Generative Artificial Intelligence and Indigenous Knowledge Systems in the Global South: A Systematic Review on Data Colonialism and Epistemic Justice

Cristian David Neira-Dàvila[1], Juan Camilo Aguirre-Rivera[2], Jhon Pablo Marín Yusti[3]

[1] *Unidad central del Valle - https://orcid.org/0009-0003-3929-5096*
[2] *Universidad de Manizales - https://orcid.org/0000-0002-0607-6271*
[3] *Unidad central del Valle - Universidad de Caldas https://orcid.org/0000-0002-4642-7013*

Received: 09/ August/2025 Revised: 25/ August 2025 Accepted: 29/ September/2025

Abstract

Introduction: The accelerated expansion of Generative Artificial Intelligence (GenAI) is reshaping contemporary ecologies of knowledge production, validation, and circulation, particularly in contexts marked by historical inequalities such as Latin America and the Global South. When these technologies are deployed over Indigenous Knowledge Systems (IKS), they not only mediate processes of cultural preservation and documentation but also intervene in broader disputes over data coloniality, epistemic sovereignty, and the risks of digital epistemicide. In this scenario, the relationship between GenAI and Indigenous knowledge becomes a strategic space for questioning who controls data infrastructures, under which frameworks Indigenous worlds are represented, and what conditions allow these knowledge systems to continue naming the world on their own terms.

Objectives: To examine how applications of Generative Artificial Intelligence (GenAI) within Indigenous Knowledge Systems (IKS) can contribute to epistemic and cultural justice in Latin America and the Global South.

Methods: A systematic review was conducted following the PRISMA protocol, with searches in Scopus and Web of Science between 2023 and 2025, complemented by Scopus AI–assisted retrieval. After screening, full-text reading, and applying inclusion/exclusion criteria, a final corpus of 15 studies was obtained. Thematic analysis was guided by four central categories: GenAI, Indigenous knowledge, Epistemologies of the South, and epistemic/cultural justice.

Results: The analysis identified four interrelated axes: A) Preservation and revitalization of knowledge; B) Culturally situated educational and evaluative tools; C) Data governance frameworks and knowledge cocreation; and D) Decolonial imagination of creativity. Together, these axes reveal a field that is both promising and shaped by dynamics of extractivism and epistemic dispossession.

Conclusions: GenAI can only be considered an ally of epistemic justice when anchored in data sovereignty, community cogovernance, and slow temporalities of deliberation, which stand in tension with the technoscientific acceleration associated with the shock of the future. It is concluded that horizons such as Slow AI and sociotechnical infrastructures designed from and with Indigenous peoples are necessary conditions for contesting data coloniality and reauthorizing their knowledge systems in the Global South.

Keywords: Generative Artificial Intelligence (GenAI); Indigenous Knowledge Systems (IKS); Epistemologies of the South; Epistemic Justice; Cultural Justice; Data colonialism; Indigenous data sovereignty.

1. Introduction

The accelerated expansion of Generative Artificial Intelligence (GenAI) has profoundly reshaped contemporary ecologies of knowledge production, validation, and circulation in the digital era. The incorporation of generative text, image, and multimodal models into fields such as education, heritage conservation, and scientific research has been documented as part of a broader process of digital transformation of academic and institutional infrastructures. This transformation expands capacities for search, analysis, and preservation, but it also redistributes centers of decision-making and cognitive authority in asymmetric ways (Ghaith & Hutson, 2024; Martínez-Camacho et al., 2025). In this scenario, Indigenous peoples and other historically marginalized groups are confronted with technologies whose deployment can enable processes of linguistic revitalization, cultural documentation, and expanded circulation of knowledge, while simultaneously reproducing epistemic hierarchies inherited from coloniality (Perera et al., 2025; Spano & Zhang, 2025) and opening new forms of digital or algorithmic epistemicide.

In this context, recent literature on AI and Indigenous peoples shows that these systems do not behave as neutral tools, but as artifacts embedded within knowledge regimes structured by the logics of the coloniality of power, knowledge, and being (Correa & Martens, 2025). Thus, in the realm of cultural representation, it has been argued that generative models operate as devices that project shadows over subordinated cultures by recombining dominant epistemic patterns that obscure non-Western ontologies, aesthetics, and narratives, even when they materially include vocabularies or images from other contexts (Karpouzis, 2024). This phenomenon is empirically corroborated in studies analyzing the performance of models such as Llama-2, GPT-4, or PaLM-2 when confronted with artifacts, symbols, and practices of Indian subcultures, showing a consistent correlation between social marginality and epistemic marginality within datasets and generative outputs. Such findings suggest that biases are not mere technical malfunctions but expressions of a global data architecture that hierarchizes which worlds are intelligible and which remain distorted (Seth et al., 2024).

These dynamics intersect with increasingly urgent debates on data sovereignty and Indigenous governance of knowledge. An illustrative example is the contrast between FAIR principles (Findable, Accessible, Interoperable, Reusable), which promote the openness, accessibility, and interoperability of scientific data, and the CARE framework (Collective Benefit, Authority to Control, Responsibility, Ethics), which centers collective benefit, authority over control, responsibility, and ethics. This contrast reveals how dominant "open science" logics conflict with the relational, collective, and spiritually grounded nature of Indigenous Knowledge Systems (IKS) (Taitingfong et al., 2024).

Other studies focusing on cultural heritage and Indigenous data in Aotearoa and Nunavut highlight how Western intellectual property regimes and global digital infrastructures fail to adequately protect practices of knowledge creation, stewardship, and transmission. As a result, communities have begun developing their own cultural licenses, community protocols, and autonomous governance models for the use of their data and cultural expressions, including those mediated by GenAI (Spano & Zhang, 2025). In this context, proposals have emerged advocating contractual clauses that allow technological collaborations to be terminated due to cultural misalignment, introducing explicit safeguards against projects that may threaten community wellbeing (Ruster & Brown, 2020).

In addition, fields such as conservation, restoration, and heritage reconstruction constitute particularly sensitive terrains for observing these tensions. The adoption of GenAI in the restoration and reconstruction of movable heritage introduces technical possibilities for generating plausible images and immersive narratives, but also raises the risk of producing representations that violate cultural protocols, distort historical memory, or displace the epistemic authority of custodial communities and local experts (Ocón et al., 2025).

Converging findings from archaeologists, curators, and conservators reveal both the potential of AI to enhance documentation, digital mapping, and predictive analysis of deterioration, and concerns regarding the excessive virtualization of heritage and the possible subordination of community-based criteria of authenticity to the technical parameters of generative models (Ghaith & Hutson, 2024). These concerns resonate strongly in the Latin American context, where university libraries have emerged as key nodes mediating between global editorial

infrastructures, AI-based platforms, and local regimes of knowledge production and access (Martínez-Camacho et al., 2025).

Moreover, the structural dimension of these processes becomes more visible when situated within decolonial frameworks that analyze AI as part of an expanded matrix of coloniality. From a Latin American perspective, it has been shown that the development and deployment of AI systems extend the logics of data colonialism, digital feudalism, and the coloniality of data, reorganizing access, extraction, and valuation of information on a planetary scale with disproportionate effects on Indigenous communities and other groups in the Global South—effects amounting to algorithmic forms of epistemicide (Correa & Martens, 2025). Parallel analyses of AI governance in sub-Saharan Africa reveal that most regulatory and public policy frameworks are still designed under assumptions and priorities of the Global North, with limited participation of local actors and technological infrastructures that rarely incorporate data produced within these territories, thereby constraining the possibility of exercising digital sovereignty and steering AI toward social and epistemic justice (Ayana et al., 2024).

Similarly, in Latin America, the development of national AI policies has been shaped by strong influence from multilateral organizations, development banks, and external regulatory models. This creates persistent tension between adopting global discourses of ethics and "trustworthy AI" and addressing situated risks such as the deepening of historical inequalities, the weakening of collective knowledge rights, and the marginalization of non-hegemonic epistemologies (Flórez, 2025). In response, scholars argue that decoloniality cannot merely be an auxiliary adjective within frameworks of "responsible AI," but must be a constitutive requirement of any reliability proposal aspiring to redistribute power within data value chains, incorporate local languages, ontologies, and values, and secure the effective participation of communities across all stages of the sociotechnical cycle (Wakunuma et al., 2025).

Within this landscape, epistemic and cultural justice emerge as essential normative axes for reorienting the relationship between GenAI and Indigenous knowledge systems. From the standpoint of epistemic justice, the challenge is not merely correcting biases in model outputs but contesting the conditions under which certain forms of knowledge are recognized as legitimate data, who is considered a credible epistemic subject, and which forms of recording, archiving, and recombination are acceptable in relation to IKS (Correa & Martens, 2025; Taitingfong et al., 2024). Cultural justice, in turn, demands that decisions on the digitization, circulation, and generative use of Indigenous images, narratives, and symbols be grounded in community protocols, data sovereignty frameworks, and governance structures that prioritize collective flourishing over the extractive logic of data platforms and markets (Spano & Zhang, 2025; Ruster & Brown, 2020).

Alongside these critical diagnoses, emerging proposals seek to envision more just sociotechnical futures. Some frameworks for digital transformation of IKS emphasize the centrality of situated pedagogies, culturally sensitive archival practices, and rigorous control of metadata to prevent the reification or hollowing out of community meanings in the digital environments where GenAI operates (Mthethwa, 2025). Others insist on the need to articulate these proposals with concrete institutional infrastructures capable of practically contesting the terms under which GenAI is integrated in Global South contexts (Martínez-Camacho et al., 2025; Ayana et al., 2024), such as networks of university libraries, community data platforms, and research consortia. In dialogue with this corpus, the present systematic review foregrounds the question of how GenAI applications within Indigenous knowledge systems can contribute to (or hinder) processes of epistemic and cultural justice in Latin America and the Global South, situating data coloniality as a structuring axis of the problem.

2. Methods

This systematic review was conducted following the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Page et al., 2021), which required making the research question explicit, defining inclusion and exclusion criteria, documenting the study selection flow (see Figure 1), and structuring each stage in accordance with protocol standards.

The review was guided by the following question: How can applications of Generative Artificial Intelligence (GenAI) within Indigenous Knowledge Systems (IKS) contribute to epistemic and cultural justice in Latin America and the Global South? This formulation delimited both the conceptual focus (the

intersection of GenAI, Indigenous Knowledge Systems, Epistemologies of the South, and epistemic/cultural justice) and the temporal and documentary scope of the study. Accordingly, the review examined publications indexed in Scopus and Web of Science (WoS) between 2023 and 2025, a period in which GenAI gained global relevance both technically and socio-epistemically.

The search strategy was structured using a replicable algorithm in Scopus and WoS, combining descriptors related to GenAI, Indigenous knowledge systems, and epistemic/cultural justice frameworks in the Global South:

TITLE-ABS-KEY ((“Indigenous knowledge” OR “Indigenous Knowledge Systems” OR “IK” OR “IKS” OR “Indigenous cultural heritage” OR “Indigenous data sovereignty” OR “Tribal Sovereignty”) AND (“GenAI” OR “Generative AI” OR “Generative Artificial Intelligence” OR “GAI”)). Using this algorithm, 17 records were retrieved from Scopus and 10 records from WoS. After identifying duplicates…

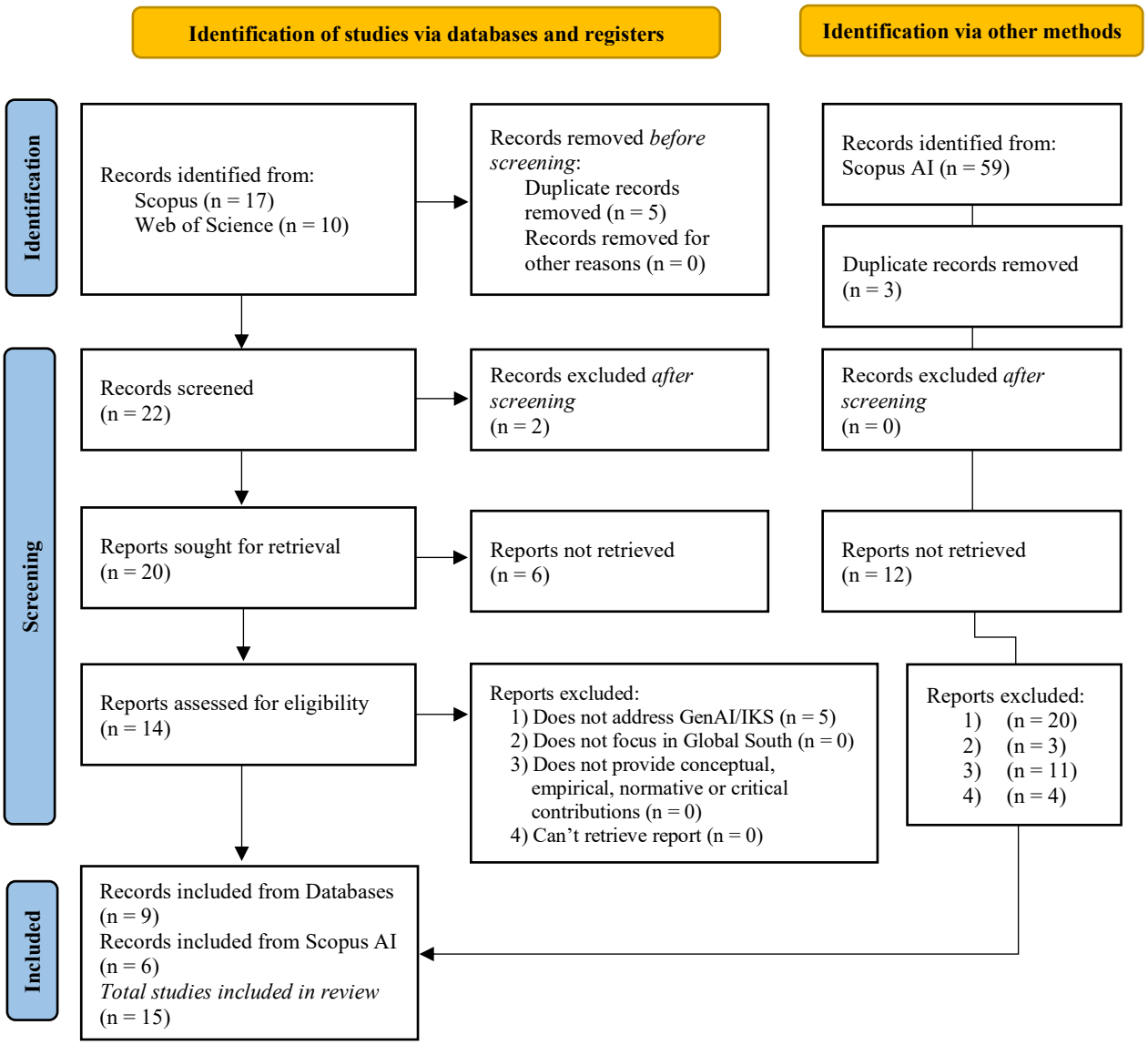

***Figure 1*:** Flowchart of the review process according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Page et al., 2021).

In WoS (half of the records), after removing 2 additional records during the title and abstract screening phase and excluding 6 records that could not be retrieved, 14 full-text documents were reviewed in depth. Of these, 5 were discarded for not fully meeting

the inclusion criteria, resulting in 9 studies that formed the initial core of the systematic corpus.

The inclusion criteria were as follows:

1. Thematic relevance, defined as the simultaneous presence of the categories GenAI and IKS;
2. A contextualized focus on the Global South (understood as an epistemic positionality, not merely a geographical location);
3. Conceptual, empirical, normative, or critical contributions relevant to the analysis;
4. Availability of full-text access.

In contrast, the exclusion criteria were:

1. Lack of thematic relevance—i.e., works that did not address GenAI/IAG and IKS simultaneously;
2. Exclusive focus on Global North contexts without problematizing the Global South as an epistemic positionality or addressing data coloniality;
3. Insufficient substantive contribution, including documents without meaningful conceptual, empirical, normative, or critical developments (e.g., short editorials, technical notes, or conference abstracts without analytical depth);
4. Inaccessibility of the full text.

Given the rapid evolution of the GenAI field, an additional AI-assisted search was conducted using Scopus AI as a discovery tool. This AI-mediated search enabled the identification of further studies related to the intersection of the key categories, yielding 59 potentially relevant records. After screening titles, abstracts, and introductions using the same inclusion and exclusion criteria described above, 6 studies were selected and added to the corpus obtained from Scopus and WoS, resulting in a final dataset of 15 studies. These were synthesized through in-depth reading and thematic coding guided by the four central categories, which informed the four analytical axes developed in the results section.

3. Results

The selected corpus shows that applications of GenAI and other forms of "creative" AI are unfolding across a constellation of settings in which Indigenous knowledge is not merely a context of application, but the terrain of dispute between data extractivism, epistemic sovereignty, and the possibilities of reauthorizing frameworks of cognitive justice from the Global South. Taken together, these works outline a map organized around four interwoven axes: (A) Generative technologies aimed at the preservation and revitalization of knowledge; (B) Culturally situated educational and assessment devices; (C) Data governance frameworks and knowledge cocreation; and (D) Decolonial imaginaries for reconfiguring creativity, value, and academic disciplinarity. These axes, by overlapping and mutually challenging one another, demonstrate that GenAI operates simultaneously as a technical tool, an apparatus of epistemic classification, and a site of political struggle.

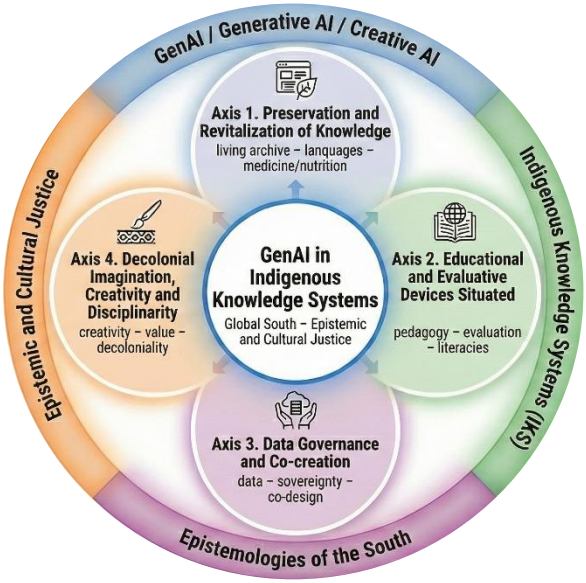

***Figure 2*:** Map of axes of articulation between GenAI and IKS from the global south.

The first axis explores the potential of GenAI and related digital infrastructures to document, recombine, and recontextualize traditional knowledge in fields such as ethnopharmacology, nutrition, oral repertoires, and library systems. Bhadra et al. (2025) propose "reimagining" ethnopharmacology through generative models capable of articulating scientific corpora with local medicinal knowledge, outlining a five-phase strategic framework that ranges from data digitization and curation to expert validation and regulated clinical deployment.

Thus, far from understanding GenAI as merely an

accelerator of drug discovery, the authors stress that any meaningful automation requires informed consent protocols, traceability of provenance, recognition of communities as co-producers of knowledge, and benefit-sharing mechanisms for the peoples who hold said knowledge. A related study examining digital nutrition platforms that incorporate traditional nutritional wisdom as input for personalized recommendations finds that the legitimacy of such systems depends on integrating principles of food sovereignty, algorithmic transparency, and community control into the translation of Indigenous culinary practices into nutritional metrics and user profiles (Suárez & Adibi, 2025).

This shift toward the preservation of living memory archives deepens in work focused on languages, stories, and intangible heritage. Analyses of GenAI as a tool for supporting oral history preservation show how automating tasks such as segmentation, transcription, and translation can reduce access gaps and generate narrative archives in endangered languages (Dueck, 2024). However, these benefits are deemed ethically acceptable only when anchored in cultural governance frameworks that respect confidentiality protocols, circulation restrictions, and collective rights over the narratives produced. Similarly, proposals on preserving Indigenous knowledge through digital technologies and AI tools highlight the role of libraries, repositories, and interactive platforms in sustaining community knowledge, exploring chatbots and virtual assistants capable of providing situated access without violating cultural protocols or intellectual property rights (Tella et al., 2025). These findings compel us to understand GenAI systems and digital repositories not as neutral containers of data, but as infrastructures that must incorporate the relationality intrinsic to Indigenous knowledge systems—that is, their ties to territories, languages, spiritualities, and lineages—in contrast to the decontextualizing tendencies of mainstream digital platforms.

The second axis is situated in the educational and evaluative domain, where results reveal a tension between uses of GenAI that reproduce colonial logics and experiences that mobilize it to strengthen the pedagogical agency of Indigenous communities. Kumu Connect is a paradigmatic example of place-based generative educational technology, where Hawaiian narratives, metaphors, and practices inform the architecture of the platform itself rather than merely its surface content (Baker-Ramos et al., 2025). Here, GenAI is conceived as a mediator that amplifies local genealogies and kaiapuni practices in learning processes rather than imposing universal curricular frameworks.

Similarly, Nyaaba et al. (2024) propose a conceptual framework for using GenAI in culturally responsive automatic generation of science assessment items, introducing a *Culturally Responsive Science Assessment* model that incorporates Indigenous language, local knowledge, and spirituality into scenario and question design. They show that automatic item generation only becomes meaningful when teachers—including Indigenous educators—critically deliberate about the biases, dominant metaphors, and ontological assumptions embedded in model outputs.

These dynamics become more complex when examining the pedagogical and disciplinary configurations that either reinforce or challenge coloniality of knowledge. Analyses of whether GenAI narratives in educational contexts reinscribe colonialism or reauthorize Indigenous perspectives show that uncritical incorporation of generative systems in schools and universities can reestablish colonial epistemic hierarchies by assuming that models determine what counts as legitimate knowledge while relegating Indigenous ontologies to mere local examples or ornamental cultural resources (Alterator et al., 2025).

Popow (2025) proposes a postdigital pedagogy for navigating epistemic transformations driven by GenAI, emphasizing the need for critical literacies that help students and teachers recognize AI as an apparatus of epistemic classification, situate its material conditions of production (infrastructures, labor, data), and practice epistemic disobedience in response to its recommendations. In the same axis, Tetteh et al. (2025) provide decisive empirical evidence showing how gender and parental education levels affect Indigenous teaching patterns, illustrating that any technological introduction operates within pedagogical ecologies already shaped by sociocultural inequalities. GenAI, therefore, does not enter a vacuum; it enters educational systems where coloniality, gender, and class intersect.

The third axis shifts attention from visible GenAI applications to the data infrastructures and governance

frameworks that make them possible. Lindgren et al. (2025) examine the struggle within technical communication against extractive linguistic modeling, proposing the articulation of FAIR principles (focused on data reuse) with Indigenous CARE principles (Collective Benefit, Authority to Control, Responsibility, Ethics) to reorient the documentation, annotation, and circulation of corpora used to train large language models. Likewise, analyses of data mining as a colonial practice from the perspective of Indigenous data sovereignty reveal that even when GenAI is not explicitly invoked, massive data collection infrastructures reproduce dispossession by separating data from the communities and territories that give them meaning (Roberts & Montoya, 2023). This diagnosis translates into normative proposals such as centering Indigenous data sovereignty in system design, establishing original codes of research, and delimiting "no-computation zones" where certain knowledge is neither digitized nor integrated into models.

In this same axis, the concept of tribal knowledge cocreation in GenAI environments is developed. Based on a case study with Native American communities, researchers argue that any integration of tribal knowledge into generative systems must involve prolonged codesign processes, differentiated access control, and community-defined criteria for determining which knowledge can or cannot be subject to algorithmic recombination (Wang et al., 2024). This shifts the emphasis from mere participation to authority over the limits of computation itself. Altogether, these findings suggest that epistemic justice in GenAI cannot be reduced to avoiding bias in model outputs; rather, it requires reconfiguring who decides what data are collected, how they are classified, what principles govern their sharing, and which knowledge forms remain intentionally outside computational circuits.

Finally, the fourth axis brings together proposals that aim to reconfigure creativity, cultural value, and academic disciplinarity from decolonial and Afro/Indigenous-centered frameworks. Jaramillo-Dent and Arora (2025) develop an anthropophagic approach to creativity and GenAI based on Ecuadorian artistic communities, identifying tensions between collectivism/individualism, process/product, and meaning/novelty, and proposing the design of anthropophagic technologies that are traceable, rejectable, augmentable, and capable of recentering community values and Indigenous roots in contrast to Global North imaginaries of copyright and originality. Aligned with this, the *creative data justice* framework expands this intuition by showing that data justice requires considering not only the distribution of benefits, but also the regimes of visibility, legibility, and silencing that AI imposes on Indigenous and community cultural practices. It calls for designing systems that make data genealogies traceable, enable strategic opacity, and allow the refusal of certain "creative" uses of corpora (Arora, 2025).

These perspectives converge with proposals aimed at reconfiguring academic disciplinarity and literacy practices. Eybers (2024), reflecting on academic literacy within Afrocentric and decolonial frameworks, shows how hegemonic curricula relegate African epistemologies to the status of "objects of study" and how academic writing practices can be reoriented so that Black and Indigenous students produce texts rooted in their cosmologies and genealogies of meaning. Similarly, Popow (2025) and Alterator et al. (2025) argue that pedagogical responses to GenAI will only be epistemically just if they shift the focus from mere digital literacy to a disciplinary reconfiguration that recognizes Southern epistemologies as sources of theory, method, and validity criteria—not simply as case studies or applications.

Thus, the findings of the corpus indicate that it is not only instrumental uses of GenAI that are being described; rather, a horizon is emerging in which generative technologies can contribute to epistemic and cultural justice insofar as they are embedded within broader projects of data sovereignty, knowledge cogovernance, and the reinvention of the very categories of creativity, evidence, and academic rigor from the Global South.

4. Discussion

The results of the review confirm that the applications of GenAI and other forms of "creative" AI unfold within a field marked by profound tensions between projects aimed at cultural revitalization, situated pedagogy, and data sovereignty, and assemblages shaped by technocolonialism, extractivism, and the reinforcement of epistemic hierarchies. In this sense, the question of whether GenAI can contribute to

epistemic and cultural justice in the Global South only allows for answers conditioned by the architectures of power, governance frameworks, and ontologies of knowledge within which it operates.

A first strong point of convergence lies in the need to decenter AI as the sole measure of intelligence and as a hegemonic epistemic reference. The proposal of "abundant intelligences" shifts the logic of cognitive scarcity, suggesting that AI should be understood as one intelligence among many within plural ecologies of knowledge, rather than as a universal validation criterion (Lewis et al., 2025). Complementarily, the Ḿn̓úxvʔit model centers haiłzaqv governance frameworks and conceives technology as a guest within an already existing epistemic order, rather than the reverse (White et al., 2024).

In contrast, analyses of technodigital colonialism and cognitive imperialism show that when GenAI is integrated without questioning the modern-colonial matrix, models tend to reinforce individualistic, extractive, and racialized ontologies, even when statistical bias-correction techniques are applied (Cambraia & Pyrrho, 2024; Ofosu-Asare, 2025). This tension becomes even more acute when Indigenous feminist perspectives are incorporated, highlighting how profiles of the "technological expert" and innovation narratives are shaped by gender, race, and geopolitics, with direct effects on which problems are prioritized and which forms of life are deemed expendable (Richardson, 2023). The review thus suggests that epistemic justice is not achieved by simply "adding" Indigenous content to existing infrastructures, but by reconfiguring the very foundations that define what counts as valid knowledge.

In the educational and evaluative sphere, empirical evidence is similarly ambivalent. Cultural audits of large language models reveal patterns of alignment with WEIRD worldviews even when model responses are not overtly discriminatory: these are biases reflected in the kinds of references privileged, the examples provided, and the presumed notions of academic or professional success (Mushtaq et al., 2024). Conversely, experiences using GenAI in graduate-level case study design indicate that when GenAI is employed as a tool for critical co-design and users retain agency to interrogate outputs, it is possible to generate materials more closely aligned with diverse student experiences (Jayasinghe et al., 2025).

In specific territorial contexts, the use of immersive virtual reality for cultural heritage in Amarilis (Huánuco) demonstrates that technologies can strengthen ties to territory and local memory when co-designed with community actors and when their reduction to tourist commodities is avoided (Jara et al., 2025). In contrast, the use of AI techniques to compute "identity distances" in the Waorani community highlights that quantitative operationalization of "culture" can solidify deficit-oriented narratives of acculturation and loss if not accompanied by participatory interpretive processes and critical interpretive frameworks (Espín-León et al., 2020).

Likewise, analyses of digital orientalism in generative image systems show that automatic production of visualities of "the Indigenous" or "the exotic" tends to condense stereotypes and clichés, reproducing colonial gazes that are later reinserted into educational, tourist, or media materials (Varghese & Rani, 2024). Taken together, these findings indicate that GenAI's contribution to epistemic justice in education depends less on the technical device itself and more on the pedagogical direction guiding its use.

Turning to the axis of data governance and sovereignty, several key insights emerge that are crucial for answering the research question. In contexts where data related to Indigenous peoples are understood as living treasures tied to genealogies, territories, and collective responsibilities, systems engineering ceases to be a merely technical matter and becomes a space of political and epistemic negotiation (Bowen & Hinze, 2022; White et al., 2024). In this regard, Participatory Data Design underscores that involving users in usability testing is insufficient; it is necessary to co-design what data are collected, for what purposes, who controls them, for how long, and under what conditions they may circulate.

In parallel, analyses of data sovereignty in India show how decolonization discourses can be appropriated by postcolonial states to legitimize regimes of surveillance and censorship, even when framed as defenses against Big Tech (Dutta & Mazumdar, 2025). This suggests that data sovereignty cannot be reduced to server geolocation or state control; rather, it requires data justice frameworks rooted in social movements,

community organizations, and critical research networks.

In the field of Indigenous knowledge repositories, the proposal to articulate FAIR and CARE principles provides a relevant normative framework for any generative assemblage intending to train on IK (Chigwada & Ngulube, 2025). Cases such as Mukurtu, SAICH, and Ulwazi show that it is possible to design platforms where communities define metadata, access levels, licenses, and future uses, so that data visibility and reuse remain subordinate to principles of collective benefit, authority to control, responsibility, and situated ethics. Translated to the GenAI context, this implies that no "open" or "FAIR" corpus is automatically legitimate as training data: the necessary condition for fair use is that data governance structures have been co-defined by the communities involved and that effective mechanisms exist for veto, review, and benefit sharing.

Complementarily, debates on "human-centered AI" in public policy suggest that dominant notions of human-centricity tend to reduce the human to "the user" and minimal rights, leaving out the community and social dimensions of justice (Sigfrids et al., 2023). Thus, without governance frameworks that integrate users, communities, society, and even planetary horizons, human-centered AI risks functioning as an ethics-washing device.

Finally, in the realm of imaginaries and decolonial design proposals, the notion of "liberatory collections" grounded in Black community archives proposes reimagining repositories not merely as memory deposits but as infrastructures of struggle, care, and self-representation—logics scarcely compatible with generative models that extract without reciprocity (Croskey et al., 2025). Similarly, the speculative figure of a decolonized digital library based on a conversational system that engages critically with the user, makes its limitations explicit, recognizes Indigenous data sovereignty, and renders visible the human labor and ecological footprint involved outlines an alternative form of GenAI: a reflective companion rather than a neutral oracle (Cox & Jimenez, 2025).

Another emerging proposal is that of Slow AI, which introduces temporality as an ethical dimension: in contrast to accelerated and homogenized deployments, a slow, iterative AI grounded in prolonged listening with marginalized communities can better grasp contexts and reduce the risks of technocratic imposition (Aristi et al., 2025). This proposal offers insight into what Toffler (1970/1971) described as "future shock": the disorientation produced by rates of change so rapid that they exceed the capacity of individuals and collectives to make sense of and govern technological transformations.

5. Conclusions

This systematic review of 15 studies shows that applications of GenAI in dialogue with Indigenous Knowledge Systems (IKS) constitute an emerging yet highly contested field, where initiatives for cultural preservation and revitalization coexist with proposals for data co-governance and persistent dynamics of epistemic extractivism. The four-axis analytical map demonstrates that GenAI contributes to epistemic and cultural justice only when embedded within infrastructures that guarantee community data sovereignty, effective participation of Indigenous peoples throughout the entire design, training, and evaluation cycle, and robust mechanisms for informed consent and traceability. At the same time, the review identifies significant gaps: a scarcity of projects led by communities and institutions in the Global South; limited exploration of intersecting variables such as gender, age, language, and diaspora; and few evaluations framed through Indigenous ontological and axiological perspectives. These gaps restrict the possibility of defining criteria for epistemic justice grounded in Indigenous worldviews and of assessing, from those standpoints, the effects of GenAI on their territories, languages, and ways of life.

Thus, within this broader digital epoch, the central issue may not be determining whether GenAI is "good" or "bad" in the abstract, but confronting the *future shock* anticipated by Toffler in the 1970s—an acceleration of technoscientific change that overwhelms collective capacities to make sense of and govern its consequences. In response to this horizon, proposals such as Slow AI—rooted in slower temporalities, prolonged listening, and situated co-design—remind us that what is at stake is not merely the architecture of models but the continuity of what makes us human: the ability to care for memories, relationships, and stories.

In this context, Indigenous peoples—living memory of plural origins and histories—cannot remain at the margins of the material and epistemic conditions that allow them to name the world in their own terms. Only insofar as these technologies submit to the rhythms, languages, and ways of life not designed by them, but by the peoples who inhabit them, can GenAI cease to function as yet another vector of future shock and epistemicide, and instead become a cautiously monitored ally in the struggle for epistemic and cultural justice in the Global South.

Refrences

[1] Alterator, S., Maddox, A., Southerton, C., & Schutt, S. (2025). Re-authorising colonialism or re-authorising indigenous perspectives: the risks of generative AI in indigenous education contexts. *British Journal of Sociology of Education,* 1-19. https://doi.org/10.1080/01425692.2025.2586732

[2] Aristi, S., Martinez Nimi, H., Chacon Quintero, J.C., Martinez Baca, A., & Ono, K. (2025). Slow AI in empowering and understanding marginalized communities: case study in Campana-Altamira, Mexico. In: Stephanidis, C., Antona, M., Ntoa, S., Salvendy, G. (eds), *HCI International 2025 Posters* (Part VIII, pp. 232-242). Springer. https://doi.org/10.1007/978-3-031-94171-9_20

[3] Arora, P. (2025). Creative data justice: a decolonial and indigenous framework to assess creativity and artificial intelligence. *Information, Communication & Society, 28*(13), 2231-2247. https://doi.org/10.1080/1369118X.2024.2420041

[4] Ayana, G., Dese, K., Daba, H., Habtamu, B., Mellado, B., Badu, K., Yamba, E., Landry, S., Ondua, M., Nsagha, D., Nkweteyim, D., & Dzevela, J. (2024). Decolonizing global AI governance: assessment of the state of decolonized AI governance in Sub- Saharan Africa. *Royal Society Open Science, 11*, 231994. https://doi.org/10.1098/rsos.231994

[5] Baker-Ramos, R., Gelder, W., Cho, L., Kolakaluri, J., & Hester, J. (2025). Kumu connect: design thinking for place-based generative educational technology in hawaiian immersion schools. CONFERENCIA

[6] Bhadra, S., Pundir, C., Mukherjee, M., Kar, A., Banerjee, S., Charoensup, R., Duangyod, T., & Mukherjee, P. (2025). Reimagining ethnopharmacology with generative AI: Towards inclusive, ethical, and data-driven traditional medicine. *Pharmacological Research, 221*, 108002. https://doi.org/10.1016/j.phrs.2025.108002

[7] Bowen, J., & Hinze, A. (2022). Participatory data design: managing data sovereignty in IoT solutions. *Interacting with Computers, 34*(2), 60–71, https://doi.org/10.1093/iwc/iwac031

[8] Cambrai, L., & Pyrrho, M. (2025). Generative artificial intelligence and the risk of thecnodigital colonialism. *Frontiers in Political Science, 7*, 01-07. https://doi.org/10.3389/fpos.2025.1628139

[9] Carghese, J. K., & Rani, P. (2024). Digital orientalism in machine vision: a cross-plataform analysis of AI-generated representations of Indian Culture. *The Russian Sociological Review, 23*(4), 113-139. https://doi.org/10.17323/1728-192x-2024-4-113-139

[10] Chigwada, J., & Ngulube, P. (2025). Integrating CARE and FAIR principles in the digital preservation of indigenous knowledge to achieve the sustainable development goals. In: P. Ngulube (Ed.), Harnessing Indigenous Epistemologies for Sustainable Progress (pp. 197-230) https://doi.org/10.4018/979-8-3693-8447-3.ch010

[11] Correa, H., Martens, C. (2025). Colonial structures in AI: a Latin American decolonial literature review of structural implications for marginalised communities in the Global South. *AI & Society*. https://doi.org/10.1007/s00146-025-02547-9

[12] Cox, A., & Jimenez, A. (2025). Chatting with the Decolonized Digital Library. *Library Trends 73*(4), 628-644. https://dx.doi.org/10.1353/lib.2025.a968499

[13] Croskey, P., Offert, F., Jacbos, J., & Thalar, K. M. (2025). Liberatory collections and thical AI: reimagining AI development from black community archives and datasets. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency* (pp. 900–913). ACM. https://doi.org/10.1145/3715275.3732058

[14] Dueck, G. W. (2024). Using AI to help preserve indigenous oral histories. In *2024 IEEE International Humanitarian Technologies Conference (IHTC)*. https://doi.org/10.1109/IHTC61819.2024.10855026

[15] Dutta, S., & Mazzumdar, S. (2025). Our data, ourselves: participation, justice, and alternative futures of data sovereignty in India. *Big Data & Society, 12*(3), 1-14. https://doi.org/10.1177/20539517251365235

[16] Eybers, O. O. (2023). Afrocentricity and decoloniality in disciplinarity: a reflective dialogue on academic literacy development. *Cristal, Critical Studies in Teaching & Learning, 11*(2), 48-64 https://doi.org/10.14426/cristal.v11i2.2003

[17] Espín-León, A. M., Jimeno-Morenilla, A., Pertegal-Felices, M. L., & Azorín-López, J. (2020). Distance computation through artificial intelligence as an analysis tool of the amazon indigenous people. A case study in the Waorani community. *Sustainability, 12*(22), 9513. https://doi.org/10.3390/su12229513

[18] Flórez, M. L. (2025). The shaping of AI policies in Latin America: a study of international influence and local realities. In: Goossens, J., Keymolen, E., Stanojević, A. (Eds.), *Public Governance and Emerging Technologies* (pp. 263-287). Springer. https://doi.org/10.1007/978-3-031-84748-6_12

[19] Ghaith, K., Hutson, J. (2024). A qualitative study on the integration of artificial intelligence in cultural heritage conservation. *Metaverse, 5*(2), 2654. https://doi.org/10.54517/m.v5i2.2654

[20] Jara, A. K., Saavedra, E. J., & Duran-Aquino, R. (2025). Inmersive virtual reality for tourism promotion in Amarilis, Huánuco: technological integration for cultural heritage. In *Proceedings of the 22nd LACCEI International Multi-Conference for Engineering, Education, and Technology*. Latin American and Caribbean Consortium of Engineering Institutions. https://doi.org/10.18687/LACCEI2024.1.1.923

[21] Jaramillo-Dent, D., & Arora, P. (2025). An Antropofagia approach to AI and creativity: Lessons from Latin America to rethink collectivity, process and meaning in creative value. *International Journal of Cultural Studies, 28*(6), 1210-1230. https://doi.org/10.1177/13678779251346959

[22] Jayasinghe, S., Arm, K., & Gamage, K. A. (2025). Designing culturally inclusive case studies with generative AI: strategies and considerations. *Education Sciences, 15*(6), 645. https://doi.org/10.3390/educsci15060645

[23] Karpouzis, K. (2024). Plato's Shadows in the Digital Cave: Controlling Cultural Bias in Generative AI. *Electronics, 13*(8), 1457. https://doi.org/10.3390/electronics13081457

[24] Lewis, J.E., Whaanga, H. & Yolgörmez, C. (2025). Abundant intelligences: placing AI within Indigenous knowledge frameworks. *AI & Sociology, 40*, 2141–2157. https://doi.org/10.1007/s00146-024-02099-4

[25] Lindgren, C. A., Yunes, E., Itchuaqiyaq, C. U. (2024). Technical communication's fight against extractive large language modeling by applying FAIR and CARE principles of data. *Journal of Business and Technical Communication, 39*(1), 11-25. https://doi.org/10.1177/10506519241280587

[26] Martínez-Camacho, H., Saavedra-Alamillas, C., Pacheco-Mendoza, J., & Machin-Mastromatteo, J. D. (2025). Generative artificial intelligence and university libraries in Latin America. *IFLA Journal*, 51(3), 647-659. https://doi.org/10.1177/03400352251348246

[27] Mthethwa, L. C. (2025). In the transit of indigenous knowledge systems through digital transformation framework. In: T. Maggie & M. Thulisile (Eds.), *Revaluation and preservation of indigenous knowkedge systems in modern society* (pp. 491-509). IGI Global. https://doi.org/10.4018/979-8-3693-7964-6.ch024

[28] Mushtaq, A., Rafay, M., Imran, M., Ghaznavi, I., & Qadir, J. (2025). Toward inclusive educational AI: auditing frontier LLMs through a multiplexity lens. *arXiv*, 1-11. https://doi.org/10.48550/arXiv.2501.03259

[29] Nyaaba, M., Zhai, X., & Faison, M. Z. (2024). Generative AI for culturally responsive science assessment: a conceptual framework. *Education Sciences, 14*(12), 1325. https://doi.org/10.3390/educsci14121325

[30] Ocón, D., Yin, C., & Luna, J. (2025). Artificial insights or historical fidelity? Crafting an ethical framework for the use of GenAI in the restoration, reconstruction and recreation of movable cultural heritage. *AI & Society*. https://doi.org/10.1007/s00146-025-02454-z

[31] Ofosu-Asare, Y. (2025). Cognitive imperialism in artificial intelligence: counteracting bias with indigenous epistemologies. *AI & Sociology, 40*, 3045–3061. https://doi.org/10.1007/s00146-024-02065-0

[32] Perera, M., Vidanaarachchi, R., Chandrashekeran, S., Kennedy, M., Kennedy, B., & Halgamuge, S. (2025). Indigenous peoples and artificial intelligence: a systematic review and future directions. *Big Data & Society, 12*(2), 1-22. https://doi.org/10.1177/20539517251349170

[33] Popow, M. (2025). Decolonizing knowledge in the postdigital era: Pedagogical strategies for navigating AI-driven epistemic transformations. *Educational Philosophy and Theory,* 1-13. https://doi.org/10.1080/00131857.2025.2552764

[34] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S… Alonso-Fernández, S. (2021). Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología, 74*(9), 790–799. http://doi.org/10.1016/j.recesp.2021.06.016

[35] Roberts, J., & Montoya, L. N. (2023). In Consideration of indigenous data sovereignty: data mining as a colonial practice. In K. Arai (Ed.), *Proceedings of the future technologies conference (FTC) 2023* (vol. 2., pp. 180-196). Springer. https://doi.org/10.1007/978-3-031-47451-4_13

[36] Richardson, T. A. (2023). Indigenous, feminine and technologist relational philosophies in the time of machine learning. *Ethics and Education, 18*(1), 6–22. https://doi.org/10.1080/17449642.2023.2188714

[37] Ruster, L., & Brown, G. (2020). Termination for cultural misalignment: setting up contract terms to ensure community well-being in the development of AI. *International Journal of Community Well-Being, 3*, 523–537. https://doi.org/10.1007/s42413-020-00081-8

[38] Seth, A., Ahuja, S., Bali, K., & Sitaram, S. (2024). DOSA: A dataset of social artifacts from different Indian geographical subcultures. *arXiv*. https://doi.org/10.48550/arXiv.2403.14651

[39] Sigfrids, A., Leikas, J., Salo-Pöntien, H., & Koskimies, E. (2023). Human-centricity in AI governance: A systemic approach. *Frontier in Artificial Intelligence, 6*, 01-09. https://doi.org/10.3389/frai.2023.976887

[40] Spano, I., & Zhang, Y. (2025). Indigenous data sovereignty in intangible cultural heritage governance: a complementary approach to public–private partnerships. *Interenational Journal of Cultural Property, 32*(2), 167-193. https://doi.org/10.1017/S0940739125100064

[41] Suarez, C., & Adibi, S. (2025). Integrating traditional nutritional wisdom into digital nutrition platforms: toward culturally adaptive and inclusive health technologies. *Nutrients, 17*, 1978. https://doi.org/10.3390/nu17121978

[42] Taitingfong, R., Martinez, A., Hudson, M., Lovett, R., Maher, B., Prehn, J., Rowe, R. K., Boileau, K., Franks, A., Khan, S., Walker, J. D., & Carroll, S. R. (2024). Aligning policy and practice to implement CARE with FAIR through Indigenous Peoples' protocols. *Acta Borealia, 41*(2), 80–90. https://doi.org/10.1080/08003831.2024.2410112

[43] Tella, A., Jatto, E. O., & Ajani, Y. A. (2025). Preserving indigenous knowledge: leveraging digital technology and artificial intelligence. *IFLA Journal, 51*(3), 703-721. https://doi.org/10.1177/03400352251342505

[44] Tetteh, A., Armah, M., Foli, J. Y., Okebukola, P., Peter, E., & Aboagye-Darko, N. O. (2025). Evaluating the impact of gender and parent education on indigenous teaching patterns. *The Journal of Educational Research*, 1–13. https://doi.org/10.1080/00220671.2025.2557292

[45] Toffler, A. (1971). *Future shock*. Banta Books. (Original work published 1970).

[46] Wakunuma, K., Ogoh, G., Akintoye, S., & Okaibedi, D. (2025). Decoloniality as an essential trustworthy AI requirement. In: Eke, D.O., Wakunuma, K., Akintoye, S., Ogoh, G. (Eds.), *Trustworthy AI*. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-75674-0_12

[47] Wang, Y-F., Chen, Y-C., Huand, Y-C., Redwing, C., & Tsai, C-H. (2024). Tribal knowledge cocreation in generative artificial intelligence systems. In *Proceedings of the 25th Annual International Conference on Digital Government Research*. Association for Computing Machinery. https://doi.org/10.1145/3657054.3657129

[48] White, E., Artelle K. A., Brown, E., Brown, K, Chan, D. E., & Housty, W. M̓únxv?it model for centering Indigenous knowledge and governance. *Conservation Biology, 38*(6), e14398. https://doi.org/10.1111/cobi.14398