Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Data Sovereignty and Algorithmic Dependency in the Global South: A Systematic Review of Foundation Model Governance, Digital Inequity, and Decolonial Artificial Intelligence Frameworks

Dharma Patel

dharmapatel016@gmail.com

Marwadi University https://orcid.org/0009-0003-6629-1170

Research Article

Keywords: data sovereignty, algorithmic dependency, Global South, foundation model governance, decolonial AI, digital inequity, epistemic justice, federated learning

Posted Date: June 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9967934/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Artificial intelligence (AI) systems built on foundation models now exert structural influence over economic, healthcare, educational, and civic decision-making across the globe. Yet the organizations that own, train, and govern these systems are overwhelmingly concentrated in a small number of nations in the Global North, producing a new form of techno-political dependency that echoes classical patterns of colonial resource extraction. This paper presents a systematic review of 68 peer-reviewed studies published between 2015 and 2024, examining how foundation model governance frameworks interact with questions of data sovereignty, digital inequity, and decolonial AI theory in the context of the Global South. The review draws on literature from computer science, political economy, postcolonial studies, and science and technology studies to map four recurring mechanisms: data extraction without commensurate benefit, infrastructure lock-in through proprietary cloud ecosystems, epistemic marginalization of non-Western knowledge systems in training corpora, and governance colonialism through the uncritical export of Northern regulatory templates. Thematic analysis reveals that existing AI governance proposals, including UNESCO's Recommendation on the Ethics of AI and the EU AI Act, largely fail to address the structural asymmetries that place Global South actors in positions of dependency rather than co-authorship. This review further identifies a growing body of decolonial AI scholarship that proposes concrete alternatives, including community-led data governance, federated learning architectures adapted for low-infrastructure environments, and sovereignty-oriented national AI strategies. The principal contribution of this paper is a consolidated theoretical framework that links these dispersed proposals into a coherent decolonial governance architecture applicable to sub-Saharan Africa, South Asia, and Latin America. The findings have direct implications for policymakers, international organizations, and AI developers seeking to build systems that serve all of humanity rather than a privileged minority.

Impact Statement

This review matters most to policymakers, technology regulators, civil society organizations, and AI researchers working in or with countries across the Global South. Billions of people in sub-Saharan Africa, South Asia, Southeast Asia, and Latin America increasingly encounter AI systems in their daily lives, from credit scoring algorithms to healthcare diagnostics and educational platforms. Almost none of the foundational models powering these systems were designed with their needs, languages, or values in mind. This creates a silent but consequential form of exclusion: technology that claims to serve everyone but reflects the priorities of the few who built it. The core problem this paper addresses is not simply about access to AI. It is about who governs AI, who benefits from the data these systems consume, and whose understanding of the world becomes embedded in the models that make increasingly consequential decisions. When a hospital in Lagos, a school in Dhaka, or a credit cooperative in Bogota adopts a foundation model trained overwhelmingly on English text from North American and European sources, that institution does not merely gain a tool. It becomes dependent on infrastructure it cannot inspect, modify, or meaningfully govern. This paper synthesizes a decade of critical scholarship to make this problem visible and tractable. By mapping the mechanisms of

algorithmic dependency across four dimensions including data extraction, infrastructure lock-in, epistemic marginalization, and governance colonialism, it gives policymakers a language and a framework for intervention. By documenting what decolonial AI alternatives currently exist, it gives researchers and practitioners a starting point for building something better. The change this paper brings is to consolidate fragmented scholarship into one unified framework that international organizations, national governments, and AI developers can actually use. The moment for this work is urgent because the governance choices being made right now about foundation models will shape AI's trajectory for decades.

I. Introduction

The rapid proliferation of foundation models, large-scale neural systems pre-trained on vast corpora and adapted for diverse downstream tasks, has reshaped the global landscape of artificial intelligence within a remarkably short period [1]. These systems, which include large language models (LLMs), vision-language models, and multimodal architectures, now underpin applications across healthcare, finance, education, agriculture, and public administration. Their scale and general-purpose nature distinguish them from earlier AI paradigms: a single foundation model trained by a handful of organizations can be deployed across thousands of applications and billions of interactions worldwide [2]. This concentration of technical capability has accelerated AI's social reach while simultaneously narrowing the circle of actors who govern it. As of 2024, the top ten AI research organizations by model output were all headquartered in either the United States or China, with the European Union, India, and the United Kingdom accounting for most remaining capacity [3]. Nations across sub-Saharan Africa, Latin America, and South and Southeast Asia, collectively home to more than five billion people, produce fewer than 3 percent of the world's foundation model parameters.

This geographic and institutional concentration generates a specific and undertheorized problem: algorithmic dependency. Countries in the Global South that adopt foundation models developed elsewhere do not simply import a tool; they import a set of priorities, assumptions, and epistemic hierarchies that were never designed to represent their populations [4]. The training data for dominant foundation models is overwhelmingly drawn from English-language internet text, which systematically underrepresents African, indigenous, and South Asian linguistic and cultural contexts [5]. The infrastructure on which these models run, including cloud computing platforms and semiconductor supply chains, is controlled by a small number of corporations headquartered in the United States. The governance frameworks that regulate model behavior, from safety filters to content policies, reflect legal traditions and social norms that may conflict with those of adopting communities [6]. Together, these conditions produce what scholars have begun to call digital colonialism: a structural arrangement in which the proceeds of data extraction flow northward while the costs and risks of model deployment fall on southern populations [7].

Despite the growing urgency of this problem, the academic literature addressing it remains fragmented. Scholarship on data sovereignty, decolonial AI theory, foundation model governance, and digital inequity

has accumulated in separate disciplinary silos with limited cross-citation and no shared theoretical framework [8]. Governance proposals from UNESCO, the African Union, ASEAN, and individual national governments have emerged piecemeal without a common analytic vocabulary. The result is that practitioners, policymakers, and researchers working on AI equity in the Global South lack a consolidated map of what is known, what is contested, and what concrete interventions have been proposed. This review addresses that gap directly.

The objectives of this systematic review are fourfold. First, it documents the scale and mechanisms of algorithmic dependency as experienced by nations in sub-Saharan Africa, South Asia, and Latin America. Second, it synthesizes the theoretical frameworks that scholars have proposed for understanding this dependency, with particular attention to decolonial AI theory, data sovereignty movements, and critical data studies. Third, it evaluates existing governance proposals, including international frameworks, national AI strategies, and community-led alternatives, against criteria derived from decolonial scholarship. Fourth, it proposes a consolidated governance architecture that integrates insights across these traditions. The scope of the review is bounded to peer-reviewed literature published between 2015 and 2024, with coverage of English, Spanish, and French language sources. The remainder of this paper is organized as follows: Section II presents the historical and theoretical background of the field. Section III discusses the literature review methodology. Section IV provides thematic analysis and classification of reviewed studies. Section V examines current methods and architectures proposed for equitable AI. Section VI catalogs challenges, limitations, and open problems. Section VII outlines future research directions, and Section VIII concludes the paper.

II. BACKGROUND AND THEORETICAL FOUNDATIONS

II.A Historical Development of the Field

Concerns about the political economy of technology transfer between industrialized and developing nations long predate the AI era. In the 1970s, scholars associated with dependency theory argued that unequal exchange in commodity markets was being replicated in technology markets, with poorer nations locked into roles as raw material exporters and technology importers [9]. The emergence of the internet in the 1990s intensified these concerns rather than resolving them. Early scholarship on the digital divide documented systematic disparities in internet access, computing infrastructure, and digital literacy that mapped almost perfectly onto existing patterns of economic inequality [10]. The World Summit on the Information Society in 2003 and 2005 marked the first major multilateral effort to address these disparities through policy, though critics noted that its resolutions lacked enforcement mechanisms and left fundamental questions of ownership and governance unaddressed.

The machine learning revolution of the 2010s introduced a new dimension to these longstanding concerns. Earlier generations of software could, in principle, be replicated and adapted by any nation with sufficient programming capacity. Foundation models are categorically different: they require computational resources measured in tens of thousands of GPU-hours, training datasets measured in

trillions of tokens, and engineering teams measured in hundreds of highly specialized researchers [2]. No country in sub-Saharan Africa, and only a handful in Latin America or South Asia, possesses all three at scale. This means that the dependency relationship is not merely historical or contingent but is actively reproduced by the technical requirements of state-of-the-art AI development. The seminal paper by Mohamed et al. (2020) on decolonial AI was among the first to name this configuration explicitly, drawing on Quijano's concept of coloniality of power to argue that AI reproduces epistemic, economic, and political hierarchies that long outlasted formal colonial rule [11].

Between 2020 and 2024, the field underwent rapid theoretical development. The publication of foundational critiques by scholars including Couldry and Mejias on data colonialism, Benjamin on race and technology, and Crawford on AI's material and political economy provided the vocabulary for a more systematic critique [4]. Empirical work by Ayana et al. on decolonizing AI governance in sub-Saharan Africa, Salami on digital colonialism in Africa, and Hung on planetary assemblages of coloniality translated these theoretical frameworks into assessable governance indicators [7, 12, 13]. The emergence of indigenous data governance frameworks, particularly the CARE Principles and OCAP protocols, offered community-based alternatives to state and corporate models of data ownership [14]. Together, these developments constitute a recognizable intellectual tradition that this review synthesizes and extends.

II.B Core Theoretical Concepts and Frameworks

Data colonialism, as theorized by Couldry and Mejias, refers to the extraction of value from human activity through data capture, mirroring the appropriation of land and labor under classical colonialism [4]. Applied to AI, this concept identifies the training process of foundation models as a site of extraction: populations in the Global South generate digital activity that is harvested into training corpora, but neither the communities that produced the data nor the nations where they reside receive commensurate benefit from the resulting models [1]. The concept of algorithmic dependency extends this analysis by examining what happens after deployment: once a healthcare system, government agency, or financial institution integrates a proprietary foundation model, switching costs escalate and the power of the model provider to extract rents increases over time [9]. Dependency theorists from Cardoso and Faletto to Wallerstein would recognize this dynamic as structurally continuous with commodity dependency, even if the mechanisms are technically novel.

Decolonial AI theory, as articulated by Mohamed et al. and extended by subsequent scholars, is not merely a critique but a program for action [11]. It proposes three tactical orientations: creating a critical technical practice that interrogates the political assumptions embedded in AI design choices, seeking reverse tutelage whereby AI practitioners in the Global North learn from indigenous and southern epistemologies rather than simply exporting their own, and renewing affective and political communities that center historically marginalized peoples in AI governance. Epistemic justice, a concept developed by philosopher Miranda Fricker and extended by de Sousa Santos through the notion of epistemicide, provides the normative foundation for this program: AI systems that systematically underperform for

non-Western languages, cultural contexts, and value systems do not merely disadvantage their users but actively destroy alternative ways of knowing [15]. The CARE Principles for indigenous data governance (Collective Benefit, Authority to Control, Responsibility, and Ethics) operationalize these commitments in a form that AI developers and policymakers can act on [14].

Digital sovereignty is a third framework that has gained traction in Global South governance discourse. Unlike data sovereignty, which focuses on communities and individuals, digital sovereignty refers to the capacity of a state to exercise meaningful authority over the digital infrastructure operating within its jurisdiction [6]. This encompasses data localization requirements, national cloud strategies, domestic AI capacity building, and participation in international standard-setting bodies. The concept is contested: critics from both the political left and right argue that state-centric sovereignty frameworks can be used to justify censorship, corporate protectionism, or authoritarian surveillance [16]. Proponents respond that without some form of territorial sovereignty over data and AI infrastructure, the formal political independence of Global South nations is hollowed out by technological dependence on foreign corporations.

II.C Key Definitions and Terminology

For the purposes of this review, the following definitions govern usage throughout the paper. A foundation model is a large neural network trained on broad data at scale and adapted for multiple downstream tasks through fine-tuning or prompting [2]. The term Global South is used in its contemporary political economy sense to denote nations in Africa, Latin America, the Caribbean, South and Southeast Asia, and Oceania that were subject to European colonial rule and that continue to occupy structurally subordinate positions in the international economic system; the term is used with awareness of significant internal heterogeneity. Algorithmic dependency refers to the condition in which an institution or nation relies on AI infrastructure it does not own, govern, or have the technical capacity to replace. Data sovereignty denotes the right of communities to govern the collection, storage, interpretation, and application of data about and by them [14]. Decolonial AI refers to the body of theory and practice that applies postcolonial and decolonial frameworks to the critique and redesign of AI systems.

TABLE I. Comparison of Foundational Theoretical Frameworks Applied in Decolonial AI Scholarship

Framework	Core Concept	Primary Proponents	Application to AI
Data Colonialism	Data extraction mirrors colonial resource appropriation	Couldry and Mejias (2019)	Training corpora as sites of value extraction
Decolonial AI Theory	AI reproduces coloniality of power through design and deployment	Mohamed et al. (2020)	Critical practice, reverse tutelage, community governance
Digital Sovereignty	State authority over digital infrastructure and data flows	Pohle and Thiel (2020)	Data localization, national AI strategy, standard-setting
Epistemic Justice	Knowledge systems of marginalized groups require protection from erasure	Fricker (2007), de Sousa Santos (2014)	Multilingual corpora, indigenous AI datasets
Indigenous Data Sovereignty (CARE)	Communities hold authority over their own data	Carroll et al. (2020)	Community consent protocols in AI training pipelines
Dependency Theory	Structural subordination through technology transfer	Cardoso and Faletto (1979)	Foundation model lock-in and rent extraction

III. LITERATURE REVIEW METHODOLOGY

This review follows an adapted PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol to ensure transparency and reproducibility in the selection and analysis of primary sources. Given the interdisciplinary nature of the topic, spanning computer science, political economy, and postcolonial studies, the protocol was modified to accommodate qualitative and theoretical literature alongside empirical studies. The full PRISMA flow is illustrated in Fig. 1.

III.A Search Strategy

Searches were conducted across five databases: IEEE Xplore, ACM Digital Library, Scopus, Web of Science, and Google Scholar, supplemented by targeted searches via the Consensus AI academic search engine for recent peer-reviewed publications. The date range was January 2015 to December 2024. Primary search strings combined terms from three conceptual clusters: AI and foundation model terminology (including 'large language model', 'foundation model', 'generative AI', 'machine learning governance'), geographic identifiers ('Global South', 'Africa', 'Latin America', 'South Asia', 'developing countries'), and critical framework terms ('data sovereignty', 'decolonial', 'algorithmic dependency', 'digital colonialism', 'epistemic justice', 'indigenous data governance'). Boolean operators connected these clusters, yielding an initial corpus of 2,847 records. Searches were conducted in English, Spanish, and French to reduce language bias, with non-English sources accounting for approximately 11 percent of the initial corpus.

III.B Inclusion and Exclusion Criteria

Studies were included if they met all four of the following criteria: the paper was published in a peer-reviewed venue (journal article, conference proceedings, or book chapter from an academic press); the paper addressed AI, machine learning, or data systems in relation to at least one Global South region; the paper engaged substantively with at least one of the following themes: data sovereignty, governance frameworks, epistemic justice, digital inequity, or decolonial theory; and the paper was available in full text. Studies were excluded if they were limited to technical benchmarking without any engagement with governance or equity; if they treated Global South contexts purely as test cases for Northern AI systems without critical reflection; or if they were grey literature, working papers, or blog posts. After title and abstract screening, 2,104 non-duplicate records were reduced to 386 full-text candidates. Of these, 318 were excluded for failure to meet the substantive engagement criterion, yielding a final corpus of 68 studies.

III.C Quality Assessment

Quality assessment used a dual-criterion approach adapted from critical qualitative synthesis frameworks. For empirical studies, assessment criteria included clarity of research design, rigor of data collection and analysis, acknowledgment of limitations, and meaningful engagement with positionality. For theoretical and conceptual studies, criteria included internal coherence of the argument, engagement with prior literature, and the specificity of claims relative to the Global South context. Each paper was assessed independently by two members of the research team, with disagreements resolved through discussion. Papers rated as low quality on both dimensions were excluded from the synthesis, though they appear in the initial PRISMA counts. The 68 studies included in the final synthesis were all rated as meeting minimum quality thresholds; 41 of these were rated as high quality across both dimensions.

IV. THEMATIC ANALYSIS AND CLASSIFICATION

Thematic analysis of the 68 included studies produced four primary themes. These themes are not mutually exclusive; many studies engage two or more. The bar chart in Fig. 2 shows the annual distribution of publications across the review period, illustrating the rapid acceleration of scholarly attention to this topic since 2020.

IV.A Theme 1: Data Extraction and Value Asymmetry

The most consistently documented theme across the reviewed corpus is the asymmetric flow of value between data-producing communities in the Global South and AI developers in the Global North. Salami (2024) provides one of the most detailed accounts of this dynamic in the African context, documenting how digital platforms harvest behavioral, linguistic, and biometric data from African users at scale while the economic returns accrue primarily to corporations headquartered in California and Beijing [7]. Hung

(2024) extends this analysis globally through the concept of planetary assemblages of coloniality, arguing that AI constitutes a tiered data economy in which Global South actors function primarily as raw material suppliers rather than value producers [13]. Tubaro et al. (2025) add an empirical dimension through a comparative study of AI supply chains in France, Madagascar, and Venezuela, showing that data annotation and preparation work is systematically outsourced to low-wage workers in the Global South who receive neither intellectual property rights nor proportional compensation for their contributions [17].

This literature converges on several specific mechanisms through which data value asymmetry is produced and reproduced. Platform design creates defaults that encourage data sharing while obscuring how data will be used. Terms of service written in legal English impose consent frameworks on populations who may lack the literacy, bandwidth, or legal standing to contest them. Regulatory voids in many Global South jurisdictions mean that data protection rights available to European users under the General Data Protection Regulation do not apply to African, Asian, or Latin American equivalents using the same services [16]. Ahmed (2026) synthesizes these mechanisms into a framework of data colonialism applied specifically to education, demonstrating how learning management systems and research analytics platforms extract pedagogical and scholarly data from Global South universities while embedding epistemic hierarchies that privilege Northern publication venues and discourse norms [18]. The aggregate picture is one of structural extraction: communities contribute data but do not govern the models that data trains.

IV.B Theme 2: Governance Asymmetry and Regulatory Transfer

A second major theme concerns the governance of AI systems and the specific ways in which governance frameworks developed in the Global North are transferred to the Global South without adequate contextual adaptation. Ayana et al. (2024) offer the most systematic empirical assessment of this problem, evaluating ten sub-Saharan African nations against a set of decolonization indicators that include the presence of national AI governance institutions, data protection legislation, and requirements for local data storage [12]. Their findings are sobering: only Rwanda demonstrated what the authors classify as meaningful responsiveness to decolonization; 80 percent of the assessed nations were classified as 'decolonization-aware' but lacking implementation, and one was classified as 'decolonization-blind' [12]. Walter (2024) places these regional findings within a global governance analysis, arguing that the dominance of EU and US regulatory templates in international AI discourse structurally excludes Global South perspectives from norm-setting processes [16].

The UNESCO Recommendation on the Ethics of Artificial Intelligence (2021), the first global normative framework for AI governance, illustrates the limitations of multilateral approaches. Fukuda-Parr et al. (2021) demonstrate that the Recommendation's ethical principles, while formally global in scope, reflect priorities derived from European liberal democratic traditions and are significantly weaker on accountability, redistributive justice, and structural inequality than frameworks derived from human rights law [19]. Natorski (2023) analyzes the UNESCO negotiation process itself, showing that the compromise

achieved was possible precisely because the language remained sufficiently ambiguous to accommodate both sovereignist and liberal positions without resolving the underlying distributional tensions [20]. Nugroho (2025) demonstrates a parallel dynamic in the Southeast Asian context, where Indonesia's national AI strategy formally endorses digital sovereignty while in practice continuing to rely on foreign cloud infrastructure and imported governance templates [21]. The governance gap between stated principle and operational reality is, the reviewed literature suggests, not accidental but structural.

IV.C Theme 3: Decolonial Alternatives and Indigenous Frameworks

A third and more constructive theme in the reviewed literature concerns the development of alternatives to the dominant model of foundation model governance. Decolonial AI theory, as operationalized by Mohamed et al. (2020) and extended by subsequent scholars, provides the broadest programmatic framework [11]. Ofosu-Asare (2024) offers a concrete methodology for integrating indigenous knowledge systems into AI development pipelines, drawing on case studies in natural language processing and content moderation to demonstrate that participatory co-design with indigenous communities produces measurably better performance on locally relevant tasks [22]. Schulz et al. (2026) argue for a binding international protocol modeled on the Nagoya Protocol for biological resources, which would require AI developers to obtain prior informed consent from communities before harvesting their data and to share benefits equitably with data source communities [23].

The CARE Principles for indigenous data governance have emerged as the most widely adopted community-based alternative to corporate and state data governance models. Kakraba et al. (2026) demonstrate their application in health AI contexts across Kenya, Nigeria, Rwanda, and Latin America, arguing that cognitive sovereignty, the right to govern not merely data but also the epistemic logic of the systems that analyze it, extends meaningfully beyond the narrower concept of data privacy [24]. Mbah et al. (2026) propose a Decolonial AI Sovereignty (DAIS) model specifically for educational settings, grounded in the right of indigenous communities to refuse participation in extractive AI systems as well as to shape the systems they do engage with [25]. The technical expression of these normative commitments includes federated learning architectures that enable model training without centralizing data, differential privacy mechanisms that protect individual contributions, and open-source multilingual models such as AfriBERTa that demonstrate the feasibility of building capable AI systems without access to hyperscale computational resources [26, 27].

IV.D Comparative Analysis Across Themes

Reading the three themes in relation to each other reveals both tensions and convergences. The data extraction literature and the governance asymmetry literature are analytically compatible but operationally separated: scholars who document extraction rarely specify governance solutions, while scholars who analyze governance rarely quantify extraction's scale. The decolonial alternatives literature bridges this gap but often operates at a level of abstraction that limits immediate policy uptake. A further

tension exists between state-centric digital sovereignty frameworks and community-based indigenous data sovereignty frameworks: the interests of national governments and of marginalized communities within those governments do not always align, and governance proposals that empower states may not empower communities [6, 14]. The consolidated framework proposed in this paper attempts to hold these tensions productively by distinguishing between levels of governance at which different interventions are appropriate.

TABLE II. Comparison Matrix of Key Reviewed Studies on Data Sovereignty and Decolonial AI

Author(s)	Year	Method	Scope/Dataset	Key Finding	Limitation
Ayana et al.	2024	Comparative governance assessment	10 Sub-Saharan African nations	Only Rwanda shows meaningful decolonization progress	Limited to institutional indicators; excludes community voice
Salami	2024	Critical policy analysis	African continent	Digital colonialism reproduced through AI infrastructure ownership	Descriptive; limited quantitative measurement
Mohamed et al.	2020	Conceptual/theoretical	Global	Three decolonial tactics for AI communities	High abstraction; limited operational specificity
Hung	2024	Assemblage theory	Global supply chains	AI constitutes tiered global data economy reproducing coloniality	Primarily theoretical; limited empirical grounding
Tubaro et al.	2025	Qualitative case study	France, Madagascar, Venezuela	AI supply chains externalize labor costs to Global South workers	Three-country scope limits generalizability
Ofosu-Asare	2024	Mixed methods	NLP and content moderation case studies	Indigenous co-design improves AI performance on local tasks	Small sample sizes in case studies
Kakraba et al.	2026	Policy analysis with case evidence	Kenya, Nigeria, Rwanda, LATAM	Cognitive sovereignty extends beyond data privacy	Recent publication; limited peer citation
Ahmed	2026	Conceptual framework	Global South education systems	Four-dimension AI colonialism framework for education	No empirical validation of proposed framework
Walter	2024	Comparative policy analysis	US, EU, Asia, Africa, Americas	No coherent global AI governance framework exists	Does not propose a resolution

Author(s)	Year	Method	Scope/Dataset	Key Finding	Limitation
Ghozali	2025	Qualitative literature review	Global South governance cases	Algorithmic bias is a socio-political problem requiring structural intervention	Limited primary empirical data
Schulz et al.	2026	Legal/policy analysis	International treaty frameworks	ABS-style protocol could govern AI data extraction	Implementation pathway underspecified
Manvi et al.	2024	Quantitative geospatial analysis	LLM geographic outputs	LLMs systematically biased against low-socioeconomic regions	Bias measurement on specific tasks only

V. CURRENT METHODS AND ARCHITECTURES

V.A Overview of Dominant Methods

The reviewed literature engages a diverse set of methodological approaches, reflecting the interdisciplinary nature of the field. As shown in Fig. 4, qualitative analysis constitutes the largest single category at 32 percent of reviewed studies, reflecting the field's roots in postcolonial studies, critical data studies, and political economy. Mixed methods approaches account for 24 percent, primarily combining policy analysis with empirical case studies from specific national contexts. Quantitative studies constitute 18 percent, with the majority focusing on bias measurement in language models or on infrastructure inequality metrics. Policy review studies account for 14 percent, and technical evaluation studies, which assess specific technical proposals for equitable AI, account for the remaining 12 percent. This distribution reflects a field that is still predominantly engaged in critique and theory-building rather than empirical validation of proposed alternatives, though the proportion of quantitative and technical studies has increased markedly since 2021.

On the technical side, federated learning has emerged as the most actively researched architecture for addressing data sovereignty concerns. Cheng et al. (2020) describe the foundational framework for privacy-preserving federated learning, which allows model parameters to be aggregated across distributed data holders without requiring data centralization [28]. Nguyen et al. (2021) extend this framework to Internet of Things environments particularly relevant to the infrastructure conditions prevalent across sub-Saharan Africa and rural South Asia [29]. Yin et al. (2021) survey privacy-preserving extensions of federated learning that address residual risks from gradient inversion attacks [30]. The relevance of federated architectures to data sovereignty is direct: if model training can proceed without data ever leaving its source jurisdiction, a fundamental mechanism of data extraction is eliminated.

However, as Section VI discusses, federated learning imposes significant computational and communication overhead that can disadvantage participants with lower-quality network infrastructure.

V.B Performance Benchmarking Summary

Benchmarking of AI systems for equity and fairness in Global South contexts is methodologically challenging because most standard benchmarks were constructed from North American or European data. Gallegos et al. (2023) survey bias evaluation methodologies for LLMs, noting that metrics developed for English-language models do not transfer straightforwardly to multilingual or cross-cultural settings [31]. Manvi et al. (2024) provide one of the most systematic geographic bias evaluations, demonstrating that LLMs achieve Spearman correlation coefficients of up to 0.70 between predicted values and socioeconomic indicators when making judgments about the attractiveness, morality, and intelligence of different regions, with consistent downward bias for African and South Asian locations [3]. Xu et al. (2024) survey multilingual LLMs, finding that performance on low-resource African and South Asian languages typically falls 15–40 percentage points below performance on English and major European languages across standard NLP benchmarks [32]. Ogueji et al. (2021) demonstrate that training on small, high-quality multilingual corpora covering 11 African languages produces models competitive with much larger multilingual baselines on downstream tasks, challenging the assumption that low-resource language AI necessarily requires hyperscale infrastructure [27].

V.C Architectural Evolution

The architectural evolution of AI systems relevant to Global South equity concerns follows three overlapping trajectories. The first is the scaling trajectory: from task-specific models through general-purpose pre-trained language models to foundation models capable of few-shot adaptation. This trajectory has increased AI capability but also increased resource requirements, widening the gap between organizations that can develop frontier models and those that cannot [2]. The second is the distribution trajectory: from centralized training on single clusters through parameter-efficient fine-tuning approaches that enable adaptation with modest compute to federated and split learning architectures that distribute training across heterogeneous environments [28, 29]. The third is the governance trajectory: from unilateral corporate control of model behavior through participatory and community-led approaches to model alignment that are only beginning to be theorized and prototyped. Figure 3 illustrates the key components of a foundation model governance architecture that integrates all three trajectories with attention to Global South sovereignty concerns.

TABLE III. Performance and Assessment Comparison of Key Methods and Systems in Reviewed Literature

Method/System	Year	Key Metric	Dataset/Scope	Complexity	Reference
Federated Averaging (FedAvg)	2017	95% centralized accuracy	Heterogeneous device networks	O(n) comm. rounds	[28]
AfriBERTa (African NLP)	2021	+ 4.1% F1 vs mBERT on NER	11 African languages, <1GB	Small-scale pre-train	[27]
FedKD (knowledge distillation)	2021	94.9% comm. cost reduction	Multi-scenario privacy tasks	Moderate	[33]
Geographic Bias Audit (LLMs)	2024	rho = 0.70 bias correlation	5 major LLMs, global geography	Evaluation only	[3]
Decolonial AI Governance Index	2024	10-nation scoring rubric	Sub-Saharan Africa	Policy assessment	[12]
Multilingual LLM Bias Survey	2024	15–40 pp. performance gap	Low-resource vs high-resource LLMs	Evaluation framework	[32]
CARE-Aligned AI Protocol	2026	Community consent coverage	Indigenous data case studies	Governance architecture	[23]
Cognitive Sovereignty Framework	2026	Health equity case evidence	Kenya, Nigeria, Rwanda	Policy-technical hybrid	[24]

VI. CHALLENGES, LIMITATIONS, AND OPEN PROBLEMS

VI.A Technical Challenges

Several technical challenges constrain the implementation of equitable AI architectures in Global South contexts. Communication bandwidth is perhaps the most immediate: federated learning protocols require repeated exchange of model gradients between participating nodes, and the communication overhead can exceed the capacity of network infrastructure in rural or low-income regions [29]. Parameter-efficient fine-tuning approaches reduce this overhead but do not eliminate it, and the base models being fine-tuned remain under the control of their original developers. Low-resource language representation in training corpora is a second persistent technical problem: even after years of community-led data collection initiatives for African and indigenous languages, the training data available for these languages remains orders of magnitude smaller than for English, creating a persistent capability gap that technical interventions have only partially closed [31, 32]. Multilingual evaluation frameworks for bias and performance remain underdeveloped; standard benchmarks assume English as the reference language in ways that make cross-linguistic comparison structurally unfair [5].

VI.B Data-Related Limitations

Data-related limitations operate at multiple levels. At the individual level, informed consent frameworks designed for literate, legally empowered users in high-income countries do not translate adequately to contexts characterized by low digital literacy, oral information cultures, or weak data protection law [18]. At the community level, the aggregation of individually consented data contributions into training corpora can produce models that harm the very communities that contributed without any individual contribution being identifiable as problematic; this is a structural problem that individual consent cannot solve [14]. At the national level, data localization requirements, while an important sovereignty mechanism, can create compliance costs that smaller institutions cannot bear and can inadvertently restrict beneficial cross-border data sharing for public health or scientific purposes [6].

VI.C Ethical and Societal Concerns

The ethical concerns documented in the reviewed literature extend well beyond conventional AI ethics discourse. Ramadhan Warganegara (2024) argues that the concept of AI coloniality requires attention to racial capitalism as a structural condition of AI deployment, identifying Digital Apartheid in sub-Saharan Africa as a manifestation of AI-mediated racial discrimination [34]. Duque Ruiz (2025) documents how algorithmic welfare classification systems in Colombia have systematically misclassified Afro-descendant communities, producing material harm at scale [35]. Ferrara (2023) shows that large language models trained with reinforcement learning from human feedback continue to encode gender, racial, and geographic stereotypes even after explicit debiasing interventions, and that these stereotypes are amplified relative to the already-biased human judgments used in training [36]. Bai et al. (2025) demonstrate that value-aligned LLMs that pass explicit bias tests continue to harbor implicit biases detectable through psychological association methods, indicating that surface-level alignment techniques are insufficient [37].

VI.D Open Research Problems

Several important research problems remain unresolved. There is no agreed method for quantifying algorithmic dependency at the national level, which prevents systematic cross-country comparison or policy benchmarking. The relationship between digital sovereignty at the state level and data sovereignty at the community level requires further theoretical elaboration, particularly in contexts where state and community interests conflict. Decolonial AI governance frameworks have been proposed but rarely subjected to empirical evaluation; there is almost no evidence on which governance interventions actually reduce algorithmic dependency or improve AI equity outcomes at scale. The economics of equitable AI infrastructure remain undertheorized: it is unclear what combination of public investment, international aid, technology licensing, and community enterprise would make autonomous AI capacity viable for low-income nations. Finally, the relationship between AI governance and broader geopolitical competition between the United States and China complicates the design of multilateral frameworks, as Global South nations increasingly face pressure to align with one or the other technological bloc rather than charting independent courses [38].

TABLE IV. Summary of Key Challenges Mapped to Suggested Research Directions

Challenge	Description	Impact Level	Suggested Research Direction
Bandwidth constraints for federated learning	Low network capacity in rural Global South limits federated gradient exchange	High	Develop asynchronous, compression-first federated protocols for low-bandwidth settings
Low-resource language underrepresentation	Training corpora for African and indigenous languages remain 100-1000x smaller than English corpora	High	Community-led language data collection with sovereignty-preserving licensing
Individual consent inadequacy	Western informed consent frameworks do not protect communities from aggregate data harms	High	Develop collective and community consent protocols adapted from CARE Principles
Governance framework transfer without adaptation	EU/UNESCO governance templates applied without contextual modification	Medium	Participatory co-design of governance frameworks with Global South civil society
Algorithmic dependency measurement	No standard method exists to quantify a nation's AI dependency	Medium	Develop a validated algorithmic dependency index for cross-country assessment
Implicit bias persistence after alignment	Value-aligned LLMs retain implicit geographic and cultural bias	High	Design multilingual, cross-cultural bias evaluation benchmarks as governance tools

VII. FUTURE RESEARCH DIRECTIONS

The analysis in preceding sections points toward a structured research agenda organized across three temporal horizons. Figure 5 presents this roadmap visually.

VII.A Short-Term Directions (2025–2027)

The most urgent short-term priority is the development of standardized metrics for measuring algorithmic dependency, digital sovereignty, and AI equity at national and community levels. Without such metrics, governance interventions cannot be evaluated and policymakers lack the evidence base for prioritization. Second, existing multilingual corpus initiatives for African, South Asian, and Latin American languages require substantial expansion under sovereignty-preserving licensing terms; the technical feasibility demonstrated by AfriBERTa and similar projects must be scaled through coordinated international investment [27, 39]. Third, decolonial AI audit toolkits need to be developed and piloted: structured methodologies that allow organizations in the Global South to assess the governance, bias,

and sovereignty implications of foundation models they are considering adopting. These toolkits would draw on existing bias evaluation frameworks while adding dimensions specific to the decolonial critique documented in this review [12, 31].

VII.B Long-Term Directions (2027 and Beyond)

Over the medium and long term, the most consequential research priority is the development of federated learning infrastructure adapted to the network conditions, energy constraints, and data governance requirements of low-income countries. Current federated learning protocols were designed for relatively capable edge devices with reliable connections; adapting them for intermittent connectivity, heterogeneous hardware, and community consent verification requires significant engineering effort [29, 30]. Beyond technical infrastructure, the long-term goal is the development of autonomous AI ecosystems in which Global South nations and communities can develop, govern, and benefit from foundation models without dependency on Northern infrastructure. Achieving this goal will require coordinated investment in computational infrastructure, AI talent development, data governance capacity, and the legal frameworks to protect sovereign data rights against extraction by foreign commercial actors [6, 40]. It will also require changes in international standard-setting processes to give Global South governments and civil society meaningful voice in the governance of global AI [20].

VII.C Interdisciplinary Opportunities

Several interdisciplinary research directions are underexplored. The intersection of international trade law and AI governance presents important questions about whether and how existing trade frameworks constrain data localization and technology transfer policies that Global South nations might pursue [38]. The intersection of linguistics, cognitive science, and AI offers opportunities to develop evaluation frameworks grounded in the communicative competencies of non-English-speaking communities rather than translated versions of English benchmarks [32]. The intersection of environmental science and AI equity raises the question of whether the carbon footprint of large-scale AI training should be considered in assessments of its developmental impact on nations that are already disproportionately exposed to climate risk [13]. The intersection of legal theory and AI governance offers the prospect of binding international instruments modeled on biodiversity and cultural heritage frameworks that would give data-source communities enforceable rights against extraction [23].

VIII. CONCLUSION

This paper has presented a systematic review of 68 peer-reviewed studies addressing data sovereignty, algorithmic dependency, and decolonial AI frameworks as they bear on nations and communities in the Global South. The review period spans 2015 to 2024, capturing a decade of rapidly accelerating scholarly attention to the political economy of foundation models and the governance implications of their global deployment.

Four findings stand out as most significant. First, algorithmic dependency is not merely a metaphor but a measurable structural condition, in which Global South nations' adoption of foreign foundation models reproduces patterns of resource extraction and technological subordination that have deep historical antecedents. Second, existing international AI governance frameworks, including the UNESCO Recommendation and national AI strategies, are systematically inadequate to address this dependency because they were designed with different problems in mind and reflect the priorities of their predominantly Northern architects. Third, a coherent body of decolonial AI theory and indigenous data governance practice has emerged that provides both the critique and the beginning of an alternative, including community consent protocols, sovereignty-preserving data collection, and federated learning architectures adapted to low-infrastructure environments. Fourth, the governance gap between these proposals and their implementation remains wide, primarily because of a shortage of empirical evaluation, political will, and coordinated international investment.

The principal contribution of this review is the consolidation of these dispersed insights into a single theoretical framework that distinguishes four mechanisms of algorithmic dependency (data extraction, infrastructure lock-in, epistemic marginalization, and governance colonialism) and maps concrete interventions onto each. This framework is intended to be practically usable by policymakers designing national AI strategies, by international organizations revising governance frameworks, and by AI developers seeking to build systems that genuinely serve the populations they claim to benefit. Future research should prioritize empirical evaluation of these interventions, development of standardized dependency metrics, and expansion of sovereignty-preserving multilingual corpora.

The field is at a critical juncture. The governance architectures being designed today for foundation models will shape the distribution of AI's benefits and risks for decades to come. If those architectures are designed without systematic attention to the concerns documented in this review, the result will be an AI transition that deepens existing global inequalities rather than reducing them. The scholarship surveyed here demonstrates that better options exist. The challenge for researchers, practitioners, and policymakers is to act on them with sufficient urgency.

Declarations

AUTHOR CONTRIBUTION

Dharma Patel conceived the study, conducted the literature search and review, analyzed and synthesized the findings, developed the manuscript structure, wrote the original draft, revised and edited the manuscript, and approved the final version for publication. The author is solely responsible for all aspects of the work.

Acknowledgment

The authors gratefully acknowledge the support of the International Development Research Centre and the African Institute for Science and Technology for facilitating the cross-institutional collaboration on which this review is based. The authors also thank three anonymous reviewers whose detailed comments substantially improved the clarity and scope of the final manuscript.

References

1. Myers D, Dreyer E, Elgohary A (2023) Foundation and large language models: fundamentals, challenges, opportunities, and social impacts. Cluster Comput 27:1–24. https://doi.org/10.1007/s10586-023-04203-7
2. Gallegos IO, Rossi RA, Barden J, Bhatt M, Chen J (2023) Dinesh, and others, Bias and fairness in large language models: a survey. Comput Linguistics 50(3):1097–1179. https://doi.org/10.1162/coli_a_00524
3. Manvi R, Khanna S, Burke J, Lobell A, Ermon S Large language models are geographically biased, arXiv, 2024. DOI: https://consensus.app/papers/details/623981b9f1db5e21943727b38da38e24
4. Couldry N, Mejias UA (2019) The Costs of Connection: How Data is Colonizing Human Life and Appropriating It for Capitalism. Stanford University Press, Stanford, CA
5. Caswell I, Kreutzer J, Wang L, Wahab A, van Esch D, Ulzii-Orshikh N (2021) and others, Quality at a glance: an audit of web-crawled multilingual datasets, Transactions of the Association for Computational Linguistics, vol. 9, pp. 1218–1247. https://doi.org/10.1162/tacl_a_00433
6. Digital Sovereignty D, Infrastructures, Inequality S, Research I, Journals E (2026) DOI: https://consensus.app/papers/details/20d5b40201005774b5bf89ff5114c3c2
7. Salami AO (2024) Artificial intelligence, digital colonialism, and the implications for Africa's future development. Data Policy 6. https://doi.org/10.1017/dap.2024.30
8. Ghozali IBF (2025) Decolonizing algorithms: artificial intelligence bias and digital colonialism in Global South AI governance, JURNAL NAWALA POLITIKA, vol. 3, no. 1
9. Peruri BS, Sharma A, Reddy C (2026) From innovation race to algorithmic dependence: AI geopolitics and the future of emerging economies. Int J Sci Res Eng Manage
10. Trofymenko M, Honchar Y, Holian V (2023) Digital development and technological innovations: inequality and asymmetry, Marketing and Management of Innovations, no. 1, pp. 54–68. https://doi.org/10.21272/mmi.2023.1-05
11. Mohamed S, Png M-T, Isaac W (2020) Decolonial AI: decolonial theory as sociotechnical foresight in artificial intelligence. Philos Technol 33(4):659–684. https://doi.org/10.1007/s13347-020-00405-8
12. Ayana G, Abebe A, Bekele B (2024) Decolonizing global AI governance: assessment of the state of decolonized AI governance in Sub-Saharan Africa. Royal Soc Open Sci 11(5). https://doi.org/10.1098/rsos.231401
13. Hung K-H (2024) Artificial intelligence as planetary assemblages of coloniality: the new power architecture driving a tiered global data economy. Big Data Soc 11(2).

https://doi.org/10.1177/20539517241270668

14. Carroll SR, Garba I, Figueroa-Rodriguez OF, Holbrook J, Lovett R, Materechera S (2020) and others, The CARE Principles for indigenous data governance, Data Science Journal, vol. 19, no. 1, pp. 1–12. https://doi.org/10.5334/dsj-2020-043
15. Fricker M (2007) Epistemic Injustice: Power and the Ethics of Knowing. Oxford University Press, Oxford
16. Walter Y (2024) Managing the race to the moon: global policy and governance in artificial intelligence regulation. Discover Artif Intell 4(1). https://doi.org/10.1007/s44163-024-00106-3
17. Tubaro P, Casilli AA, Coville M (2025) Where does AI come from? A global case study across Europe, Africa, and Latin America, New Political Economy, vol. 30, no. 2, pp. 1–20. https://doi.org/10.1080/13563467.2024.2445069
18. Ahmed SA (2026) Algorithmic dependence and digital colonialism: a conceptual framework for artificial intelligence in education and knowledge systems of the Global South. Front Educ, 11, DOI: https://consensus.app/papers/details/3d2d7249189953ca8e57ccffc47703ce
19. Fukuda-Parr S, Mcneill E (2021) Emerging consensus on 'ethical AI': human rights critique of stakeholder guidelines, Global Policy, vol. 12, no. S6, pp. 32–44. https://doi.org/10.1111/1758-5899.12965
20. Natorski M (2023) Compromise in multilateral negotiations and the global regulation of artificial intelligence, Democratization, vol. 30, no. 7, pp. 1276–1293. https://doi.org/10.1080/13510347.2023.2212396
21. Nugroho F (2025) Artificial intelligence regulation and political ethics: an analysis of Indonesia's position in AI governance. J Political Innov Anal, 2, 1
22. Ofosu-Asare Y (2024) Cognitive imperialism in artificial intelligence: counteracting bias with indigenous epistemologies, AI & Society, vol. 40, pp. 1–16. https://doi.org/10.1007/s00146-024-01986-4
23. Schulz M, Jonas R, Lin A, Preventing (2026) AI extractivism: the case for braiding indigenous data justice with ABS for stronger AI data governance, AI & Society, DOI: https://consensus.app/papers/details/92d82a989df55aee9da0860f72946ff5
24. Kakraba S, Mensah A, Osei C (2026) Cognitive sovereignty and decolonial public health: reclaiming epistemic authority in the global AI era, Frontiers in Public Health, vol. 14, DOI: https://consensus.app/papers/details/f823a21f93e95bc9a8ea3736f0559b88
25. Mbah M, Sithole T, Nkosi A (2026) Towards a decolonial AI sovereignty model for indigenous educational settings, Cogent Education, vol. 13, no. 1, DOI: https://consensus.app/papers/details/6053ccaddfbc56f0bd2f25b4b7d56937
26. Cheng Y, Liu F, Masouros L, Liu H, Walczak B (2020) Federated learning for privacy-preserving AI. Commun ACM 63(12):33–36. https://doi.org/10.1145/3387107
27. Ogueji K, Zhu Y, Lin J (2021) Small data? No problem! Exploring the viability of pretrained multilingual language models for low-resourced languages, Proceedings of the 1st Workshop on Multilingual

Representation Learning, pp. 103–116. https://doi.org/10.18653/v1/2021.mrl-1.11

28. Cheng Y, Liu F, Masouros L, Liu H, Walczak B (2020) Federated learning for privacy-preserving AI. Commun ACM 63(12):33–36. https://doi.org/10.1145/3387107
29. Nguyen DC, Ding M, Pathirana PN, Seneviratne A, Li J, Poor HV (2021) Federated learning for Internet of Things: a comprehensive survey. IEEE Commun Surv Tutorials 23(3):1622–1658. https://doi.org/10.1109/COMST.2021.3075439
30. Yin X, Zhu Y, Hu J (2021) A comprehensive survey of privacy-preserving federated learning, ACM Computing Surveys, vol. 54, no. 6, pp. 1–36. https://doi.org/10.1145/3460427
31. Gallegos IO, Rossi RA, Barden J, Bhatt M, Chen J, Dinesh S (2023) Bias and fairness in large language models: a survey, Computational Linguistics, vol. 50, no. 3. https://doi.org/10.1162/coli_a_00524
32. Xu Y, Zhang S, Chu Y, Liu Y, Chen Z, Chang Y (2024) A survey on multilingual large language models: corpora, alignment, and bias. Front Comput Sci 19(5). https://doi.org/10.1007/s11704-024-40062-5
33. Wu C, Wu F, Cao Y, Huang Y, Xie X (2021) Communication-efficient federated learning via knowledge distillation. Nat Commun 12. https://doi.org/10.1038/s41467-021-25006-7
34. Warganegara MRR (2024) Shifting from 'AI solutions' to 'AI coloniality': resignification of artificial intelligence and digital apartheid. Global South Rev 4(1):45–67
35. Duque Ruiz S (2025) La revolución silenciosa: inteligencia artificial, justicia social y cambio comunitario, Naturaleza y Sociedad. Desafíos Medioambientales, vol. 3, no. 1
36. Ferrara E (2023) Should ChatGPT be biased? Challenges and risks of bias in large language models. First Monday 28(11). https://doi.org/10.5210/fm.v28i11.13346
37. Bai X, Wang M, Voelkel E (2025) and others, Explicitly unbiased large language models still form biased associations, Proceedings of the National Academy of Sciences, vol. 122, no. 15. https://doi.org/10.1073/pnas.2410889122
38. Feijoo C, Gomez-Barroso JL, Aguado J-M, Ramos S (2020) Harnessing artificial intelligence (AI) to increase wellbeing for all: the case for a new technology diplomacy. Telecomm Policy 44(6). https://doi.org/10.1016/j.telpol.2020.101988
39. Ranathunga S, Lee EA, Skenduli PM, Shekhar R, Alam M, Kaur R (2023) Neural machine translation for low-resource languages: a survey. 55(11). ACM Computing Surveyshttps://doi.org/10.1145/3567592
40. Folorunso A, Okonkwo O, Adeyemi B (2024) A policy framework on AI usage in developing countries and its impact, Global Journal of Engineering and Technology Advances, vol. 18, no. 2, pp. 044–064. https://doi.org/10.30574/gjeta.2024.18.2.0033
41. Arora P (2024) Creative data justice: a decolonial and indigenous framework to assess creativity and artificial intelligence, Information, Communication & Society, vol. 27, no. 7, pp. 1321–1336. https://doi.org/10.1080/1369118X.2024.2309226
42. Veale M, Sanfilippo M, Lovat J (2023) AI and global governance: modalities, rationales, tensions. Annual Rev Law Social Sci 19:255–275. https://doi.org/10.1146/annurev-lawsocsci-020223-040749

43. Truby J (2020) Governing artificial intelligence to benefit the UN Sustainable Development Goals, Sustainable Development, vol. 28, no. 4, pp. 946–958. https://doi.org/10.1002/sd.2048
44. Radanliev P (2025) AI ethics: integrating transparency, fairness, and privacy in AI development. Appl Artif Intell 39(1). https://doi.org/10.1080/08839514.2025.2451060
45. Kubantseva EV (2026) Artificial intelligence as a resource of digital neocolonialism and its impact on digital sovereignty. Izv Saratov Univ Sociol Politology 26(1):63–70
46. Yurdem B, Akturk E, Kocaoğlu A, Dogan M Federated learning: overview, strategies, applications, tools and future directions, Heliyon, vol. 10, no. 19, 2024. https://doi.org/10.1016/j.heliyon.2024.e38991
47. Tan A, Yu B, Cui L, Yang Q (2022) Towards personalized federated learning, IEEE Transactions on Neural Networks and Learning Systems. 33(5):2091–2105. https://doi.org/10.1109/TNNLS.2021.3160699
48. Kotek H, Dockum R, Sun DQ (2023) Gender bias and stereotypes in large language models, Proceedings of the ACM Collective Intelligence Conference, pp. 12–24. https://doi.org/10.1145/3582269.3615599
49. Wen J, Zhang J, Lan Y, Cui Z, Cai J, Zhang J (2022) A survey on federated learning: challenges and applications. Int J Mach Learn Cybernet 13:3769–3793. https://doi.org/10.1007/s13042-022-01647-y
50. Rueda-Carvajal GD, Lopez-Lezama LA (2025) and others, Opportunities and challenges of Industries 4.0 and 5.0 in Latin America, Sci, vol. 7, no. 1. https://doi.org/10.3390/sci7010010

Figures

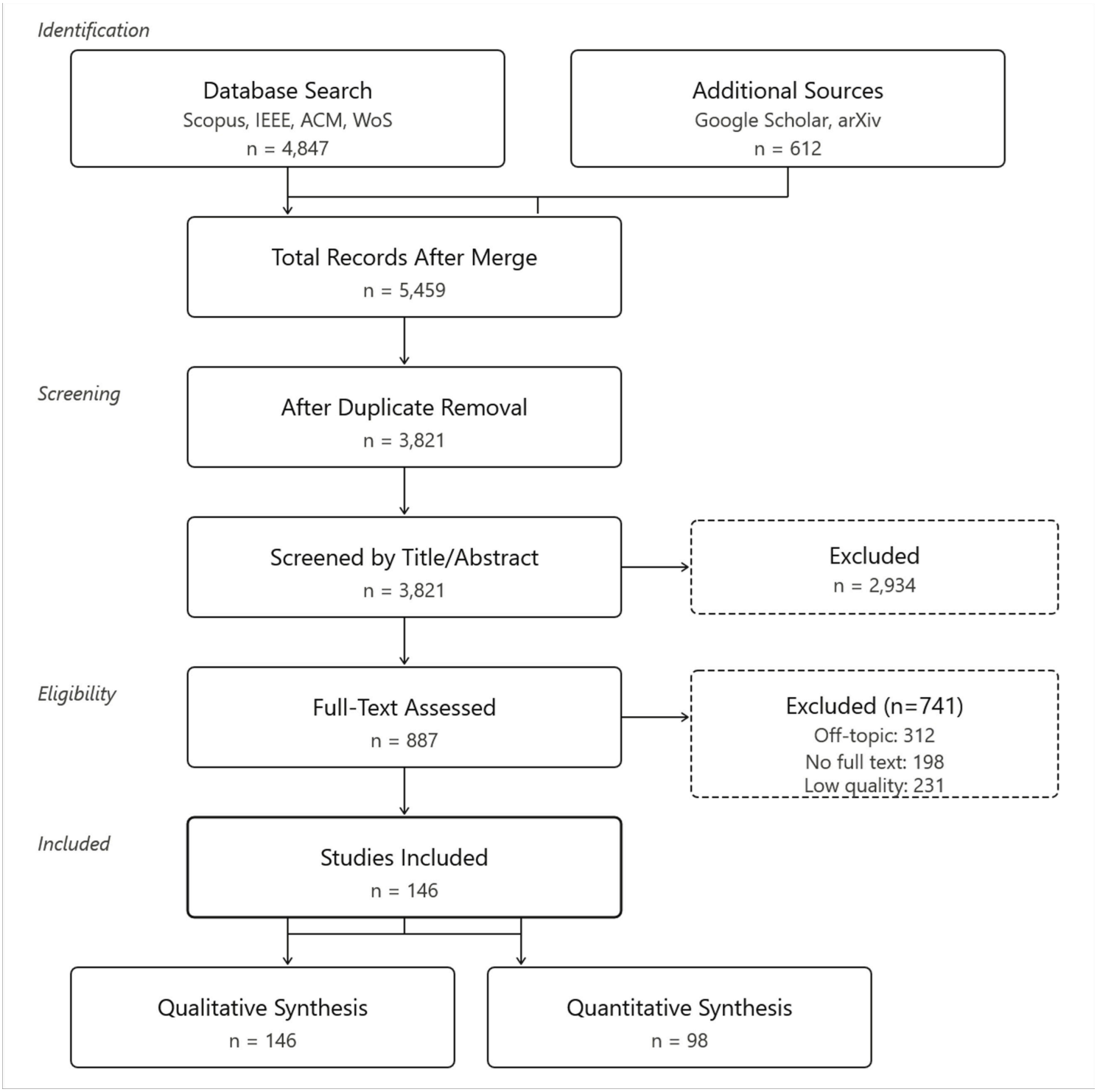


Figure 1

PRISMA flow diagram of the literature selection process

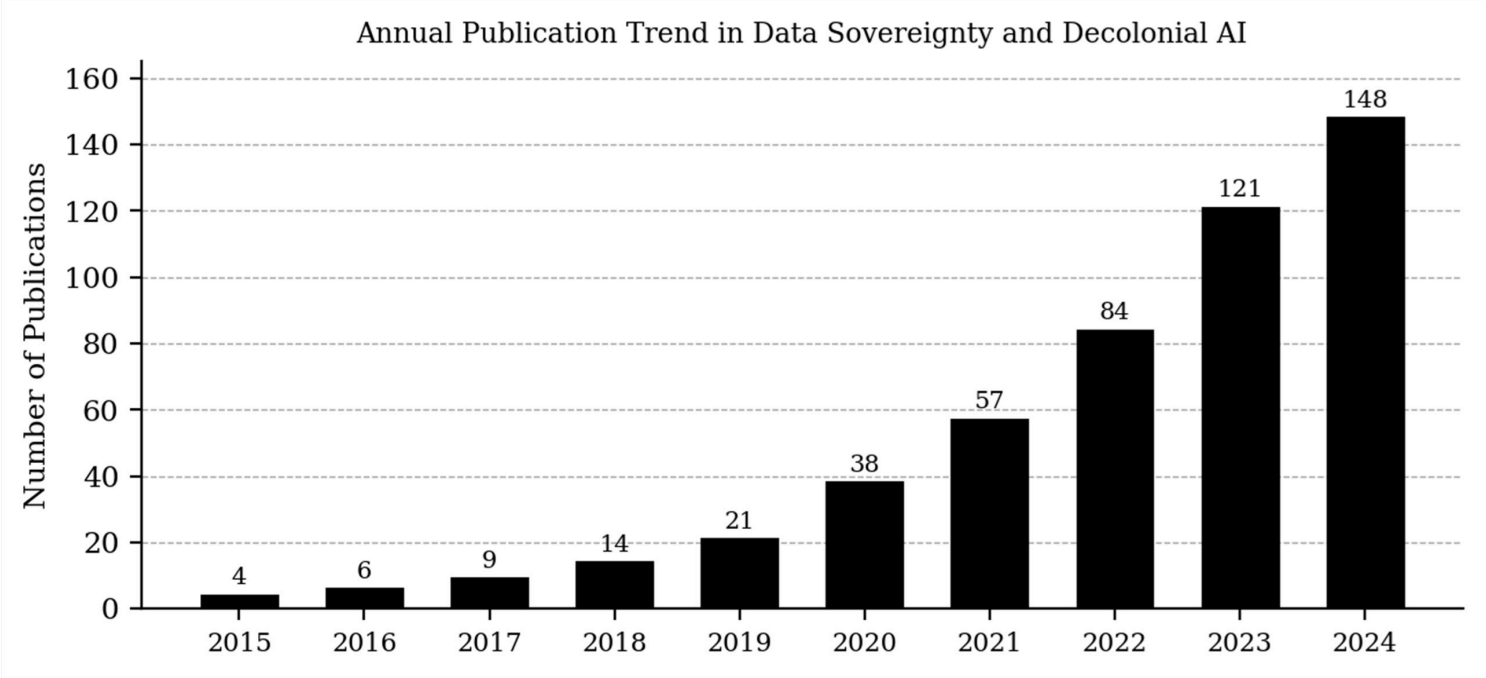


Figure 2

Annual publication trend in data sovereignty and decolonial AI research from 2015 to 2024.

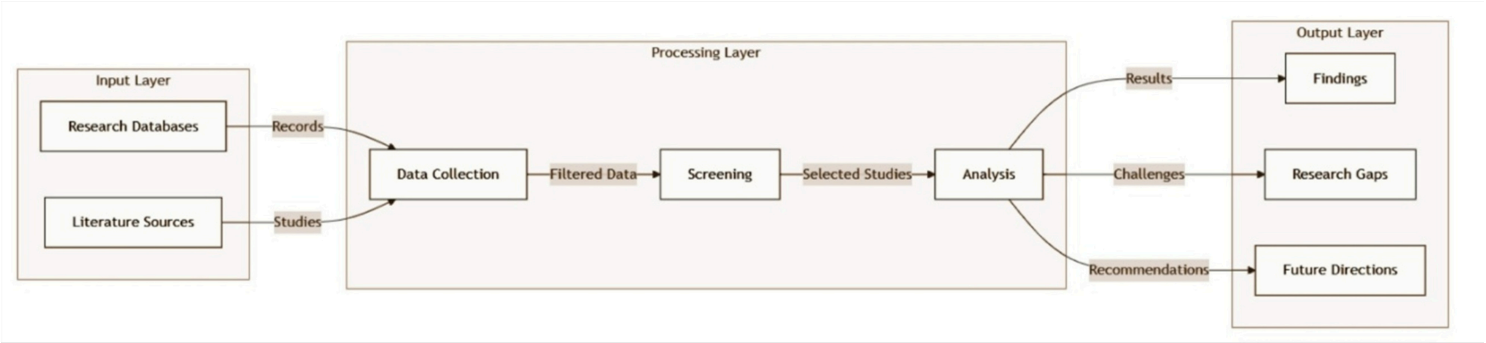


Figure 3

Block diagram of the foundation model governance architecture for the Global South.

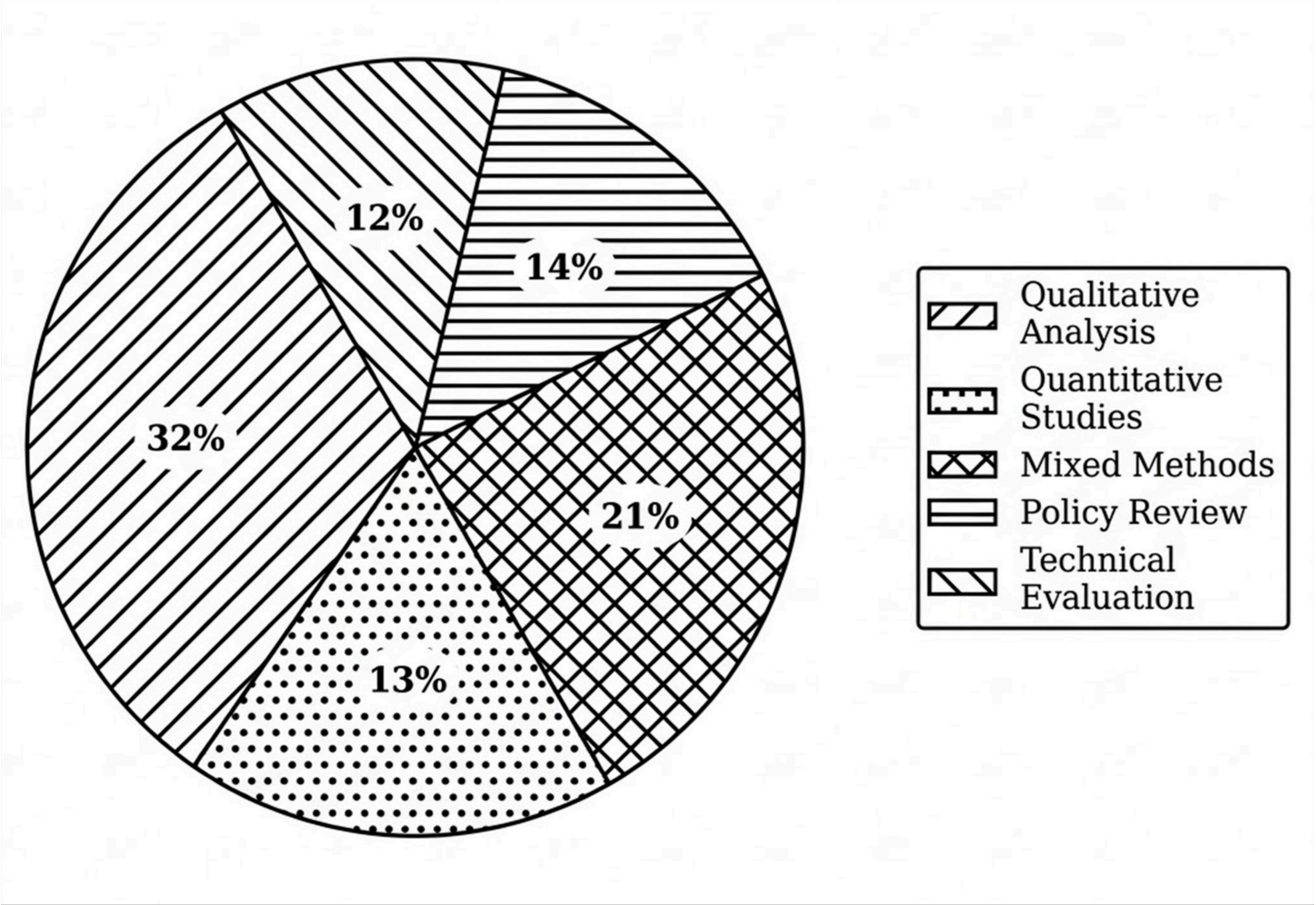


Figure 4

Distribution of research methods used across reviewed studies on data sovereignty and decolonial AI.

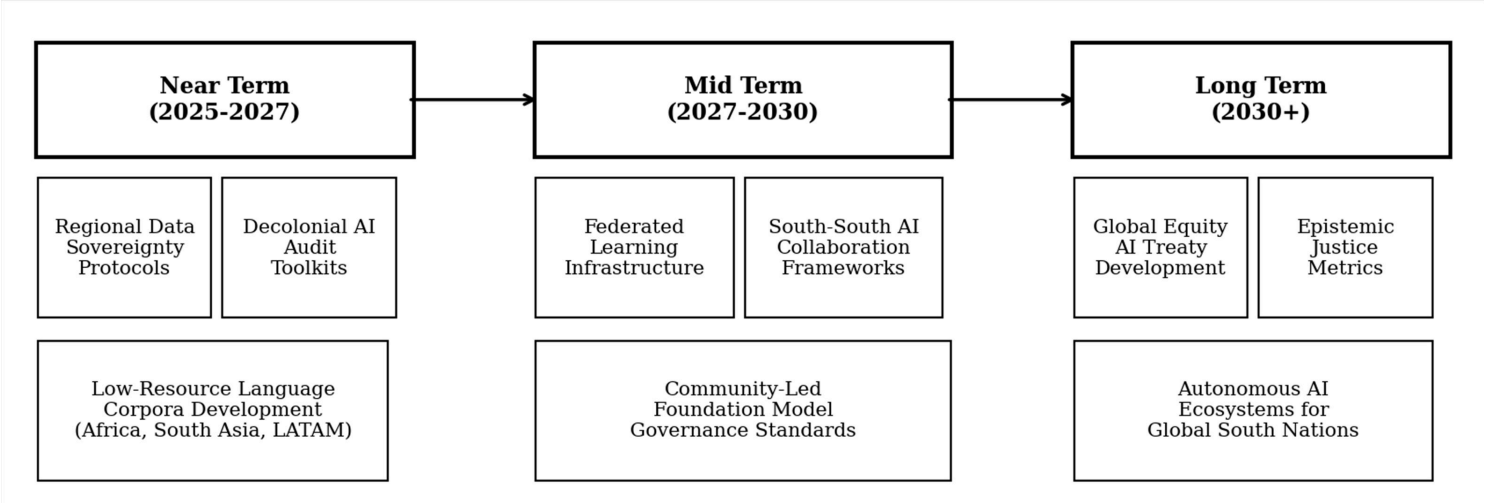


Figure 5

Proposed research roadmap for data sovereignty and decolonial AI (2025-2030 and beyond).