by adding isolated cultural examples, but with attention to how resources are classified, whose categories organize the archive or curriculum, and whether the language is presented as a living system of knowledge. Guiding questions concern representational accuracy, acknowledgement of dialectal, generational, and gendered variation, and whether the resource challenges or reproduces the dominance of colonial and nationally privileged languages.

2.5.2. Linguistic rights and equitable access

This dimension concerns access to opportunities, resources, and institutional support. A digitally documented language may remain inaccessible where materials require expensive devices, continuous connectivity, or literacy in a dominant language, so equity requires attention to affordability, offline and disability access, teacher availability, maintainability by schools and local organizations, and public funding for languages without commercial markets.

2.5.3. Community agency and knowledge authority

This dimension concerns participation in decision-making and control over linguistic heritage: whether community members determine project priorities, approve data use, validate outputs, establish access rules, and share in benefits, participation throughout mapping, data collection, model development, educational adaptation, deployment, and revision, not consultation after a system has been designed, including ownership and control of recordings, datasets, models, and platforms, and the ability to correct, restrict, or withdraw materials.

2.5.4. Intergenerational and culturally sustaining learning

This dimension concerns whether the intervention contributes to continuing language use, identity formation, and transmission, enabling learners to engage with both inherited knowledge and contemporary life. The objective is not to freeze the language in a supposedly authentic past, but to support community authority over continuity and change: interaction among generations, use in contemporary communicative domains, and opportunities for

young people to create and circulate new content rather than only consume existing materials.

These dimensions provide the analytical framework for the remainder of the article: the typology classifies AI applications by preservation function, the scenarios examine how the dimensions take different forms across contexts, the roadmap translates them into implementation activities and indicators, and the risk analysis considers the institutional conditions required to sustain them. The framework does not assume that every intervention must contribute equally to all four dimensions; the central requirement is that technological choices be assessed against the community's preservation ecology and educational priorities rather than technical performance alone.

3. A Function-Based Typology of AI Applications for Endangered African Linguistic Heritage

AI-supported preservation of endangered African languages and scripts requires a differentiated understanding of technological use, since “AI for African languages” encompasses very different technologies and governance requirements: a critically endangered oral language may need emergency documentation; a revitalization movement, pronunciation archives and mobile learning resources; a manuscript tradition, handwriting recognition and differentiated access; a script community, fonts, keyboards, and publishing support. The appropriate intervention therefore depends on the preservation ecology, available data, and institutional capacity.

The typology developed in this section is an authorial synthesis

of scholarship on language documentation, low-resource NLP, digital heritage, and community data governance rather than a derivation from any single existing classification. It was constructed by comparing the principal functions that AI-supported tools may perform and the distinctive educational and governance questions associated with each.

Four primary functional categories are distinguished: documentation and capture, through which vulnerable speech, texts, manuscripts, and writing practices are recorded and converted into processable formats; interpretation and linguistic mediation, through which language materials are organized, connected, contextualized, and made accessible across semantic, orthographic, and cultural boundaries; education and everyday use, through which validated language resources are transformed into learning materials and tools that support continuing use; and coordination, infrastructure, and governance, through which institutions and communities map resources, identify gaps, establish access conditions, and coordinate long-term investment.

These categories are analytical rather than mutually exclusive: a recording may simultaneously support oral documentation, pronunciation learning, dictionary development, and community archiving. The typology's purpose is to clarify each tool's principal preservation function, expected educational contribution, and governance requirements rather than to assign every tool a single permanent category.

Table 1 shows that AI-supported preservation extends well beyond machine translation and large language models, and that educational significance does not correspond to technical complexity: a reliable pronunciation archive, locally usable keyboard, or searchable set of validated stories may contribute

more directly to learning and intergenerational transmission than an experimental general-purpose model.

⟨Table 1⟩ Function-Based Typology of AI Applications for Endangered African Linguistic Heritage

Primary function	AI application	Preservation problem addressed	Suitable context	Principal output	Multicultural-education contribution	Principal governance requirement
Documentation and capture	AI-assisted oral documentation	Loss of elder speech, oral histories, and culturally embedded knowledge	Critically endangered and primarily oral languages	Validated recordings, draft transcripts, contextual metadata, thematic indexes	Representation of marginalized voices and intergenerational access to oral knowledge	Informed and continuing consent, speaker authority, recording and access restrictions
Documentation and capture	Speech indexing and pronunciation archives	Limited access to fluent pronunciation and difficulties searching oral collections	Oral and partially written languages; dispersed learners	Audio-linked words, phrases, narratives, and pronunciation collections	Listening access, pronunciation learning, recognition of dialectal and generational variation	Voice rights, restrictions on synthesis and imitation, speaker attribution
Documentation and capture	OCR and handwriting recognition	Printed and handwritten materials remain inaccessible, non-searchable, or physically vulnerable	African scripts, Ajami traditions, historical documents, community archives	Searchable text, corrected transcription, digital editions, training datasets	Recovery of written knowledge and expanded access to script and manuscript literacy	Custodial authority, approved training use, differentiated access, expert correction
Interpretation and linguistic mediation	Living lexicons	Conventional wordlists detach words from social and cultural meanings	Revitalization programmes, community education, ecological and cultural knowledge projects	Audio dictionaries, examples, images, semantic domains, cultural annotations	Recognition of community knowledge systems and culturally contextualized vocabulary learning	Community validation, protection of sensitive meanings, rules for internal and public entries
Interpretation and linguistic mediation	Transliteration and multiscript tools	Historical, religious, official, and contemporary scripts remain disconnected	Languages written in multiple scripts or orthographies	Parallel-script texts, search links, instructional materials, conversion support	Access across generations and literacy traditions without erasing script identities	Avoidance of script hierarchy, preservation of original forms, community-led orthographic decisions
Interpretation and linguistic mediation	Cultural knowledge graphs and community archives	Relationships among language, place, history, ecology, performance, and identity remain fragmented	Heritage, documentation, education, and archival projects	Linked collections of words, places, narratives, songs, images, and manuscripts	Critical knowledge construction and recognition of community categories and relationships	Community annotation, culturally appropriate metadata, tiered access, correction and withdrawal procedures
Education and everyday use	Generative AI for draft educational materials	Shortage of classroom-ready and community-relevant learning resources	Schools, community classes, adult education, youth and diaspora learning	Draft readers, exercises, posters, lesson plans, audio activities, teacher guides	Culturally sustaining learning, curriculum representation, contemporary language use	Use of validated source materials, speaker and teacher review, disclosure of AI-generated content
Education and everyday use	Constrained learning assistants	Limited access to teachers, fluent speakers, and regular practice opportunities	Young, urban, adult, and diaspora learners	Quizzes, guided dialogue, audio practice, flashcards, resource navigation	Equitable access to practice, learner participation, and support for continuing language use	Restriction to approved corpora, visible source attribution, error reporting, limits on open generation
Coordination, infrastructure, and governance	Language-technology observatories and resource maps	Fragmented information about language vitality, digital resources, institutional capacity, and technological gaps	National, regional, university, archive, and community planning	Resource inventories, readiness profiles, gap analyses, partnership maps	More equitable allocation of educational and technological resources	Community authorization, protection of vulnerable collections, avoidance of extractive data exposure

Source. Developed by the author through a conceptual synthesis of scholarship on language documentation and revitalization, low-resource language technology, digital heritage, multicultural education, and community data governance, including Himmelmann (1998), Woodbury (2011), Besacier et al. (2014), Galla (2016), Muehlberger et al. (2019), Carroll et al. (2020), and related literature.

3.1. Documentation and Capture

Documentation and capture applications address the initial problem that vulnerable linguistic materials may be unrecorded, deteriorating, or inaccessible for search and educational use, aiming to create durable, contextualized, governable records that support later interpretation and community use.

Three application families illustrate this category, detailed in Table 1. AI-assisted oral documentation—sound cleaning, diarization, segmentation, preliminary transcription—can reduce the preparatory labour of documentary linguistics (Himmelfmann 1998; Woodbury 2011); because most endangered African languages lack the accurately transcribed speech datasets required for reliable automatic processing (Besacier et al. 2014), outputs must be treated as provisional materials for review by speakers and linguists. Speech indexing and pronunciation archives offer greater immediate value than general speech recognition for severely under-resourced languages, preserving tonal and dialectal distinctions—voice data requiring consent protections beyond the original recording (Section 2.4). OCR and handwriting recognition extend documentation to African scripts and community-held manuscripts (Muehlberger et al. 2019); output requires correction by readers familiar with the specific tradition, and digitization does not imply unrestricted publication. Across all three families, the contribution is recognition, excluded voices and literacies become available for community teaching, measured by validation and educational use rather than volume alone.

3.2. Interpretation and Linguistic Mediation

Documentation produces records, but records do not automatically become meaningful or educationally usable. Interpretation and linguistic mediation applications organize materials, connect them across forms, and preserve contextual relationships that may be lost through decontextualized data processing.

Living lexicons connect words to pronunciation, recordings, and cultural explanation, presenting language as a system of knowledge rather than interchangeable labels; AI may propose entries and semantic groupings, but machine-generated definitions are prompts for speaker review, and governance should allow entries to remain restricted rather than assuming all lexical knowledge belongs in an open database. Transliteration and multiscript mediation connect Latin, Ajami, and indigenous-script traditions, reducing exclusion of learners literate in a different tradition, but transliteration is not a neutral conversion: scripts carry historical and political meanings, so responsible systems preserve the original representation and avoid installing one orthography as default. Cultural knowledge graphs preserve the relationships linking words to places, persons, and institutions; their categories should emerge from community knowledge, and access rules and withdrawal conditions should travel with the digital object when materials are reused (Table 1).

3.3. Education and Everyday Use

The preservation value of recorded and digitized resources

remains limited if they do not support learning and use. Education and everyday-use applications transform validated materials into forms usable in classrooms, community programmes, homes, and diaspora learning; they are closest to the paper's multicultural-education objectives but also carry significant risks of linguistic error and cultural distortion.

Generative AI can convert validated recordings and wordlists into draft readers and teacher guides, but general-purpose models trained largely on dominant languages may invent words or mistranslate cultural concepts, risks gravest where few fluent speakers can identify errors, so generation should operate only on a bounded, community-authorized corpus, with every resource reviewed before use. Constrained learning assistants that retrieve and reorganize only verified material are correspondingly safer than open-ended chatbots, which may present invented forms with unwarranted confidence; such assistants should identify their sources and never present themselves as substitutes for speakers or teachers. Because dependence on continuous connectivity reproduces exclusion, offline functionality and community access points are core requirements, and, consistent with Section 2.3, platforms should enable young people to contribute content subject to community review (Table 1).

3.4. Coordination, Infrastructure, and Governance

Individual tools are often developed without reliable information about existing materials, institutional capacity, or community priorities. Coordination applications address the fragmentation that occurs when recordings, dictionaries,

manuscripts, educational resources, fonts, and technical projects are distributed across different organizations.

A language-technology observatory documents language vitality, resources, and institutional capacity, directing investment toward community priorities; its contribution is distributive, revealing which languages remain excluded from support. Mapping also creates risk, publishing the location of vulnerable collections may facilitate unauthorized extraction, so observatories require differentiated visibility and explicit access rules (Table 1).

3.5. Connecting Technical Applications to Multicultural Education Outcomes

The four functional categories demonstrate that technical outputs and educational outcomes are related but not identical: a recording's educational value depends on whether it is validated, accessible, and used for learning; a keyboard's on whether learners and teachers can use it across common devices; an OCR model's on whether it recovers marginalized knowledge without overriding manuscript authority.

Applications should therefore be evaluated through the four question sets developed in Section 2.5: whether the tool accurately represents marginalized voices, scripts, and knowledge systems while preserving variation; who can use its outputs, and under what technical and economic conditions; who selected the application, approved the data, validated the output, and controls future reuse; and whether it creates opportunities to hear, speak, read, write, teach, and circulate the language in meaningful contemporary domains.

These criteria prevent AI preservation from being evaluated only through dataset size or model accuracy: a technically sophisticated but inaccessible or externally controlled tool may contribute little to revitalization, whereas a technically modest resource may have substantial value if it strengthens teaching and intergenerational transmission.

The typology also confirms that there is no universal sequence in which applications must be adopted: emergency documentation may precede educational development in one context, while a community with substantial recorded material may prioritize a keyboard or curriculum resource. The next section therefore moves from functional categories to four illustrative preservation scenarios—Ongota, Yaakunte/Yaaku, Nluu, and Ajami manuscript traditions—with African script technology treated separately as a cross-cutting infrastructural layer.

4. Illustrative Preservation Scenarios and a Cross-Cutting Script-Infrastructure Layer

The function-based typology presented in Section 3 identifies what different AI-supported tools may contribute to endangered-language preservation. This section shifts the analysis from technological functions to preservation ecologies. It considers how combinations of documentation, linguistic mediation, educational-resource development, and governance may be adapted to different linguistic and cultural contexts.

The four scenarios examined here—Ongota, Yaakunte/Yaaku, Nluu, and African Ajami manuscript traditions—are not complete

empirical case studies, nor do they represent the full linguistic diversity of Africa. They are illustrative scenarios selected for contrasting preservation conditions: extreme urgency with very few fluent speakers; community-led revitalization connected to ecological memory; reparative recovery involving elder authority; and manuscript preservation involving marginalized literacy traditions.

Each scenario is examined through a common sequence: the principal preservation and educational problem; the realistic role AI-supported tools may perform; the multicultural-education outcomes supported; the governance conditions required; and, as summarized in Table 2, feasible indicators. All five discussions presuppose the shared governance baseline developed in Section 2.4 and Phase 2 of the roadmap (Section 5)—layered, use-specific consent and human validation—not restated within each scenario.

The section then considers African script technologies as a cross-cutting infrastructural layer: fonts, keyboards, OCR, handwriting recognition, transliteration, spelling support, and mobile publishing are enabling conditions that may serve multiple language and script communities rather than constituting a single equivalent case (see Section 1.1).

4.1. Ongota: Emergency Oral Documentation and Rapid Community Return

Ongota, also known as Birale, is spoken in southwestern Ethiopia and has long been described as severely endangered, with active use contracted to a very small number of speakers shifting toward neighbouring Ts'amakko (Savà & Tosco 2000).

The central preservation problem is urgency: with few fluent speakers, projects have limited time to record variation in pronunciation and oral history, and the immediate objective should be high-quality, community-governed records rather than a general-purpose language model.

A realistic AI-supported workflow could assist with improving audibility while retaining original files, segmenting long recordings, and preparing draft metadata for human review, functions that reduce organizational labour without producing authoritative transcripts, since limited data and speaker variation make automatic processing unreliable, so output remains provisional material requiring review by speakers and linguists.

This article proposes rapid community return as the governing principle for such a scenario—an author-developed operational principle rather than an established concept in the literature: rather than remaining with external researchers until a grammar or archive deposit is complete, approved materials should be returned in usable forms—speaker—labelled audio collections, pronunciation sets, family copies, offline audio lessons—during the life of the project.

The multicultural-education significance lies in recognition: Ongota voices should survive not only as research data but as the community's own educational inheritance. The scenario also concerns knowledge authority, speakers should determine which genres are recorded and who may listen, with some materials open to general educational use and others limited to families or designated researchers.

The proposed intervention is deliberately modest in technical ambition but potentially substantial in educational

value: the question is not whether AI can reproduce fluent Ongota, but whether AI-assisted processing can help preserve irreplaceable voices and return them responsibly to the people whose heritage they represent.

4.2. Yaakunte/Yaaku: Community Revitalization, Ecological Memory, and Youth Agency

Unlike Ongota, which begins from the urgency of recording a small speaker base, the Yaakunte/Yaaku scenario begins from active community interest in revitalization. The Yaaku community of central Kenya has experienced substantial shift toward Maa, and Yaakunte has been classified as severely endangered or moribund (Dimmendaal & Voeltz 2007); reported community-led initiatives nevertheless suggest that interrupted transmission does not eliminate the possibility of renewed teaching (Carrier 2011). The preservation ecology is closely connected to the Mukogodo Forest and to cultural knowledge of landscape and oral history, so revitalization should connect language learning to the relationships through which words acquire meaning, not isolated vocabulary lessons.

An AI-supported pilot could combine a speaker-validated living lexicon linking words to ecological categories; an audio-linked curriculum organized around family, forest, and seasons; youth-produced digital storytelling documenting interviews; and community-governed ecological knowledge mapping. The contribution is especially visible in youth agency: young people act as interviewers and curriculum co-designers while elders contribute oral history and ecological interpretation, the reciprocal exchange described in

Section 2.3, giving youth a role in determining how their heritage language enters contemporary life.

Not every element of ecological knowledge should be digitized: some may require restriction, and the design should permit entries classified as public, community-only, or preserved but not displayed. The project should also avoid presenting one speaker as the uncontested form of Yaakunte: where speakers differ, the resource should represent variation rather than prematurely impose a standard, treating orthographic decisions as choices requiring community deliberation.

The purpose of the scenario is thus not to build a technology that symbolically represents a nearly lost language, but to support a social learning environment in which language, ecological memory, youth participation, and community identity reinforce one another.

4.3. Nluu: Reparative Language Recovery, Pronunciation, and Elder Authority

The Nluu case illustrates a different preservation ecology. Nluu, associated with ǀKhomani San communities in South Africa, experienced profound disruption through colonial dispossession and the dominance of Afrikaans; its decline connects to histories in which San languages and knowledge were devalued and excluded from formal institutions (Mlaba et al. 2025; Shah & Brenzinger 2017). This article therefore proposes a reparative orientation—the analytical position adopted here rather than a claim that one project can repair the wider historical harms experienced by San communities—requiring that preservation recognize elder speakers as

intellectual authorities, restore community access, and avoid reproducing extractive relationships.

The immediate technical priority should not be unrestricted automatic speech recognition: given limited annotated speech and phonetic complexity, a general-purpose system may be neither feasible nor useful under current conditions. A more appropriate pilot would use bounded tools—a pronunciation archive organized by speaker, phrase-level retrieval, and a constrained learning assistant based only on verified resources—since Nluu includes distinctions difficult to represent through text alone, so the archive should preserve actual speaker recordings rather than synthetic speech.

AI may assist with clipping, indexing, and searching recordings, and may support pronunciation comparison, framed as learning support rather than automatic judgment of authenticity, since an algorithm trained on very few speakers could wrongly treat one pronunciation as the only legitimate norm. Educational resources should be developed with elder speakers and community organizations, and should not represent Nluu solely as a language of the past.

A constrained learning tool could provide speaker-recorded vocabulary, guided repetition, and cultural explanation; it should not generate unrestricted Nluu sentences, and where content is incomplete it should say no validated answer is available rather than invent a form. The contribution involves recognition and the repair of institutional exclusion: Nluu resources can challenge narratives that treat San languages as extinct and support community members reconnecting with a language from which their families were separated.

Elder authority is central: recordings should retain speaker

names where consent permits rather than being anonymized, and decisions about synthetic voice and posthumous use require especially careful consent. The scenario demonstrates that AI-supported preservation should not seek technological scale at the expense of legitimacy: a small but carefully governed pronunciation system may contribute more to recovery than an inaccurate model claiming broad competence.

4.4. African Ajami Traditions: Manuscript Recovery, Multiliteracy, and Governed Access

Ajami refers broadly to African languages written using adapted forms of Arabic script, serving religious learning, poetry, and historical writing across many regions (Mumin & Versteegh 2014; Ngom 2016), challenging accounts that equate African literacy with Latin-script schooling. The preservation problem is both material and epistemic: many manuscripts are physically vulnerable and poorly catalogued, and digitization can expand access to African intellectual history while removing materials from the interpretive authority of their custodians. A responsible model must begin with custodial partnership rather than scanning technology; this article proposes a five-part workflow.

4.4.1. *Custodial consultation and collection-level assessment*

Before digitization, manuscript holders and relevant communities should determine custodial status, whether digitization is appropriate, and what benefits and copies will be returned.

4.4.2. High-quality digitization and contextual metadata

Manuscripts should be scanned in ways that preserve page structure and marginalia, with metadata documenting language, script variety, and provenance using local categories rather than external archival terminology.

4.4.3. AI-assisted recognition and human transcription

OCR and handwriting-recognition tools may assist in identifying characters and proposing transcriptions, but Ajami systems vary across languages and writers, so output must be corrected by readers familiar with the relevant orthographic practice (Muehlberger et al. 2019; Yousuf et al. 2025).

4.4.4. Layered digital editions and multiscript access

One promising model is a layered digital edition presenting the original image, a diplomatic transcription, a transliteration, and audio readings, optional rather than hierarchical layers, so the Latin version neither replaces nor outranks the Ajami original (Fall et al. 2020).

4.4.5. Educational return and differentiated access

Digitized materials should be returned to custodians in forms supporting their priorities—local copies, printed facsimiles, teacher resources—since public access should not be digitization's automatic outcome: layered access may distinguish public catalogue information, community-member access, and preserved materials not displayed.

The significance of the Ajami scenario lies in multiliteracy: manuscripts reveal forms of African literacy largely absent from formal curricula, and transliteration can connect learners

in Latin-script schools with family literacy traditions without requiring either to displace the other. Governance should include recognition and compensation of manuscript owners and transcribers; external institutions should not retain exclusive control over images or AI models derived from them.

The Ajami scenario demonstrates that AI may improve manuscript recognition and accessibility, but preservation must remain accountable to the people and institutions that have maintained these intellectual traditions.

4.5. A Cross-Cutting Layer: African Script Technology Infrastructure

The preceding four scenarios focus on specific community or heritage ecologies. African script technology requires different treatment: scripts such as N’Ko, Adlam, Bamum, Tifinagh, Vai, and Ethiopic are associated with different languages and levels of digital support, and should not be treated as one community. They are brought together for a narrower comparative purpose: each illustrates the gap between formal recognition and practical usability, since a script may be encoded in Unicode yet remain difficult to use because of limited keyboards or poor rendering (Unicode Consortium 2025).

The scripts discussed are illustrative rather than exhaustive, selected for their historical diversity and varying relationships to revitalization and identity.

A cross-cutting script-technology package should be understood as modular infrastructure rather than a standardized product—fonts, keyboards, OCR datasets, and

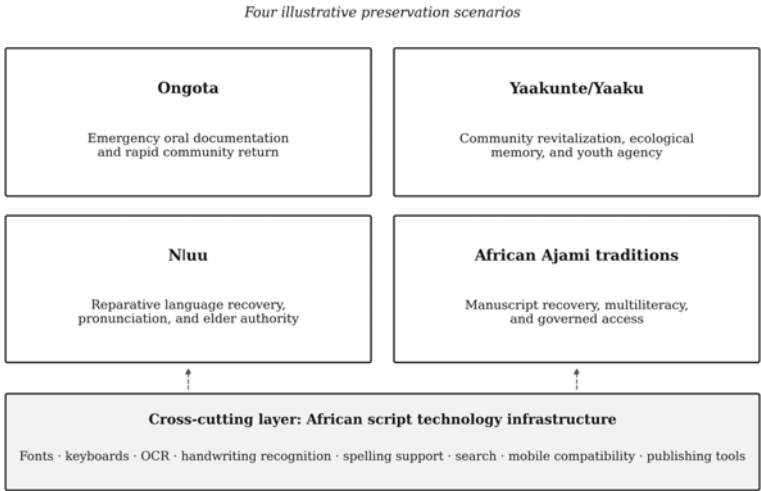
locally maintainable repositories—adapted to each script and community, since orthographic rules and governance cannot be inferred from Unicode encoding alone. Script infrastructure contributes to recognition and identity formation: learners cannot meaningfully exercise a right to use a script they cannot type or search on common devices, and absent tools silently push users toward dominant scripts.

A responsible development process would include consultation with script users, testing of open tools across devices, and long-term maintenance arrangements. AI may assist with OCR and spelling suggestions, but AI-generated normalization should not erase historical variation: users should be able to view original forms and reject machine correction.

The aim is not simply to place African scripts inside digital systems. It is to make them practically usable for writing, learning, publishing, searching, communication, and cultural production.

4.6. Comparative Synthesis

Figure 2 visualizes the analytical structure of the comparison: the four community- and heritage-centred scenarios appear as parallel panels, connected from below by the script-technology infrastructure layer that may serve several of them simultaneously.



Note. The four boxes represent community- or heritage-centred scenarios; the cross-cutting layer supports multiple script communities. Source: Developed by the author.

〈Figure 2〉 Four Illustrative Preservation Scenarios and a Cross-Cutting Infrastructure Layer

Table 2 consolidates the comparison developed in this section: for each of the four scenarios and the cross-cutting layer, it summarizes the preservation ecology, priority problem, realistic AI-supported functions, multicultural education outcomes, governance conditions, and feasible indicators. The final row is included for comparative convenience only: as Sections 1.1 and 4.5 explain, script infrastructure operates at a different analytical level from the four scenarios.

〈Table 2〉 Comparative Framework for the Four Illustrative Scenarios and the Cross-Cutting Infrastructure Layer

Analytical unit	Primary preservation ecology	Priority problem	Realistic AI-supported functions	Principal multicultural-education outcomes	Core governance conditions	Feasible indicators
Ongota	Severely endangered oral language	Loss of elder speech and oral knowledge	Audio processing, segmentation, alignment, indexing, provisional transcription, metadata assistance	Recognition of marginalized voices; community access; introductory intergenerational learning	Layered consent, speaker authority, rapid community return, restrictions on AI training and voice use	Validated and returned recordings; community researchers trained; educational resources developed
Yaakunte/Yaaku	Community revitalization connected to ecological and cultural knowledge	Interrupted transmission and limited learning materials	Living lexicon, audio curriculum, digital storytelling, ecological knowledge mapping	Youth agency; culturally sustaining learning; intergenerational exchange; ecological literacy	Community validation, representation of variation, differentiated access to ecological knowledge	Youth-created content; learner participation; intergenerational activities; use of validated resources
Nluu	Reparative language recovery with very few elder speakers	Loss of pronunciation, learning opportunities, and institutional recognition	Pronunciation archive, audio-linked orthography, phrase retrieval, constrained learning tools	Identity affirmation; pronunciation learning; recognition of elder knowledge; recovery from institutional exclusion	Elder authority, attribution, consent for voice reuse, careful representation of variation	Use of pronunciation resources; younger learner participation; offline access; speaker-approved materials
Ajami traditions	Dispersed and vulnerable manuscript literacies	Material deterioration, weak discoverability, and exclusion from formal literacy narratives	Digitization support, layout analysis, handwriting recognition, transcription assistance, transliteration, layered digital editions	Multiliteracy; recovery of African intellectual history; connection of formal and religious education	Custodial partnership, specialist correction, differentiated access, return of digital copies and benefits	Corrected transcriptions; local specialists trained; educational editions; functioning access controls
African script infrastructure	Cross-cutting digital-literacy infrastructure	Gap between encoding and everyday usability	Fonts, keyboards, OCR, handwriting recognition, spelling support, conversion, publishing tools	Script literacy; participation in digital communication; cultural and linguistic recognition	Script-specific consultation, open and maintainable tools, respect for orthographic plurality	Device compatibility; classroom and community use; publications produced; local maintenance capacity

Source. Developed by the author through comparative synthesis of the preservation ecologies and literature discussed in Sections 2–4.

The comparison demonstrates that the appropriate role of AI changes with the preservation ecology, urgent documentation for Ongota, ecological memory for Yaakunte/Yaaku, reparative recovery for Nluu, custodially governed access for Ajami traditions, and practical infrastructure for script communities. The scenarios demonstrate a method of contextual adaptation rather than transferable models; the seven-phase roadmap in the next section translates this logic into a structured but adaptable process.

5. A Seven-Phase Roadmap for Community-Governed AI Preservation

The preceding sections distinguished the functions AI-supported tools may perform and the ecologies within which they may be used. The next challenge is implementation: without a structured process, recordings may be deposited without local access and tools created without governance arrangements. This section proposes a seven-phase roadmap, an analytical model developed by the author through the synthesis of language documentation, multicultural education, digital heritage, and community data governance, together with the requirements identified in the four scenarios.

The roadmap is not a universally validated model but a design framework identifying functions responsible projects are likely to require: phases may overlap or be revisited, and governance and sustainability operate across the whole cycle. Three principles organize it: preservation should begin with a defined problem rather than an available tool; AI should assist rather than displace human expertise; and success should be evaluated through both technical and multicultural-education outcomes.

Table 3 separates technical outputs from multicultural-education outcomes. A searchable corpus is not itself intergenerational transmission, a keyboard is not automatically digital participation, and a digitized manuscript is not necessarily educational inclusion: such outputs acquire multicultural-education significance when they are validated, accessible, used in learning, and governed by the communities

whose knowledge they represent.

〈Table 3〉 Seven-Phase Roadmap for Community-Governed AI Preservation

Phase	Core activities	Principal actors	Technical or resource output	Multicultural-education outcome	Illustrative indicators
1. Mapping, scoping, and consultation	Identify speakers, learners, scripts, materials, language domains, educational needs, digital gaps, institutions, and internal diversity	Speakers, elders, youth, teachers, community organizations, researchers, local institutions	Community-reviewed language and preservation profile	Recognition of community priorities and diverse learner needs; participation in defining the problem	Diversity of participants; documented priorities; inventory of resources and gaps; evidence of community approval
2. Governance and ethical protocol design	Define consent, authority, access, ownership, attribution, AI-use permissions, benefit sharing, correction, withdrawal, and dispute procedures	Community governance body, speakers, custodians, cultural authorities, institutions, legal or ethical advisers	Community-approved governance protocol and decision structure	Linguistic rights, democratic participation, knowledge authority, and protection from extraction	Signed or recorded protocols; differentiated access levels; community veto and withdrawal mechanisms; compensation arrangements
3. Documentation and digitization	Record speech, scan manuscripts, document scripts, collect existing scripts, create metadata, and preserve originals	Speakers, elders, scribes, custodians, archivists, linguists, community researchers	Governed and contextualized documentation collection	Representation of marginalized voices, literacy traditions, and knowledge systems	Amount and diversity of material collected; percentage with consent and metadata; community copies created; preservation quality
4. AI-assisted processing and human validation	Apply segmentation, transcription assistance, OCR, handwriting recognition, transliteration, indexing, and metadata tools; review all outputs	Technologists, speakers, teachers, linguists, scribes, manuscript specialists, community reviewers	Validated corpus, transcription set, searchable collection, or script resource	Epistemic accuracy; protection from technological misrepresentation; recognition of community expertise	Error and correction rates; proportion human-validated; representation of linguistic variation; traceability of revisions
5. Resource development and educational transformation	Convert validated materials into lessons, dictionaries, readers, audio resources, script tools, exhibits, and constrained learning applications	Teachers, speakers, curriculum developers, youth, designers, cultural workers, technologists	Community-reviewed educational and cultural resource package	Culturally sustaining pedagogy, curriculum representation, identity affirmation, and intergenerational learning	Teacher and speaker approval; range of learner levels; contemporary and heritage content; accessibility and usability testing
6. Community deployment and institutional integration	Use resources in schools, homes, community centres, archives, religious institutions, cultural events, media, and digital platforms	Community organizations, schools, universities, archives, ministries, religious institutions, technology partners	Active use across educational, community, and institutional domains	Equitable participation, youth agency, everyday language use, and institutional recognition	Number and diversity of users; frequency of use; classroom and community adoption; learner participation; youth-created content
7. Sustainability, maintenance, and evaluation	Maintain tools, update content, train personnel, secure financing, preserve files, review governance, and assess outcomes	Community bodies, universities, schools, libraries, ministries, funders, technical partners	Maintained and renewable preservation ecosystem	Long-term access, institutional equity, community capacity, and cultural continuity	Recurring funding; local maintenance capacity; resource updates; continued use; governance reviews; educational and transmission outcomes

Source. Developed by the author through a conceptual synthesis of language documentation, revitalization, multicultural education, digital heritage, and community data-governance scholarship and the comparative preservation scenarios examined in Sections 3 and 4.

5.1. The Seven Phases in Overview

Table 3 presents the seven phases with their activities, actors, outputs, and outcomes; the prose account is therefore brief. The cycle begins with mapping and community

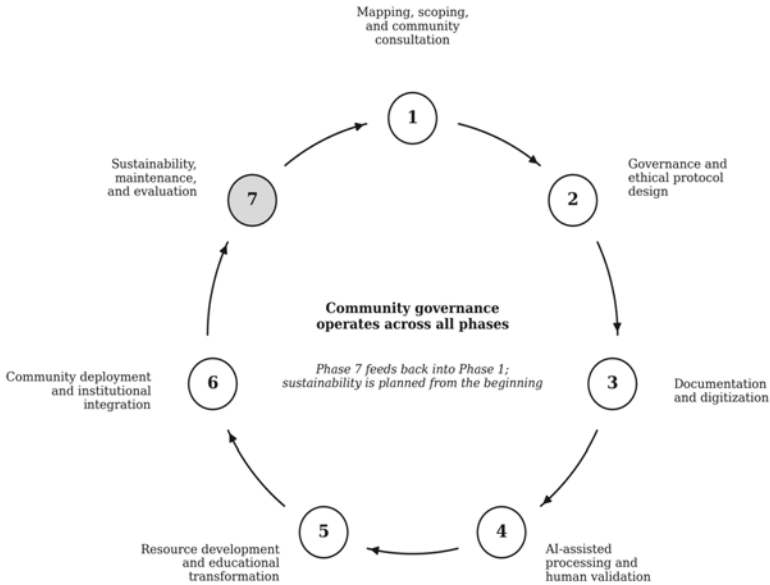
consultation (Phase 1), whose output is a community-reviewed preservation profile rather than an externally written assessment, consultation designed to reveal internal diversity, since elders, youth, and diaspora learners rarely share a single priority. Governance and ethical protocol design (Phase 2) then establishes, before data are collected, who may approve the project, who controls derivatives, which materials are restricted, whether data may train AI systems or be voice-synthesized, and how disagreement is handled.

Documentation and digitization (Phase 3) create purposefully broad primary materials with metadata recorded in local categories and originals preserved for correction. AI-assisted processing and human validation (Phase 4) bound the technology by the limits of the data and keep every output's status visible. Resource development (Phase 5) converts validated materials into reviewed teaching resources; community deployment (Phase 6) embeds them in schools and institutions with explicit responsibilities and offline access; and sustainability, maintenance, and evaluation (Phase 7), planned from Phase 1 onward, asks not whether a language has been “saved” but whether access, community authority, and meaningful use have been strengthened.

5.2. The Roadmap as an Iterative and Multicultural Education Process

The circular form of the roadmap in Figure 3 below matters because preservation requires continuous feedback: deployment may reveal errors that return the project to validation, new learners may create documentation needs, technological change

may require revised governance, and evaluation may show that a technically accurate tool has little educational value.



Note. Governance operates across all phases, and sustainability is planned from the beginning. Source: Developed by the author.

〈Figure 3〉 Seven-Phase Circular Roadmap for Community-Governed AI Preservation

The roadmap can also be read through the four dimensions of Section 2: recognition begins in mapping and continues through decisions about materials; rights and access are established through governance and resource design; community agency operates through consultation and withdrawal; and intergenerational learning becomes visible in deployment while depending on all earlier phases.

The roadmap should therefore be adapted rather than

copied: each of the scenarios examined in Section 4 would weight the phases differently, as summarized in Table 2.

What remains constant is not a particular technology but a set of responsibilities: begin with community-defined needs, establish governance before extraction, validate AI output, and plan for long-term maintenance. The following section examines the risks that may undermine these responsibilities.

6. Risks, Governance Architecture, Institutional Partnerships, and Policy Implications

The typology, scenarios, and roadmap identify practical possibilities, but these are accompanied by risks arising from how linguistic data are collected, stored, and institutionalized: the question is not only whether AI tools perform a technical function but how their use redistributes authority and educational opportunity.

This section identifies five major risk domains: extractive data governance, exposure of restricted knowledge, AI error and epistemic distortion, infrastructural unsustainability, and unequal institutional power, an author-developed analytical synthesis derived from critical scholarship on data colonialism, decolonial AI, and digital heritage, together with the recurring governance problems identified in the scenarios and the roadmap.

The five risks are analytically distinct but interconnected—extractive arrangements intensify institutional inequalities, weak access controls expose restricted knowledge, inadequate validation converts technical error into educational

misinformation, and infrastructural dependency undermines community control—so governance must address their relationships rather than treating each in isolation.

〈Table 4〉 Principal Risks and Governance Responses in Community-Governed AI Preservation

Risk domain	Mechanism of harm	Possible multicultural-education consequence	Principal governance response	Illustrative indicators
Extractive data governance	Community recordings, texts, manuscripts, or annotations are controlled or reused by external actors beyond the original purpose	Community knowledge becomes visible without strengthening educational authority, access, or linguistic rights	Layered consent, data-use agreements, community authority over reuse, benefit sharing, withdrawal rights	Explicit AI-training permissions; documented ownership and access rules; community control over secondary use
Exposure of restricted knowledge	Digitization, search, translation, or model training makes sensitive materials available beyond intended audiences	Sacred, familial, ecological, or historical knowledge is decontextualized or taught inappropriately	Culturally differentiated access, community-defined metadata, exclusion from models, custodian approval	Functioning access categories; restricted materials excluded from public search and AI training; periodic access review
AI error and epistemic distortion	Speech, OCR, transliteration, translation, or generation systems produce inaccurate or culturally misleading outputs	Learners encounter invented language, distorted grammar, erased variation, or misleading cultural explanations	Human-in-the-loop validation, uncertainty labels, source traceability, constrained generation, correction procedures	Proportion of outputs reviewed; visible validation status; correction rates; representation of variation
Infrastructural unsustainability	Platforms, tools, datasets, or educational resources become inaccessible, incompatible, or unsupported	Learners and teachers lose access; community trust and educational continuity are weakened	Open formats, maintenance plans, local capacity, recurring financing, exportable data, distributed backups	Named maintenance responsibility; recurrent budget; local administrators trained; successful migration and backup tests
Unequal institutional power	Funders, universities, archives, developers, or governments control project priorities, resources, publication, and infrastructure	Community participation becomes symbolic; external expertise displaces local knowledge authority	Accountable partnership, shared decision-making, compensation, co-authorship, local access, dispute procedures	Community representation in governance; transparent budgets; attribution and payment; locally accessible outputs

Source. Developed by the author through a synthesis of data-governance, decolonial AI, digital-heritage, multicultural-education, and language-revitalization scholarship and the scenarios examined in this article.

The table distinguishes technological risk from educational harm: a system error becomes a multicultural education concern when it shapes what learners are taught or which linguistic forms acquire authority, and data inclusion does not automatically produce educational inclusion, a language may become more visible in an AI system while its speakers remain unable to access the model or determine how their heritage is taught.

6.1. Extractive Data Governance and the Limits of Digital Inclusion

The first risk concerns the extraction of linguistic and cultural materials from the communities that produce and maintain them: endangered-language resources are especially attractive for AI development because they contain rare data unavailable in widely used corpora.

These possibilities reflect the concerns of data colonialism and decolonial AI scholarship: digital systems can convert cultural knowledge into resources controlled through unequal structures (Birhane 2020; Couldry & Mejias 2019; Mohamed et al. 2020), a risk intensified for endangered languages by the scarcity of the material.

Digital inclusion and data sovereignty should therefore be distinguished: inclusion asks whether a language is represented, sovereignty who has authority over that representation. A language may enter a speech model while speakers remain excluded from decisions about ownership and benefit sharing; consent should remain differentiated by use and open to renewal, as specified in Phase 2.

Governance should also address derived resources: a community may retain authority over original recordings yet lose control if an institution claims ownership of derivative transcriptions or models.

The multicultural-education objective is to ensure that digital representation strengthens rather than bypasses community authority: whether communities can access the resources, influence their educational use, correct their representation, and refuse applications inconsistent with their priorities.

6.2. Restricted Knowledge and Culturally Differentiated Access

The second risk concerns materials not intended for unrestricted circulation: ritual speech, sacred songs, healing knowledge, religious manuscripts, ecological information, family histories, and politically sensitive narratives.

Digitization alters the conditions under which such knowledge circulates: a story shared in one setting becomes globally searchable, and generative systems may reproduce restricted content without its access conditions.

The risk cannot be resolved through a simple open/closed division: communities may require several categories, from open access through community-member and elder-only access, to preservation without display.

Access conditions should travel with the object through metadata and platform design. Community-defined labels can communicate obligations concerning attribution and ceremonial status; the Traditional Knowledge Labels developed through Local Contexts offer one useful model, adapted rather than mechanically transferred to African settings (Local Contexts, n.d.).

AI systems should respect access categories technically: restricted material should be excludable from public search, downloadable corpora, model training, automated translation, generative interfaces, and third-party data aggregation.

Culturally differentiated access affirms that recognition does not require unrestricted exposure: respect for cultural difference includes respecting the right of communities to decide that some knowledge should not be publicly represented.

6.3. AI Error and Epistemic Distortion

The third risk concerns inaccurate, incomplete, or culturally misleading AI outputs, to which endangered and low-resource languages are particularly vulnerable because available training data may be limited, uneven, or dominated by a few speakers, genres, or orthographic forms.

Errors may occur across all the applications examined here: misidentified tones or word boundaries, confused characters, imposed orthographies, distorted cultural concepts, invented vocabulary, one speaker presented as the universal norm, or false relationships among people, places, and cultural categories.

The problem is not only that AI systems make mistakes but that machine-generated outputs may acquire institutional authority: a learner may assume an app is correct, or an institution may adopt a tool whose categories override community usage.

Human-in-the-loop validation is therefore an educational and epistemic requirement, not merely an accuracy method, involving the expertise relevant to each resource—speakers, teachers, linguists, scribes, custodians, and ordinary users.

Systems should preserve a visible distinction among source material, unreviewed output, and community-approved content; a constrained assistant should state that it lacks a validated response rather than generate a plausible but unsupported answer.

This risk ultimately concerns knowledge construction: AI systems participate in determining what counts as correct and teachable, and community validation prevents machine-produced approximations from replacing marginalized linguistic knowledge.

6.4. Infrastructural Sustainability and Dependency

The fourth risk concerns the long-term viability of technologies and institutional arrangements: digital tools may become inaccessible when hosting ends or trained staff leave, a structural vulnerability rather than an inevitable failure.

Sustainability has technical, financial, and social dimensions: projects should use transferable formats and budgets extending beyond initial development, since a resource that merely remains online is not sustainable if teachers cannot use it.

Local capacity development reduces dependency, though it should not be romanticized as immediate self-sufficiency: the objective is that community organizations and African institutions can perform routine management, correct content, control access, export resources, and participate meaningfully in major technical decisions.

The multicultural-education consequence of unsustainability is interrupted access: learners, teachers, and communities may lose resources on which new educational practices have begun to depend, so sustainability should be treated as a condition of educational equity rather than only a technical concern.

6.5. Unequal Institutional Power and Accountable Partnership

The fifth risk concerns differences in power among communities, researchers, universities, archives, developers, governments, and funders. AI-supported preservation usually requires partnership, since no single participant possesses all the relevant linguistic, educational, archival, financial, and technical capacities.

Partnership does not itself guarantee equity: external

institutions may control funding, design, storage, publication, intellectual property, and infrastructure, while community participants provide recordings, interpretation, and validation with limited influence over budgets, software, or long-term use.

Accountable partnership therefore requires more than consultation: participation in defining the problem, representation in governance, transparent funding, written agreements on ownership and reuse, payment for linguistic and custodial labour, shared credit where warranted, practical return of materials, and the ability to reject or modify a proposed application (Table 4).

Technical expertise should not automatically outrank linguistic or cultural expertise: a developer may understand model architecture without understanding whether the system's categories are socially meaningful.

African universities are especially important partners because they connect linguistic, archival, and policy expertise, and are more likely to strengthen linguistic sovereignty when partnerships maintain community authority rather than presuming ownership of community data. More broadly, no single institution can fulfil all requirements: developers must design for small datasets and human validation, and governments and funders should condition support on governance and community benefit rather than dataset scale.

6.6. Governance Architecture for Responsible AI Preservation

The five risks point toward a governance architecture whose mechanisms are introduced in Phase 2 of the roadmap-layered consent, community authority, differentiated access, human

validation with traceability, benefit sharing, capacity development, correction and withdrawal, and dispute resolution, operating as ethics as infrastructure: designed into platforms, budgets, and institutional agreements rather than appended as an ethical statement.

This matters especially under urgency: pressure to collect first and resolve governance later is strongest precisely when materials collected may be the only substantial record a community will ever possess.

6.7. Policy Implications for Multicultural Education and Digital Inclusion

The governance analysis has implications not only for AI and cultural-heritage policy but for multicultural education itself, since a framework concentrating exclusively on datasets and platforms will not address the processes through which language status and intergenerational transmission are shaped.

6.7.1. *Curriculum, teacher education, and assessment policy*

Curriculum policy should provide legitimate space for endangered and minoritized languages and local knowledge beyond occasional heritage days—bilingual readers, oral-history projects, digital storytelling—presenting endangered languages as relevant to contemporary life rather than only relics of the past.

Teachers need preparation not only in using digital tools but in their limitations: critical AI literacy should enable teachers to distinguish validated content from machine output

and identify errors. Teacher education should also address linguistic hierarchy, since learners may associate dominant languages with modernity and endangered languages with backwardness, associations AI tools should challenge rather than reproduce.

6.7.2. Educational resources, public financing, and procurement

Public resource programmes should support small-language materials even where potential users are few, since commercial viability is an inappropriate sole criterion; funding should weigh cultural urgency and community commitment.

Assessment policy should move beyond counting digitized items toward learner participation, script literacy, youth-created content, and continued use after initial funding.

6.7.3. Language policy, digital infrastructure, and endangered-language prioritization

National language policies should recognize that contemporary language support includes fonts and keyboards, speech archives, OCR, digital publishing, community archives, teacher resources, and governance standards, formal recognition without usable infrastructure leaves a language symbolically valued but practically excluded.

Publicly funded projects should demonstrate community participation and human validation, and procurement should avoid locking communities into proprietary systems they cannot afford or leave.

African AI and language-technology strategies understandably prioritize widely spoken languages, but this approach may reproduce the marginalization of smaller languages that are

less commercially attractive; dedicated funding streams are needed for severely endangered languages, oral traditions, manuscript collections, and script communities.

6.7.4. *Cross-border cooperation and digital sovereignty*

Languages and manuscript traditions often cross national borders, and regional cooperation can support shared research and OCR development, provided it recognizes local variation rather than replacing community-specific governance with a single standard.

Finally, policy should distinguish between making a language available on a platform and enabling communities to govern the infrastructure representing it: digital sovereignty includes national capacity, but African control over data centres is not sufficient if minoritized communities remain excluded from decisions about their own materials.

7. Conclusion

This article has examined how artificial intelligence may support the preservation and revitalization of endangered African languages and scripts when interventions are embedded within community-governed, educationally meaningful systems. AI can assist with recording, transcription, OCR, and archival access, but these capabilities become significant only when communities retain authority over what is documented and how it is used in teaching.

The article's principal theoretical contribution is to connect AI-supported language preservation with multicultural education,

linguistic minoritization, and digital language justice: Banks explains why endangered languages are an educational concern rather than only cultural heritage, May clarifies how languages remain socially meaningful while being institutionally excluded, and their integration with culturally sustaining pedagogy and community data governance produced the article's four analytical dimensions.

These dimensions shift evaluation beyond computational performance: accuracy matters because errors produce educational harm, but an intervention's value depends on whether it expands community access and strengthens local decision-making, a modest pronunciation archive may have greater educational value than a sophisticated model communities cannot govern.

The article's applied contribution—the function-based typology, the four illustrative scenarios with their cross-cutting script-infrastructure layer, the seven-phase roadmap, and the five-risk governance framework—translates this evaluative shift into design practice.

Together, these proposals support a transition from a project logic, which evaluates success through completed products, to an ecosystem logic, which asks whether those products remain connected to speakers and long-term use. Preservation is not completed when materials are digitized: it requires continuing validation and community decision-making. Governance, in this sense, is a condition of educational usefulness, determining whether AI-supported preservation merely represents diversity or redistributes educational authority.

For multicultural education, this means endangered languages should not enter education only as symbolic evidence of

diversity: curriculum, teacher preparation, and assessment should be designed for minoritized communities rather than adapted as an afterthought. African AI strategies should likewise avoid equating linguistic inclusion with incorporating widely spoken languages into large models: digital language justice includes severely endangered languages and indigenous scripts, supported by governance standards and long-term maintenance financing.

The article has several limitations: it is conceptual and policy-oriented rather than an empirical evaluation, the scenarios are illustrative constructions grounded in documented conditions rather than fieldwork, and questions of custodial rights can only be resolved through community-specific processes.

Future research should develop the framework through community-based studies of implemented projects and comparative policy analysis—including the single-country implementation study identified in Section 1.1—and should examine tensions a conceptual framework cannot resolve in advance, over standardization, custodial authority, and the distribution of benefits, since responsible governance requires mechanisms for representing disagreement as priorities change.

Ultimately, AI may contribute either to linguistic extraction or to linguistic recovery: extraction when endangered-language materials become data governed elsewhere, recovery when it helps communities preserve voices, reconnect generations, strengthen literacy in marginalized scripts, and expand the educational presence of their languages on their own terms. Which future develops will depend less on model capability than on governance, educational integration, and the distribution of authority among the institutions and communities involved.

References

- African Union.(2024). Continental artificial intelligence strategy. African Union.
https://au.int/sites/default/files/documents/44004-doc-EN-_Continental_AI_Strategy_July_2024.pdf(Retrieved on: May 27, 2026)
- Banks, J. A.(2009). Multicultural Education: Dimensions and Paradigms, In J. A. Banks(Ed.), *The Routledge International Companion to Multicultural Education*, Routledge, pp.9-32.
- Besacier, L., Barnard, E., Karpov, A., & Schultz, T.(2014). Automatic speech recognition for under-resourced languages: A survey. *Speech Communication*, 56, pp.85-100. <https://doi.org/10.1016/j.specom.2013.07.008>
- Birhane, A.(2020). Algorithmic colonization of Africa. *SCRIPTed*, 17(2), pp.389-409. <https://doi.org/10.2966/scrip.170220.389>
- Carrier, N.(2011). Reviving Yaaku: Identity and indigeneity in Northern Kenya. *African Studies*, 70(2), pp.246-263. <https://doi.org/10.1080/0020184.2011.594632>
- Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M.(2020). The CARE principles for Indigenous data governance. *Data Science Journal*, 19(1), Article 43. <https://doi.org/10.5334/dsj-2020-043>
- Couldry, N., & Mejias, U. A.(2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.
- Dimmendaal, G. J., & Voeltz, E. F. K.(2007). Africa, In C. Moseley(Ed.), *Encyclopedia of the World's Endangered Languages*, Routledge, pp.579-634.
- Fall, E. M., Nguer, E. M., Sokhna, B. D., Khoule, M., Mangeot, M., & Cisse, M. T.(2020). Digraphie des langues ouest-africaines: Latin2Ajami: Un algorithme de translittération automatique [Preprint]. arXiv. <https://arxiv.org/abs/2005.02827>(Retrieved on: May 27, 2026)
- Fishman, J. A.(1991). *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Multilingual

Matters.

- Galla, C. K.(2016). Indigenous language revitalization, promotion, and education: Function of digital technology. *Computer Assisted Language Learning*, 29(7), pp.1137-1151. <https://doi.org/10.1080/09588221.2016.1166137>
- Grenoble, L. A., & Whaley, L. J.(2006). Saving languages: An introduction to language revitalization. Cambridge University Press.
- Himmelman, N. P.(1998). Documentary and descriptive linguistics. *Linguistics*, 36(1), pp.161-195. <https://doi.org/10.1515/ling.1998.36.1.161>
- Hinton, L., Huss, L., & Roche, G. (Eds.).(2018). The Routledge handbook of language revitalization. Routledge.
- Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M.(2020). The state and fate of linguistic diversity and inclusion in the NLP world. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (Eds.), Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pp. 6282-6293, Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.560>
- Kukutai, T., & Taylor, J.(Eds.).(2016). Indigenous data sovereignty: Toward an agenda. ANU Press.
- Local Contexts.(n.d.). Traditional knowledge labels. <https://localcontexts.org/labels/traditional-knowledge-labels/>(Retrieved on: May 27, 2026)
- May, S.(2012). Language and minority rights: Ethnicity, nationalism and the politics of language (2nd ed.). Routledge.
- Mlaba, S., Tshamano, W., & Mabela, M. L.(2025). Evaluating the accessibility of digital resources for Nluu language revitalisation: The study of internet and social media. *E-Journal of Humanities, Arts and Social Sciences*, 6(5), pp.642-652. <https://doi.org/10.38159/ehass.20256520>
- Mohamed, S., Png, M.-T., & Isaac, W.(2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33, pp.659-684. <https://doi.org/10.1007/s13347-020-00405-8>
- Muehlberger, G., Seaward, L., Terras, M., Ares Oliveira, S., Bosch, V., Bryan, M., Colutto, S., Déjean, H., Diem, M., Fiel, S., Gatos, B., Greinoecker, A., Grüning, T., Hackl, G., Haukkovaara, V., Heyer,

- G., Hirvonen, L., Hodel, T., Jokinen, M., ... Zagoris, K.(2019). Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study. *Journal of Documentation*, 75(5), pp.954-976. <https://doi.org/10.1108/JD-07-2018-0114>
- Mumin, M., & Versteegh, K.(Eds.).(2014). The Arabic script in Africa: Studies in the use of a writing system. Brill.
- Ngom, F.(2016). Muslims beyond the Arab world: The odyssey of 'Ajami and the Murīdiyya. Oxford University Press.
- Paris, D.(2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational Researcher*, 41(3), pp.93-97. <https://doi.org/10.3102/0013189X12441244>
- Paris, D., & Alim, H. S.(Eds.).(2017). Culturally sustaining pedagogies: Teaching and learning for justice in a changing world. Teachers College Press.
- Savà, G., & Tosco, M.(2000). A sketch of Ongota, a dying language of Southwest Ethiopia. *Studies in African Linguistics*, 29(1), pp.59-136. <https://doi.org/10.32473/sal.v29i2.107366> (Retrieved on: May 27, 2026)
- Shah, S., & Brenzinger, M.(2017). Writing for speaking: The Nluu orthography. In M. C. Jones & D. Mooney (Eds.), *Creating orthographies for endangered languages*, Cambridge University Press, pp.109-125.
- UNESCO(2025, February 27). UNESCO and the promotion of languages in Africa: Cultural diversity and multilingualism. <https://www.unesco.org/en/articles/unesco-and-promotion-languages-africa-cultural-diversity-and-multilingualism>(Retrieved on: May 27, 2026)
- Unicode Consortium.(2025). The Unicode standard, version 17.0.0: Chapter 19, Africa. <https://www.unicode.org/versions/Unicode17.0.0/core-spec/chapter-19/>(Retrieved on: May 27, 2026)
- Woodbury, A. C.(2011). Language documentation. In P. K. Austin & J. Sallabank (Eds.), *The Cambridge handbook of endangered languages*, Cambridge University Press, pp.159-186.
- Yousuf, O., Aminu, A., Muhammad, M. S., Usman, B., Hashim, M. K., Nivre, J., Megyesi, B., & Høgel, C.(2025). A handwritten text recognition dataset for Ajami manuscripts in Fulfulde and Hausa. In X. C. Yin, D. Karatzas, & D. Lopresti (Eds.), *Document analysis*

and recognition-ICDAR 2025, pp. 620-637, Springer. https://doi.org/10.1007/978-3-032-04627-7_36

Conflict of interest statement

The author declares no conflicts of interest regarding this manuscript.

During the preparation of this manuscript, ChatGPT (OpenAI) was used for language editing and formatting assistance. All interpretations, arguments, and final wording were reviewed and approved by the author, who takes full responsibility for the content.

필자 소개

성 명 게브루 베텔

소 속 부산외국어대학교 아프리카연구소 HK연구교수

주 소 부산광역시 금정구 금샘로485번길 65 A519

전자우편 bethelg@bufs.ac.kr

Ⅰ 국문초록

아프리카 소멸위기 언어와 문자를 위한 공동체 주도형 AI - 디지털 언어정의를 위한 다문화교육 프레임워크 -

게브루 베텔
(부산외국어대학교)

본 연구는 인공지능(AI)이 아프리카 소멸위기 언어와 문자의 보존 및 재활성화를 어떻게 지원할 수 있는지를 다문화교육과 디지털 언어정의의 관점에서 검토한다. 방법론적으로 이 연구는 언어 기록과 재활성화, 다문화 · 다언어교육, 저자원 언어기술, 디지털 유산, 탈식민 AI, 공동체 데이터 거버넌스에 관한 연구를 종합하는 개념적 · 적용적 · 정책지향적 접근을 취한다. Banks의 다문화교육론과 May의 언어적 소수화 및 언어권리 논의를 문화적으로 지속가능한 교육과 공동체 데이터 주권의 관점과 결합하여, 언어적 인정과 재현, 언어권리와 공평한 접근, 공동체 행위성과 지식 권위, 세대 간 · 문화적으로 지속가능한 학습이라는 네 가지 분석 차원을 제시한다. 이를 바탕으로 AI 활용을 기록과 수집, 해석과 언어적 매개, 교육과 일상적 사용, 조정 · 인프라 · 거버넌스의 네 가지 기능으로 유형화한다. 또한 옹고타(Ongota), 케냐의 야아쿤테/야아쿠(Yaakunte/Yaaku), 남아프리카공화국의 누(Nluu), 아자미(Ajami) 필사본 전통을 네 가지 예시적 보존 시나리오로 분석하고, 아프리카 문자 기술을 여러 공동체에 적용되는 교차적 인프라 층위로 구분한다. 나아가 공동체 상담에서 지속가능성과 평가에 이르는 7단계 실행 로드맵과 추출적 데이터 거버넌스, 제한된 지식의 노출, AI 오류와 인식론적 왜곡, 인프라의 비지속가능성, 불평등한 제도 권력이라는 다섯 가지 위험 영역을 제시한다. 이 연구는 AI 기반 보존의 가치를 기술적 성능이나 데이터 규모만으로 평가해서는 안 되며, 언어적 인정, 교육적 접근, 공동체 권한, 세대 간 학습, 문화적 연속성에 대한 기여를 중심으로 평가해야 한다고 주장한다.

|주제어| 인공지능, 아프리카 소멸위기 언어와 문자, 다문화교육, 언어적 소수화, 공동체 거버넌스, 디지털 언어정의, 언어 재활성화

투고일 2026. 06. 18 / 심사일 2026. 07. 18 / 게재확정일 2026. 07. 21
--