

## Small AI as Digital Public Infrastructure: A New Paradigm for Resource-Constrained Economies

Unyime Ufok Ibekwe <sup>1\*</sup>, Divine Ufok Ibekwe <sup>2</sup>

<sup>1</sup>Information Technology Department, Central Bank of Nigeria, Abuja, Nigeria

<sup>2</sup>Department of Computer and Information Sciences, Covenant University, Canaan Land, Ota, Ogun State, Nigeria

\*Corresponding author: [keenconsole@gmail.com](mailto:keenconsole@gmail.com)

ORCID: <https://orcid.org/0009-0000-3144-8201>

### ABSTRACT

Digital Public Infrastructure (DPI) is a key enabler of modern development policy, supporting the design and delivery of scalable public and private digital services. Current DPI discourse focuses primarily on three core components: digital identity systems, interoperable payment platforms and data exchange systems. However, recent advances in artificial intelligence (AI) create demand for an additional infrastructure layer that enables intelligent, adaptive and cost-effective service delivery, especially in resource-limited settings. This paper conceptualizes Small AI as a fourth layer of DPI capable of advancing inclusive and resilient development. Unlike frontier AI systems that depend on extensive computational resources, concentrated infrastructure and significant financial investment, Small AI systems are resource-efficient, locally adaptable and economically accessible. The paper positions Small AI not merely as a technological tool but as a form of public infrastructure that may generate broad societal benefits comparable to roads, electricity and telecommunications networks. Using a qualitative conceptual research design supported by comparative policy and institutional analysis, the study presents the Small AI Digital Public Infrastructure Framework (SAI-DPIF) organized into three functional layers: Small AI compute capacity as a shared public utility, Small AI model repositories as open knowledge commons and Small AI application ecosystems as inclusive multi-sector service delivery platforms. Together, these layers provide a conceptual basis for understanding how Small AI may contribute to institutional efficiency, financial inclusion, agricultural modernization and public service delivery. By providing a structured lens for examining these development pathways, the framework is intended to inform both policymaking and future research. To support practical adoption, the paper recommends two policy instruments: the Global Small Artificial Intelligence Infrastructure Fund (GSAIF) and the Small Artificial Intelligence Readiness Index (SAIRI) to facilitate investment decisions, implementation and cross-country assessment.

**KEYWORDS:** Artificial Intelligence, Edge AI, Digital Inequality, Digital Public Infrastructure, Inclusive development, Small AI, Sustainable development.

### I. INTRODUCTION

Over the past decade, Digital Public Infrastructure (DPI) has emerged as one of the most influential architectural blueprints in digital governance and development policy. During the 2023 G20 presidency, member states adopted a widely accepted definition of DPI as a set of open, secure and interoperable digital systems that support equitable access to public and private services for populations at scale (Sengupta et al., 2025; G20 Digital Economy Working Group, 2023a). This framework identifies three foundational pillars: digital identity, payment

systems and data exchange. It has subsequently been endorsed by major international institutions, including the United Nations, the World Bank and the International Telecommunication Union, as a scalable architecture for inclusive development (ITU, 2024; World Bank, 2023). The United Nations Secretary-General further identified DPI as one of twelve high-impact initiatives capable of accelerating progress toward the Sustainable Development Goals, with more than 100 countries committing to its advancement (United Nations Development Programme, 2023).

The developmental potential of DPI is illustrated by India's ecosystem of Aadhaar, the Unified Payments Interface (UPI) and the Data Empowerment and Protection Architecture (DEPA), which have expanded access to identity, payments and consent-based data sharing at scale (Parsheera, 2024; Thukral & Lakshit, 2025; ITU, 2024). According to the World Bank's Global Findex database, the proportion of Indian adults with bank accounts increased from 35% in 2011 to 80% by 2017, largely due to the Pradhan Mantri Jan-Dhan Yojana initiative (Demirgüç-Kunt et al., 2018). More broadly, countries with mature DPI systems tend to achieve stronger outcomes in public service delivery, social protection, digital commerce and financial inclusion. Yet despite these achievements, current DPI architectures remain largely focused on enabling access, interoperability and transactions. They provide limited support for intelligent decision-making, adaptive automation and Artificial Intelligence-enabled service delivery.

This limitation is becoming increasingly significant as Artificial Intelligence reshapes healthcare, agriculture, finance, education and public administration. However, the dominant trajectory of frontier Artificial Intelligence development is characterized by substantial infrastructure requirements, including hyperscale cloud platforms, advanced computing resources, energy-intensive data centers and highly specialized expertise (Kim et al., 2024; Lin et al., 2024). These requirements have created profound inequalities in access and capability, placing many low- and middle-income countries at risk of becoming consumers rather than producers of Artificial Intelligence innovation. Training a single frontier model can require thousands of megawatt-hours of electricity, often exceeding the capacity of fragile power systems. The International Energy Agency projects that electricity consumption from data centers could increase by up to 50% by 2030, largely driven by Artificial Intelligence adoption (IEA, 2025; Pilz et al., 2025). Similarly, UNESCO and UCL (2025) report that the annual energy consumption of generative Artificial Intelligence systems already rivals that of some low-income countries and continues to grow rapidly. These disparities are reinforced by the concentration of Artificial Intelligence resources in high-income economies, which account for only 17% of the world's population but host 77% of global data-center capacity, 87% of notable Artificial Intelligence models, 86% of Artificial Intelligence start-ups and 91% of venture capital investment (World Bank, 2025a; UNCTAD, 2025). Consequently, frontier Artificial Intelligence remains inaccessible as a broad-based development instrument for much of the Global South.

At the same time, an alternative paradigm is emerging. UNESCO and UCL (2025) highlight that small, task-specific Artificial Intelligence models can reduce energy consumption by up to 90% relative to large-scale models while maintaining comparable performance. Gartner (2025) predicts that by 2027 organizations will be three times more likely to deploy specialized small models than general-purpose large language models. Empirical studies further demonstrate that on-device Artificial Intelligence systems outperform larger architectures in training efficiency, inference speed and energy consumption (Leroux et al., 2022). The World Bank (2025b) similarly notes that affordable and accessible Small Artificial Intelligence applications are already generating measurable development benefits across emerging economies by enabling countries to bypass traditional infrastructure constraints.

Despite this potential, Small Artificial Intelligence remains weakly theorized within the Digital Public Infrastructure literature and is rarely conceptualized as infrastructure in its own right.

This paper attempts to address that gap by advancing a theory of Small Artificial Intelligence as a fourth layer of Digital Public Infrastructure, offering developing economies a distinct and more accessible pathway to Artificial Intelligence-enabled development. Conceptualizing Small Artificial Intelligence as infrastructure rather than as a collection of applications elevates it to the level of a strategic public asset requiring sustained investment, governance arrangements, technical standards and institutional capacity. To operationalize this vision, the paper introduces a Small Artificial Intelligence Digital Public Infrastructure Framework (SAI-DPIF) comprising three interconnected layers: Small AI compute capacity as a shared public utility, Small AI model repositories as open knowledge commons and Small AI application ecosystems as inclusive multi-sector service delivery platforms. The paper further examines the governance, financing and scalability mechanisms required to institutionalize Small Artificial Intelligence as public infrastructure across developing economies and concludes with policy recommendations for multilateral development institutions.

## II. REVIEW OF RELATED LITERATURE

### A. The Architecture and Logic of Digital Public Infrastructure

Digital Public Infrastructure (DPI) extends a long-standing principle in development economics that views infrastructure as a driver of productive transformation, economic growth and poverty reduction (Calderon & Serven, 2008; Estache, 2010). Like traditional public infrastructure, digital infrastructure generates population-scale positive externalities, exhibits elements of non-rivalry and often requires public or quasi-public governance to address market failures in provision (Stiglitz, 1999). DPI applies these principles to the digital economy by creating shared digital systems that enable inclusive access to services and broad participation in economic activity.

In practice, DPI is commonly organized around three core pillars: digital identity systems, digital payment systems and data exchange frameworks (Ibrahimov, 2025). Digital identity systems, exemplified by India's Aadhaar and the Modular Open-Source Identity Platform (MOSIP), which is now deployed across nine developing countries, provide foundational authentication for both public and private services (Alonso et al., 2023; Sengupta et al., 2025). Digital payment systems, including India's UPI, Brazil's Pix and Rwanda's eKash on Mojaloop, enable real-time and inclusive financial transactions at scale (Dixit et al., 2025; AfricaNenda et al., 2022). Complementing these systems, data exchange platforms such as India's Data Empowerment and Protection Architecture (DEPA) and Estonia's X-Road facilitate interoperable, consent-based data sharing across sectors (Bakhtina et al., 2023). Viewed through North's (1990) institutional theory, DPI functions as a set of formal digital rules that reduce transaction costs and improve the coordination of economic activity.

While DPI has attracted growing scholarly and policy attention, important critiques highlight the conditions under which it can deliver inclusive outcomes. Heeks (2022), through the concept of adverse digital incorporation, argues that digital infrastructure does not automatically promote inclusion and may instead reinforce existing inequalities, particularly among marginalized communities in the Global South. He distinguishes digital inclusion from digital empowerment, emphasizing that participation without meaningful benefits or agency can produce adverse outcomes. This perspective underscores the need to treat accessibility, language and usability as core design requirements rather than optional features. Similarly, Dencik et al. (2019) argue through a data justice framework that data infrastructures can reproduce existing inequalities unless principles of recognition, participation and

accountability are embedded in their design. For DPI, this implies that data exchange frameworks and consent mechanisms must address the vulnerabilities of low-literacy, low-connectivity and marginalized populations.

Rather than undermining DPI, these critiques clarify the governance and design principles necessary for inclusive digital infrastructure. They also reinforce the rationale for conceptualizing Small Artificial Intelligence (AI) as an additional DPI layer. The proposed framework addresses Heeks's (2022) concerns through offline functionality, local-language support and low-device requirements embedded within Layer 1. It further incorporates the data justice principles of Dencik et al. (2019) through consent-based governance, open licensing and community co-design mechanisms embedded within Layer 2. Yet an important limitation remains within existing DPI architectures: the absence of intelligence as infrastructure. Although current DPI systems effectively support access, interoperability and transactions at scale, they provide limited support for decision-making, adaptive automation and intelligent service delivery. It is this missing cognitive layer that Small AI seeks to provide.

## B. The Frontier AI Exclusion Problem

The exclusion of developing economies from frontier Artificial Intelligence is increasingly evident across multiple structural dimensions. A fundamental constraint is the unequal distribution of digital infrastructure. As of mid-2025, high-income countries accounted for approximately 77% of global co-location data-center capacity, whereas low-income countries represented less than 0.1% (World Bank, 2025a). This disparity is reinforced by the substantial energy requirements of frontier Artificial Intelligence systems. Strubell et al. (2019) provided one of the first systematic assessments of the energy and carbon costs of training large neural network models, finding that training BERT on a Graphics Processing Unit (GPU) can generate carbon emissions comparable to those of a trans-American flight. Subsequent studies have shown that these energy demands continue to rise as frontier models increase in scale.

Beyond infrastructure and energy constraints, frontier Artificial Intelligence requires levels of capital and technical expertise that remain inaccessible to many developing economies. The cost of large-scale Artificial Intelligence training facilities frequently ranges from hundreds of millions to more than one billion dollars (Sajjadi Mohammadabadi et al., 2025), while the specialized workforce needed to design, train and maintain frontier models is concentrated in a small number of countries. Together, these constraints create significant barriers to participation in the frontier Artificial Intelligence ecosystem.

The developmental implications of this exclusion are increasingly recognized. UNDP (2025b) warns that unmanaged Artificial Intelligence diffusion may deepen global inequalities by widening disparities in economic performance, human capabilities and governance systems. Similarly, Khan et al. (2024) observe that low-income countries remain largely marginalized in Artificial Intelligence research and deployment despite its potential to improve healthcare, education, energy systems and public-sector effectiveness. The World Bank's (2026) World Development Report further cautions that unequal access to the infrastructure, capital and skills required for Artificial Intelligence may widen development gaps between high- and lower-income countries.

These findings point to a broader structural exclusion problem rooted in deficits in compute infrastructure, energy, capital, connectivity and technical capacity. Rather than treating these constraints as temporary barriers, the SAI-DPIF recognizes them as defining characteristics of many resource-constrained environments. It therefore seeks to identify and promote a class of Artificial Intelligence systems whose technical and economic requirements are compatible with the realities of developing economies.

**C. The Small AI Inflection Point and the Theoretical Gap**

Recent developments in Artificial Intelligence (AI) economics and deployment patterns strengthen the case for Small Artificial Intelligence as a developmentally appropriate alternative to frontier systems. Gartner (2025) predicts that by 2027 organizations will use small, task-specific Artificial Intelligence models three times more frequently than general-purpose large language models. This shift is driven by growing evidence that smaller models can deliver substantial efficiency gains without significant losses in performance. UNESCO and UCL (2025) report that matching customized small models to specific tasks can reduce energy consumption by up to 90%, while empirical studies show that small language models can achieve between 70% and 95% of GPT-class performance on targeted domain-specific tasks while requiring far fewer computational resources (Sanh et al., 2019; Corradini et al., 2025).

These characteristics are particularly relevant for developing economies. According to the World Bank (2025b), many Small Artificial Intelligence applications can operate on everyday devices without extensive digital infrastructure, creating opportunities for technological leapfrogging in resource-constrained environments. As a result, Small Artificial Intelligence offers a pathway to expand access to intelligent services without the substantial infrastructure, energy and capital requirements associated with frontier models.

Despite these converging technological and economic trends, the academic literature has not yet conceptualized Small Artificial Intelligence as a form of infrastructure. While Zhang and Zhang (2025) examine large language models as quasi-public goods, their analysis does not extend to a Digital Public Infrastructure-layer architecture or to policy frameworks tailored to developing economies. This omission represents an important theoretical gap. The present paper addresses this gap by conceptualizing Small Artificial Intelligence as a Digital Public Infrastructure layer capable of supporting inclusive and scalable intelligence services in resource-constrained settings. Table 1 presents a comparative analysis of Frontier Artificial Intelligence and Small Artificial Intelligence systems.

**Table 1:** Frontier AI Systems vs. Small AI as Digital Public Infrastructure a Comparative Analysis

Dimension	Frontier AI Systems	Small AI as Digital Public Infrastructure
<b>Compute Requirements</b>	Thousands of high-end GPUs; centralized hyperscale data centers; specialized cooling and power infrastructure (Kaur, 2025; Sajjadi Mohammadabadi et al., 2025)	Low-power CPUs/NPUs; edge-optimized processors; national shared compute utilities, deployable at district and provincial scale (Wang et al., 2025; Corradini et al., 2025, Sanh et al., 2019)
<b>Energy Consumption</b>	Training a single frontier model can consume thousands of MWh, generative AI's annual energy footprint already equivalent to that of a low-income country (IEA, 2025; Kim et al., 2024; Lin et al., 2024; Pilz et al., 2025; Corradini et al., 2025 )	Small AI models tailored to specific tasks can reduce energy use by up to 90% without sacrificing performance, on-device inference uses up to 90% less energy than cloud processing for the same task ((UNESCO & UCL, 2025; Wang et al., 2025; CSIS, 2025; Leroux et al. 2022; Corradini et al., 2025 )
<b>Capital Cost</b>	High cost per training cluster; per-token API pricing excludes low-resource users and developing-country governments from AI capability (Sajjadi Mohammadabadi et al., 2025; Cottier, et al. (2024)	National shared compute utility: open-source inference at marginal electricity cost; SLMs achieve 70–95% of GPT-class performance at 1/10th the cost (Wang et al., 2025; Corradini et al., 2025; ByteLota (2026)

Dimension	Frontier AI Systems	Small AI as Digital Public Infrastructure
<b>Connectivity Dependency</b>	High-bandwidth cloud connectivity essential; significant latency in low-connectivity settings makes real-time services unreliable and cost-prohibitive for rural populations (Sajjadi Mohammadabadi et al., 2025)	Fully deployable offline or on low-bandwidth networks; on-device inference eliminates cloud dependency; compatible with feature-phone interfaces (Wang et al., 2025)
<b>Data Sovereignty</b>	User data processed in foreign jurisdictions under foreign law; structural dependency on external platforms limits national data governance and accountability (CSIS, 2025)	Local data processing preserves national sovereignty; compatible with consent-based data frameworks (e.g., DEPA); enables data localization mandates without sacrificing AI capability (Aayog, 2020; Wilkinson et al., 2025; CSIS, 2025)
<b>Provision Model</b>	Private, proprietary; market driven; governed by corporate terms of service subject to unilateral change; no democratic accountability or development mandate (World Bank, 2025a)	Public utility; open knowledge commons; interoperable application ecosystem; multi-stakeholder governance aligned with G20 DPI principles and UN Digital Public Goods frameworks (Aayog, 2020; Government of India, Press Information Bureau, 2024, 2025 & 2026; Khan et al., 2024; Mrisho et al., 2020, IndiaAI Mission, 2025)
<b>Language &amp; Context</b>	Predominantly trained on high-resource languages; poor performance on low-resource languages; limited local cultural and regulatory calibration (Kaur, 2025)	Efficiently fine-tuned on low-resource languages using LoRA and quantization; context-calibrated for local development conditions; India's Bhashini covers 22 official languages (Kaur, 2025; World Bank, 2025a; Khan et al., 2024)
<b>Development Alignment</b>	Compute infrastructure is concentrated in HICs, structurally excludes most developing economies; (World Bank, 2025a; Sajjadi Mohammadabadi et al., 2025; Khan et al., 2024)	Designed for resource-constrained, high-need contexts; Small AI as leapfrogging vehicle for LICs; (Khan et al., 2024, World Bank, 2025a)

*Source: Authors' construction. LIC = Low-Income Country; HIC = High-Income Country.*

While Table 1 highlights important differences between Frontier Artificial Intelligence and Small Artificial Intelligence, the comparison should be interpreted in light of several considerations. First, Frontier Artificial Intelligence and Small Artificial Intelligence are not competing technologies but complementary components of the broader Artificial Intelligence ecosystem. Many Small Artificial Intelligence systems depend on frontier models for pre-training or knowledge distillation. For example, DistilBERT was derived from BERT through knowledge distillation (Sanh et al., 2019), and many task-specific models build on representations learned from larger systems. Second, some public-sector applications in developing economies, such as epidemiological modelling, satellite analysis and complex legal document processing, may continue to require the capabilities of larger models, particularly as digital infrastructure and connectivity improve. Third, the advantages associated with Small Artificial Intelligence depend on adequate data quality, institutional capacity and governance, conditions that remain uneven across emerging market and developing economies.

Accordingly, Table 1 should be viewed as an assessment of how each technology class aligns with the prevailing institutional and infrastructural realities of low- and middle-income countries rather than as a definitive measure of developmental value. As the World Bank (2025a) notes, Frontier Artificial Intelligence and affordable Small Artificial Intelligence applications are likely to coexist and interact, with their relative roles shaped by specific use cases, infrastructure endowments and governance maturity..

#### D. Small AI as Public Infrastructure

Small AI refers to lightweight, resource-efficient and task-specific systems designed for affordability, local deployment and operation within constrained computational environments. Unlike Frontier Artificial Intelligence, which typically relies on centralized cloud infrastructure, extensive computing resources, proprietary access models and continuous internet connectivity, Small Artificial Intelligence can operate effectively on low-cost devices, edge systems and low-bandwidth networks while maintaining contextual relevance (Sajjadi Mohammadabadi et al., 2025; Wang et al., 2025; Corradini et al., 2025). These characteristics make Small Artificial Intelligence particularly well suited to environments where governments must deliver digital services at scale despite limited infrastructure. This aligns with the World Bank's growing emphasis on Artificial Intelligence solutions that remain affordable, practical and effective under conditions of constrained connectivity and computing capacity. It also reflects increasing recognition that Frontier Artificial Intelligence, while technologically advanced, is often poorly matched to the realities of developing economies because of its high inference costs, concentrated ownership structures and limited interoperability (CSIS, 2025). Viewed through this lens, Small AI is more than a technological tool. Its affordability, accessibility and deployability enable it to function as a foundational digital resource that supports large-scale service delivery, economic participation and inclusive digital transformation. Consequently, Small AI can be conceptualized as a form of public digital infrastructure rather than merely an application deployed within individual sectors.

##### D.1. Conditions for Classifying Small AI as Digital Public Infrastructure

A key question is whether Small AI should be conceptualized as infrastructure or merely as a category of digital technology. Addressing this requires clarifying the defining features of Digital Public Infrastructure (DPI). Drawing from public infrastructure economics and emerging DPI scholarship, four core conditions are typically identified: non-rivalry, broad accessibility (low exclusion), positive externalities and enabling platform functionality (Frischmann, 2009; Agénor, 2010; World Bank, 2024).

First, infrastructure is generally *non-rivalrous*, meaning one user's consumption does not significantly reduce availability for others. Small AI largely meets this condition because, once models are developed and deployed, additional users can access services at very low marginal cost. For example, an agricultural advisory model can serve millions of farmers simultaneously without proportional increases in development costs.

Second, DPI is characterized by *broad accessibility* and low exclusion costs. While perfect non-excludability is rare in digital systems, DPI seeks to maximize inclusion through open standards, interoperability and affordable access. Small AI advances this objective because it can run on low-cost devices, function offline or with intermittent connectivity and be distributed through open-source or public-interest deployment models. This substantially lowers access barriers compared with frontier AI systems that depend on expensive cloud infrastructure.

Third, infrastructure generates *positive externalities* that extend beyond direct users. Small AI exhibits these spillover effects by improving decision-making, productivity and

service quality across sectors. In agriculture, for instance, advisory systems not only improve farm-level outcomes but also strengthen food security, enhance knowledge diffusion and improve efficiency across value chains. Comparable spillovers are evident in healthcare, education and financial inclusion applications.

Fourth, infrastructure serves as an *enabling platform* on which other services and economic activities can be built. Roads enable transport systems, electricity enables industrial production and telecommunications enable digital communication. In the same way, Small AI provides a foundational layer of computational intelligence that can be reused across multiple applications and sectors. Governments, firms and civil society organizations can therefore build services on shared Small AI systems without independently developing full AI capabilities.

Collectively, these characteristics indicate that Small AI satisfies the core conditions of Digital Public Infrastructure. Its value lies not only in individual applications but in its role as a reusable and scalable foundation for intelligent service delivery across the economy.

## D.2. Distinguishing Small AI Infrastructure from Digital Platforms and Public Digital Services

It is necessary to distinguish Small AI infrastructure from digital platforms and public digital services. Digital platforms typically operate at the application layer by delivering end-user services within defined markets or sectors, such as e-commerce platforms, social media networks and government service portals. In a similar way, public digital services, including agricultural extension systems, telemedicine applications and digital learning platforms, are designed to address specific service delivery needs at the point of use.

Small AI, as conceptualized in this paper, operates at a lower level of abstraction. The proposed *Small AI Compute Capacity layer* and *Small AI Model Repository layer* are not end-user services. Instead, they constitute shared infrastructure in the form of compute resources, foundational models, development toolkits and interoperability standards that enable the creation of diverse downstream applications. Their function is comparable to other DPI layers such as cloud infrastructure, telecommunications networks and digital identity systems, which provide reusable foundations rather than direct services.

This distinction is important because the value of Small AI infrastructure does not reside in any single application. Rather, it emerges from its ability to support an ecosystem of intelligent services across agriculture, healthcare, education, finance and public administration through shared compute and model resources. Small AI should therefore be understood not as an embedded feature within applications, but as a foundational Digital Public Infrastructure layer that reduces the cost of innovation while expanding access to AI-enabled services in resource-constrained settings.

From this perspective, Small AI represents an emerging form of public-interest Digital Public Infrastructure. Like other infrastructure systems, it requires substantial upfront investment but generates sustained social and economic returns across multiple sectors. By integrating low-cost compute capacity, shared model repositories and interoperable ecosystems, Small AI enables inclusive, resilient and scalable digital transformation. Its infrastructural nature is defined not by individual applications but by its capacity to serve as a common foundation for a wide range of public and private services across developing economies.

### III. Theoretical Framework: Small AI as the Fourth DPI Layer

#### A. Public Goods Theory and the Case for Small AI Infrastructure

Samuelson's (1954) classical theory of public goods defines goods through two core properties: non-rivalry, where one person's use does not reduce availability for others, and non-excludability, where restricting access is difficult or costly. Infrastructure often occupies an intermediate position, since it is largely non-rival until congestion effects emerge and partially excludable depending on institutional design. Building on this, Stiglitz (1999) argues that infrastructure is frequently treated as a public good because wide access generates positive externalities that exceed private incentives. Similarly, Ostrom (1990) demonstrates that shared knowledge resources can be sustainably governed through polycentric institutions and collective stewardship. Extending these ideas, Frischmann (2006) conceptualizes infrastructure as generating "input value" by enabling downstream activity, producing social value beyond private returns and being systematically underprovided due to market failure. At the global scale, Kaul et al. (1999) further emphasize that knowledge systems, standards and digital infrastructure constitute global public goods that require multilateral governance to avoid under-provision, a logic directly applicable to Small AI model repositories and shared compute systems.

Within this framework, Small AI infrastructure can be understood as a quasi-public good. Once developed, open-source Small AI models can be replicated and distributed at near-zero marginal cost, reflecting digital non-rivalry. Through open licensing and interoperable design, access can also be made broadly non-excludable. In addition, these systems generate widespread positive externalities across healthcare, agriculture, education, financial inclusion and fraud detection, which private actors are unlikely to fully internalize. As a result, purely market-driven systems tend to underinvest in Small AI for low-resource languages, underserved populations and resource-constrained environments, where social returns exceed private incentives. Supporting this view, Zhang and Zhang (2025) model large language models as quasi-public goods and show that open artificial intelligence systems generate knowledge spillovers that markets do not fully internalize. Building on this, the present study argues that Small AI's infrastructural role is defined not only by technical properties but also by the institutional arrangements governing its deployment as a shared resource.

Three qualifications refine this public goods interpretation. First, infrastructure status depends on both technical features and governance structures (Frischmann, 2006). A proprietary Small AI system may exhibit limited public goods characteristics, whereas the same system under open-access governance may function more fully as infrastructure. The normative implications of this framework therefore depend on the governance architecture outlined in Section VI. Second, not all artificial intelligence applications qualify as public goods; many are subject to congestion, require continuous data and maintenance inputs, or generate primarily private rather than social value. Accordingly, the public goods argument applies most strongly to foundational layers such as compute infrastructure and model repositories rather than application-layer services. Third, the global public goods perspective of Kaul et al. (1999) implies that Small AI infrastructure requires international coordination and multilateral financing rather than purely national provision, thereby motivating the Global Small AI Infrastructure Fund proposed in Section VII.

#### B. Platform Economics and the Infrastructural Logic of Small AI

Platform economics provides a second theoretical foundation for conceptualizing Small AI as Digital Public Infrastructure. In this view, infrastructure generates value not only through its technical provision but through the ecosystems of users, applications and services it enables (Rochet & Tirole, 2003; Eisenmann et al., 2006). Traditional infrastructure illustrates this logic

clearly: roads facilitate mobility and commerce, electricity enables industrial production and the Unified Payments Interface (UPI) demonstrates how interoperable digital rails can unlock extensive financial technology ecosystems. In the same way, Small AI infrastructure which comprises shared compute systems, model repositories, interoperability standards and deployment frameworks derives its value from downstream applications in healthcare diagnostics, agricultural advisory services, education technologies, citizen-facing platforms and fraud detection.

Rochet and Tirole (2003) further show that platforms generate value through network externalities by enabling interactions between distinct user groups. In the Small AI–Digital Public Infrastructure context, these groups include model developers on one side and application developers and end users on the other. They also emphasize that platform governance must prevent monopolization and ensure broad, non-discriminatory access. Extending this argument, Parker et al. (2016) distinguish between pipeline systems, which create value through linear production, and platform systems, which create value by enabling interactions across an ecosystem. Their analysis shows that open access to core infrastructural resources is essential for innovation, while proprietary control over foundational layers constrains the diversity and scale of complementary applications. This reinforces the case for treating Small AI as public infrastructure rather than as a proprietary platform.

Building on large-scale comparative evidence, Evans and Gawer (2016) demonstrate that the most transformative platforms are those that enable complementary innovation through open, interoperable and accessible core systems. Their taxonomy of transaction, innovation and investment platforms maps closely onto the Small AI Digital Public Infrastructure architecture (SAI-DPIF). The compute utility layer functions as a transaction platform enabling Artificial Intelligence inference, the model repository layer operates as an innovation platform supporting application development and the Global Small AI Infrastructure Fund (GSAIF) serves as an investment platform that channels resources into public goods provision.

This platform logic implies that developing economies should prioritize shared foundational infrastructure over fragmented Artificial Intelligence applications. The experience of India’s Digital Public Infrastructure ecosystem illustrates this principle. Rather than promoting isolated payment applications, interoperable Unified Payments Interface rails were developed as shared infrastructure, enabling hundreds of fintech innovations to emerge on top of a common system (G20 Task Force, 2023). Parker et al. (2016) describe such arrangements as innovation platforms whose primary value lies in enabling complementary innovation rather than delivering standalone services. Extending this reasoning to the cognitive layer of Digital Public Infrastructure, the Small AI–Digital Public Infrastructure framework (SA-DPIF) argues that the main return on investment in compute utilities and model repositories is not the direct services they provide, but the wide range of locally adapted, context-specific and sector-diverse Artificial Intelligence applications they enable governments, firms and communities to build.

### **C. Development Economics: Leapfrogging and Hirschmanian Linkages**

Development economics has long emphasized that late-developing economies can leapfrog traditional infrastructure pathways by adopting advanced technologies without replicating earlier stages of industrial development (Abramovitz, 1986; Perez, 2002). A contemporary illustration is the rapid expansion of mobile money in Sub-Saharan Africa, where digital financial systems substituted for conventional banking infrastructure and significantly improved financial inclusion (Demirgüç-Kunt et al., 2018). Extending this logic, Khan et al. (2024) provide preliminary evidence that artificial intelligence catch-up in low-income countries is feasible through leapfrogging and absorptive capacity mechanisms, while the

World Bank (2025a) similarly argues that affordable Small AI systems can enable countries to bypass traditional infrastructural constraints. Within this framing, Small AI offers developing economies an opportunity to deploy efficient, locally governed and resource-efficient systems without first reproducing the capital-intensive computational infrastructures of advanced economies.

Hirschman's (1958) theory of backward and forward linkages further strengthens the case for Small AI as infrastructure. Investment in Small AI compute capacity can generate backward linkages by stimulating demand for complementary investments in electricity supply, connectivity infrastructure and hardware ecosystems, thereby supporting employment creation and broader productive capacity. At the same time, it produces forward linkages across sectors such as healthcare, agriculture, education and public administration by enabling intelligent applications that enhance productivity and service delivery. These interdependencies imply that the developmental returns of Small AI extend well beyond its direct costs. Shared compute systems and open model infrastructures can support multiple downstream applications simultaneously, generating economy-wide multiplier effects that isolated, project-based artificial intelligence deployments are unlikely to achieve. Together, this linkage structure supports the interpretation of Small AI not as a narrow technology procurement effort, but as a strategic public infrastructure investment that warrants coordinated national and multilateral financing.

#### IV. METHODOLOGY

##### A. Research Design and Epistemological Position

This study adopts a qualitative conceptual research design, which is appropriate where the objective is theory development rather than empirical hypothesis testing (Jaakkola, 2020; Rocco & Plakhotnik, 2009). It is grounded in an interpretivist epistemology that treats infrastructural categories as socially and institutionally constructed rather than fixed technical realities. Within this perspective, conceptual frameworks are evaluated based on theoretical coherence, adaptation, topology and framework, the four criteria identified by Jaakkola (2020). Following Jaakkola's theory-building approach, the study proceeds in four steps: it identifies a significant gap in the literature, develops an original framework through theoretical synthesis and analogical reasoning from institutional cases, refines the framework using comparative policy analysis across contexts and evaluates it against the three quality criteria. Assessment in this study is therefore limited to conceptual validity rather than empirical testing. As a result, the framework has not yet been subjected to expert interviews, Delphi validation or pilot implementation. This limitation defines a key methodological boundary and informs the future research agenda outlined in Section VIII.C, particularly the proposed engagement with Digital Public Infrastructure practitioners, national artificial intelligence mission leaders and multilateral development institution experts for subsequent validation and refinement.

While Jaakkola (2020) and Patton's (2002) purposive sampling logic provide the primary methodological foundation, two complementary traditions further inform the design. Eisenhardt's (1989) case study theory-building approach underpins the use of comparative institutional cases to generate analytically generalisable rather than statistically representative insights. This supports the inclusion of cases such as India's IndiaAI Mission, Kenya's AI Hub, Rwanda's Kigali Innovation City and the MOSIP and Mojaloop open-source ecosystems. In line with Eisenhardt (1989), theoretical sampling is used to extend and refine emerging propositions rather than maximize representativeness, with maximum variation employed to capture differences in governance, financing and scalability across contexts. In addition, Charmaz's (2014) constructivist grounded theory approach informs the iterative coding and constant comparison process used in the literature mapping. Although this is not a grounded

theory study, Charmaz's emphasis on theoretical sensitivity, reflexivity and iterative movement between data and concepts informs the thematic synthesis that underpins the framework's three-layer architecture.

### **B. Systematic Literature Mapping**

This study employed a structured dual-corpus literature mapping across two interrelated domains: Digital Public Infrastructure and Small AI. The Digital Public Infrastructure corpus included terms such as “digital public infrastructure,” “India Stack,” “digital identity systems,” “interoperable payments” and “MOSIP,” while the Small AI corpus included “small language models,” “edge AI,” “on-device AI,” “efficient AI inference” and “Small Language Models Global South.” Sources were retrieved from Scopus, Web of Science, IEEE Xplore, ACM Digital Library, Google Scholar, World Bank Open Knowledge, International Telecommunication Union databases, G20 policy documents, United Nations Development Programme, United Nations Educational, Scientific and Cultural Organization, International Energy Agency, Carnegie Endowment and Center for Strategic and International Studies between 2018 and 2025, supplemented by foundational theoretical works regardless of publication date.

In total, 97 sources were analyzed, comprising 55 peer-reviewed journal articles, 19 institutional reports, 8 government and policy documents and 15 technical reports from artificial intelligence research organizations and open-source foundations. Analysis followed Braun and Clarke's (2006) reflexive thematic analysis approach using an iterative inductive coding process. Initial codes derived from both corpora were refined through three iterative rounds and consolidated into five higher-order themes, with analysis continuing until thematic saturation was achieved. These themes were: (i) infrastructural characteristics of Small AI, (ii) market-failure justifications for public provision, (iii) governance and accountability requirements, (iv) financing mechanisms and policy precedents and (v) scalability conditions and institutional transferability. Inclusion criteria were limited to English-language publications with clear theoretical or empirical relevance to artificial intelligence infrastructure, Digital Public Infrastructure or development economics. Exclusion criteria removed purely technical artificial intelligence studies lacking governance or development relevance, as well as institutional documents without verifiable empirical grounding.

### **C. Comparative Policy Analysis**

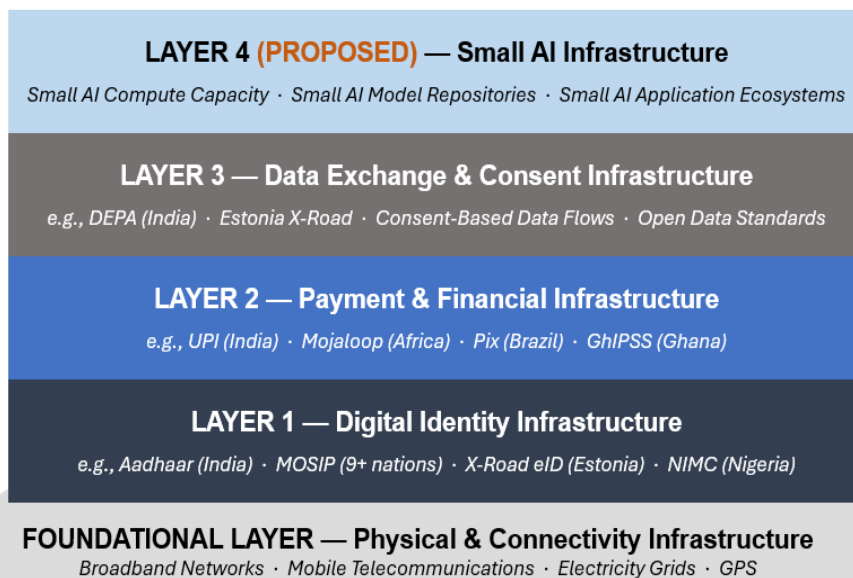
Country selection followed maximum variation sampling (Patton, 2002) to ensure diversity across five dimensions: income level (low to high), Digital Public Infrastructure maturity (nascent to advanced), artificial intelligence investment capacity (minimal to globally leading), geographic coverage (Africa, Asia, Europe, Latin America and multilateral institutions) and governance model (state-led, market-led and open-source). This design produced five analytical clusters that reflect the institutional contexts in which Small AI–Digital Public Infrastructure systems may emerge and operate. Each cluster serves a distinct analytical purpose. Digital Public Infrastructure leaders provide governance and design templates, artificial intelligence-producing economies illustrate technological trajectories, Global South adopters demonstrate implementation feasibility under constraints, multilateral initiatives offer financing and coordination precedents, and open-source ecosystems highlight alternative provisioning and innovation models. To ensure cross-case comparability, all clusters were analyzed using a structured four-dimension protocol covering: (i) definitional frameworks for artificial intelligence and Digital Public Infrastructure, (ii) provisioning models (public, private or hybrid), (iii) financing mechanisms and sustainability structures and (iv) governance, interoperability and scalability arrangements.

**D. Institutional Synthesis and Framework Assessment**

The final methodological component involved iterative framework synthesis and assessment. The Small AI–Digital Public Infrastructure Framework (SAI–DPIF) was developed through three iterative rounds of theoretical synthesis using constant comparison across theory and comparative policy findings (Charmaz, 2014). In the first round, public goods theory, platform economics, development linkage theory and leapfrogging literature were mapped to Small AI characteristics to establish its theoretical basis as infrastructure. In the second round, a three-layer architecture was developed by aligning governance, financing and scalability patterns identified in the comparative policy analysis. In the third round, the framework was evaluated against conceptual rigor and practical relevance criteria (Jaakkola, 2020) and refined accordingly. The final framework was then assessed for internal coherence, ensuring that governance structures across all layers remained consistent with the public goods and platform economics principles underpinning the framework.

**V. THE SMALL AI-DPI FRAMEWORK (SAI-DPIF)**

Before presenting the framework architecture, it is important to clarify its epistemological status. The Small AI–Digital Public Infrastructure Framework (SAI–DPIF) is a conceptual proposition developed through theoretical synthesis, analogical reasoning from Digital Public Infrastructure precedents and comparative institutional analysis, as described in Section IV. It is not an empirically validated causal model. The developmental outcomes associated with its layers, including productivity gains, financial inclusion and improved service delivery, are theoretically derived inferences informed by documented deployments in specific contexts rather than generalized empirical evidence. Consequently, the enabling functions attributed to Small AI are conditional theoretical constructs, with empirical validation identified as a future research priority (Section VIII.C). Within this conceptual framework, Figure 1 positions the Small AI layer within the extended Digital Public Infrastructure stack, showing how it builds on identity, payments and data exchange layers while introducing a cognitive layer that transforms Digital Public Infrastructure from a transactional system into an intelligence-enabled development platform. Figure 2 presents the detailed architecture.



*Source: Authors' own construction*

**Figure 1:** The Extended DPI Stack - Small AI as the Fourth Infrastructure Layer

### A. Layer 1: Small AI Compute Capacity as a Shared Public Utility

The foundational layer of the SAI–DPIF is *shared compute capacity*, comprising the physical and virtual infrastructure required for artificial intelligence model training, fine-tuning and inference. In contrast to frontier artificial intelligence systems, which depend on hyperscale data centres consuming hundreds of megawatts of electricity and requiring billions in capital expenditure (Strubell et al., 2019), Small AI compute infrastructure is designed to be distributed, modular and energy-efficient. This layer includes low-power graphics processing unit and neural processing unit clusters, inference-optimized processors, national shared compute facilities with open-access mandates, regional compute pools coordinated through multilateral arrangements and edge-computing nodes embedded in healthcare facilities, agricultural extension centres, schools and government offices. Its core principle is accessibility, meaning it must support both inference and fine-tuning while remaining financially and operationally viable for public institutions in resource-constrained settings.

The public utility framing of this layer draws directly from the theoretical foundations outlined in Section III. Applying Frischmann’s (2009) three-condition infrastructure framework, compute capacity (i) generates input value by enabling downstream applications rather than functioning as a final good, (ii) produces social returns in sectors such as healthcare, agriculture, financial inclusion and governance that exceed what any single private actor can capture and (iii) is underprovided by markets due to insufficient commercial incentives, particularly for low-resource language systems and rural deployment contexts. This market failure justifies public or quasi-public governance of foundational compute infrastructure, consistent with historical precedents such as electricity grids and telecommunications networks, which were initially treated as public infrastructure before private-sector expansion at the application layer (Parker et al., 2016). The analogy to electrification is particularly instructive. Rural electricity access required public investment in generation and grid infrastructure before private firms could scale appliance ecosystems; similarly, Small AI requires public investment in compute foundations before local developers, health workers and agricultural services can build viable applications.

Beyond physical access, Heeks’s (2022) concept of adverse incorporation highlights additional constraints related to cognitive and cultural accessibility. Infrastructure may exist but remain effectively inaccessible due to literacy barriers, interface complexity or connectivity assumptions that exclude rural and peri-urban users. Accordingly, Layer 1 embeds core design requirements for offline inference, voice-based interfaces for low-literacy contexts and local language support across major deployment regions, rather than treating them as optional enhancements.

Empirical illustrations demonstrate the feasibility of this approach. India’s IndiaAI Mission shows national-scale implementation through subsidized shared graphics processing unit infrastructure and open-access compute facilities (Government of India, 2024; Government of India, 2026). In Kenya, the PlantVillage Nuru system, developed as a public-good initiative by PlantVillage at Pennsylvania State University in collaboration with the Food and Agriculture Organization of the United Nations, the International Institute of Tropical Agriculture and the International Maize and Wheat Improvement Center, enables offline crop disease diagnosis using basic smartphones, illustrating how distributed Small AI compute can function without hyperscale infrastructure (Mrisho et al., 2020). Similarly, Farmer Lifeline Technologies has developed a solar-powered pest detection device using computer vision to identify infestations within seconds and deliver alerts via short message service, reducing crop losses by up to 30% and demonstrating how ultra-low-power compute can enable artificial intelligence inference for smallholder farmers in Sub-Saharan Africa (Royal Academy of Engineering, 2024).

## B. Layer 2: Small AI Model Repositories as Open Knowledge Commons

The second layer of the SAI–DPIF consists of *digital repositories* where small artificial intelligence models are stored, documented, versioned and made accessible for both public and private use. Conceptually, this layer is grounded in Ostrom’s theory of commons governance, which explains how shared resources can be sustainably managed through collective-choice arrangements, graduated sanctions and polycentric governance structures operating across institutional levels (Ostrom, 1990; Ostrom, 2010). Within this framing, small AI model repositories function as governed knowledge ecosystems for development-relevant domains such as health, agriculture, education, finance, governance and disaster management.

These repositories host models trained on locally relevant datasets and languages, enabling developers to build applications without retraining foundational models. Similar to open-source software repositories and digital public goods registries, they are sustained through public funding, community contributions, peer review, benchmarking and open licensing regimes that ensure reuse and accessibility. Core supporting components include model cards for transparency (Mitchell et al., 2019), dataset provenance documentation (Geburu et al., 2021), privacy-preserving dataset repositories, benchmark suites, deployment recipes for offline and edge environments and contributor-governed adaptation registries. Collectively, these mechanisms convert fragmented open-source activity into a structured public good aligned with Digital Public Goods principles (Digital Public Goods Alliance, 2023).

Empirical implementations of this model are already emerging at national scale. India’s Bhashini initiative, launched in 2022 under the National Language Translation Mission, develops speech and language artificial intelligence across all 22 scheduled Indian languages and serves as a foundational Digital Public Infrastructure for language technologies (Ministry of Electronics and Information Technology, 2022; Digital India BHASHINI Division, 2026). Similarly, AIKosh, launched in 2025 under the IndiaAI Mission, functions as a national repository for artificial intelligence models and datasets (Government of India, Press Information Bureau, 2025; IndiaAI Mission, 2025). These initiatives illustrate how model repositories can be institutionalized as core components of public digital infrastructure. In parallel, UNESCO and UCL (2025) illustrate that mixture-of-experts architectures, which combine multiple specialized small models into a unified system, offer more resource-efficient deployment than single general-purpose models, reinforcing the logic of modular, repository-based design.

A central function of this layer is localisation and language inclusion. Frontier artificial intelligence systems are predominantly trained in a limited number of global languages, excluding large populations from meaningful digital participation. Small AI systems, by contrast, can be fine-tuned on local language datasets and deployed through accessible interfaces such as voice applications, messaging platforms and local governance systems. India’s Bhashini infrastructure already enables integration with WhatsApp-based civic chatbots such as Jugalbandi and governance platforms such as SabhaSaar, which are used by thousands of village councils to transcribe and translate Gram Sabha proceedings into native languages, shifting artificial intelligence from an urban-centric innovation to a practical public service (Reboot Democracy AI, 2026; Ministry of Electronics and Information Technology, 2022). Similarly, Microsoft Research’s Project Gecko develops speech and small language models for six underrepresented Kenyan languages namely Swahili, Kikuyu, Kalenjin, Dholuo, Maa and Somali trained on 3,000 hours of community-sourced audio datasets, with models optimized to run efficiently on low-cost devices in rural areas (Microsoft Research, 2025).

Within the SAI–DPIF, Layer 2 therefore functions as an open knowledge common that democratizes access to artificial intelligence capability, expands linguistic inclusion and

enables applications aligned with Sustainable Development Goals across health, food security, education, governance and finance.

### C. Layer 3: Small AI Application Ecosystems as Interoperable Service Delivery Platforms

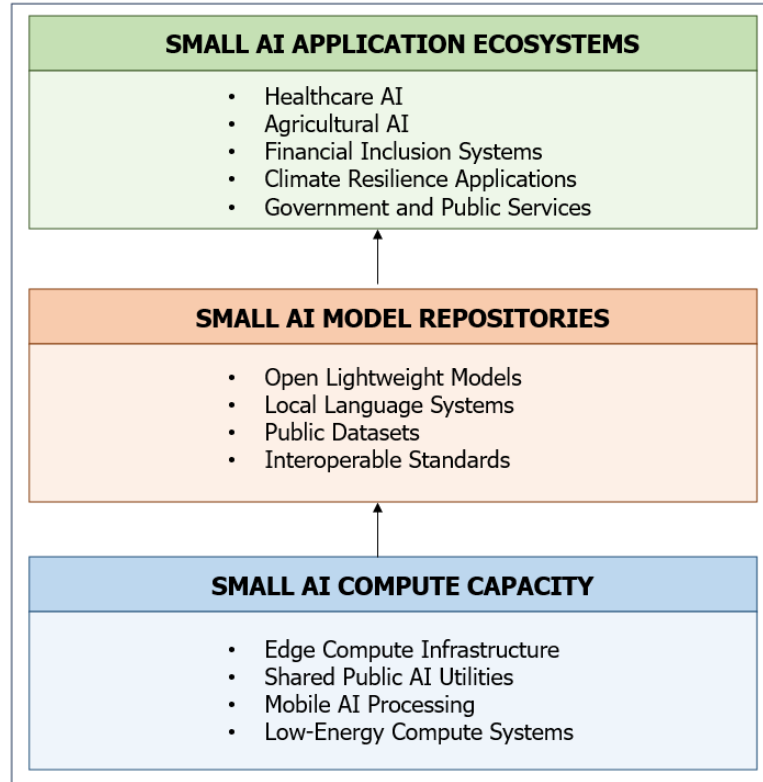
The third layer of the SAI–DPIF consists of *interoperable application ecosystems* built on shared compute infrastructure and open model repositories. Its defining design principle is inclusive, multi-sector service delivery, with applications designed to reach underserved populations across healthcare, agriculture, education, governance and public finance through interfaces suited to low-connectivity and low-literacy environments. Interoperability is enabled through shared standards, application programming interfaces, data exchange protocols and authentication systems established at the model repository layer and inherited by this application layer. This inherited interoperability prevents vendor lock-in and allows a single infrastructure base to support multiple services simultaneously, distinguishing Digital Public Infrastructure ecosystems from proprietary digital platforms. The effectiveness of this approach is illustrated by India’s Unified Payments Interface, where a single interoperable infrastructure layer enabled a large ecosystem of applications and service providers to emerge (G20 Task Force on Digital Public Infrastructure, 2024). In the same way, Small AI–DPI enables multiple intelligent public service applications to operate across sectors on shared infrastructure.

At this layer, the value of Small AI infrastructure is realized through practical service delivery applications that operate across systems and share data and functionality. In healthcare, locally trained small models can assist community health workers with offline diagnosis and decision support. In agriculture, context-specific models can provide crop disease detection, weather forecasts and market price information through voice and short message service interfaces in local languages. In education, adaptive tutoring systems in local languages can extend teacher capacity in underserved schools, while governance applications such as translation tools and citizen engagement platforms can improve access to public services for linguistic minorities. Collectively, these applications address exclusion gaps identified in artificial intelligence and development research (Hazari & Mohan, 2024; Khan et al., 2024) and respond to concerns raised by the United Nations Development Programme (2025b) regarding unequal artificial intelligence diffusion in the absence of inclusive infrastructure design. To enable such cross-sector service delivery, this layer operationalizes interoperability standards inherited from Layer 2, including identity verification, data schemas, inference application programming interfaces, consent management systems, notification services, payment integration and offline synchronization protocols.

The developmental potential of this layer is already visible in practice. Open-weight models and interoperable standards enable governments and local developers to adapt systems to local contexts without continuous dependence on external vendors for updates or retraining. In Zambia, the DawaMom platform developed by Dawa Health provides WhatsApp and short message service-based maternal health guidance for rural mothers using multilingual artificial intelligence systems trained on local health data and national protocols, supporting languages such as Bemba, Lozi, Nyanja, Tonga, Shona and Ndebele (Dawa Health, 2026). The platform has reached more than 5,000 mothers through over 10,000 chatbot interactions, with approximately 84% of users receiving point-of-care tests having no prior antenatal blood tests, indicating successful outreach to previously underserved populations (MIT Solve, 2023).

Within the broader SAI–DPI framework, Layer 3 functions as an inclusive, multi-sector service delivery ecosystem characterized by strong positive externalities and network effects. Governed through sectoral regulatory frameworks, interoperability standards bodies and

Digital Public Infrastructure certification mechanisms, it can be financed through public–private partnerships, impact investment, blended finance and GovTech procurement systems, thereby accelerating financial inclusion, agricultural productivity, public service delivery and progress toward Sustainable Development Goals at scale.



Source: Author's conceptualization

Figure 2: The Small AI-DPI Framework

Figure 2 presents the tripartite architecture of the Small AI–DPI Framework, structuring the three interdependent layers described in Sections V(A), V(B) and V(C) into a unified infrastructure stack. At the base is *Small AI Compute Capacity*, which provides the physical and virtual infrastructure for artificial intelligence processing at national and regional scale through edge-computing nodes, shared public artificial intelligence utilities, mobile processing systems and energy-efficient compute resources. The middle layer, *Small AI Model Repositories*, operates as an open knowledge commons where lightweight models, local language systems, public datasets and interoperability standards are stored, governed and made accessible to developers and institutions. At the top, *Small AI Application Ecosystems* constitute the service delivery layer, where compute capacity and model repositories converge to enable practical applications in healthcare, agriculture, financial inclusion, climate resilience and governance. The relationships between these layers are explicitly interdependent. Application ecosystems depend on governed model repositories for reusable intelligence components, while model repositories in turn rely on foundational compute capacity for training, fine-tuning and deployment. This layered dependency structure ensures that no single component functions in isolation. Together, the three layers form an integrated Digital Public Infrastructure system in which each layer reinforces the others, enabling inclusive, sovereign

and scalable artificial intelligence-enabled development across low- and middle-income countries.

## VI. GOVERNANCE, FINANCING AND SCALABILITY OF SMALL AI-DPI

### A. Multi-Tier Governance Architecture

Effective governance of Small AI Digital Public Infrastructure (DPI) requires simultaneous coordination across open access, quality assurance, algorithmic accountability, data sovereignty and financial sustainability. Unlike physical infrastructure, digital infrastructure operates across jurisdictional boundaries and introduces complex risks related to privacy, algorithmic harm and platform dependency, necessitating coordination among public and private actors. Drawing on North (1990) and comparative policy analysis, this study proposes a four-tier governance architecture spanning global, regional, national and community levels (Table 2).

At the *global tier*, institutions such as the International Telecommunication Union, United Nations Development Programme, United Nations Educational, Scientific and Cultural Organization, World Bank, International Monetary Fund and G20-aligned coalitions coordinate standards, maintain a Small AI Digital Public Goods registry, administer the Global Small AI Infrastructure Fund (GSAIF) and publish the Small AI Readiness Index (SAIRI). The International Telecommunication Union has already developed over 100 artificial intelligence-related standards, with more than 120 in development as of May 2024 (ITU, 2024). Complementing this, UNESCO's Recommendation on the Ethics of Artificial Intelligence provides the first global normative framework for AI governance (UNESCO, 2021), while the Global Digital Compact strengthens coordination by calling for interoperable standards that ensure safety, reliability, sustainability and human rights compliance (United Nations, 2024). Together, these initiatives reflect an emerging multilateral architecture aimed at preventing exclusion from the benefits of artificial intelligence.

At the *regional tier*, bodies such as the African Union, Association of Southeast Asian Nations, Caribbean Community, Pacific Islands Forum and European Union harmonize interoperability standards, coordinate data governance frameworks and pool compute resources through shared regional infrastructure. This pooling approach enables countries with limited individual capacity to access scalable compute infrastructure through public-private partnerships and shared cloud systems, thereby reducing costs and expanding access for researchers, startups and governments. This regional logic is particularly important given that many developing economies lack the scale required for independent investment in advanced digital infrastructure (Khan et al., 2024).

At the *national tier*, artificial intelligence compute authorities, analogous to telecommunications regulators would license and regulate Small AI utilities under open-access mandates, curate national model repositories and certify application ecosystems for interoperability. India's IndiaAI Mission provides a practical governance template. Approved in March 2024 with an allocation exceeding USD 1.2 billion, it treats compute, data and foundational models as shared national infrastructure rather than concentrated private assets, establishing the IndiaAI Compute Portal and AIKosh repository as open-access resources for startups, micro, small and medium enterprises and public institutions (Government of India, Press Information Bureau, 2025; IndiaAI Mission, 2025).

At the *community tier*, local governments, non-governmental organizations, farmer cooperatives, community health workers and academic institutions co-design context-specific applications, contribute local data for model adaptation and ensure last-mile service delivery. UNESCO and the Innovation for Policy Foundation (2022) emphasize that inclusive

policymaking requires multistakeholder participation to harness collective intelligence in both design and implementation. This bottom-up layer ensures that Small AI applications reflect local languages, livelihoods and institutional realities that cannot be fully specified by global or national actors.

Across all tiers, data governance must address two interrelated challenges. The first is data sovereignty, ensuring that locally generated data remains under national jurisdiction while remaining compatible with consent-based frameworks embedded in DPI Layer 3 (Wilkinson et al., 2025; NYU Center on International Cooperation, 2024). The second is algorithmic accountability, ensuring that deployed models are transparent, auditable and non-discriminatory. The Global Digital Compact supports interoperable governance frameworks while affirming national regulatory autonomy over privacy, security and human rights (United Nations, 2024), rejecting one-size-fits-all global models. In this sense, democratic oversight and public accountability remain essential safeguards for Digital Public Infrastructure.

Operationalizing algorithmic accountability requires model documentation standards, bias auditing and accessible appeal mechanisms adapted to developing-country institutional capacities. Bias audits conducted by independent actors including researchers, journalists and civil society organizations assess systems by testing outputs against diverse inputs, while model cards document training assumptions, validation procedures and performance across demographic groups (Mitchell et al., 2019). Complementary transparency tools such as artificial intelligence registries and lifecycle audits from design to deployment and monitoring enable civil society scrutiny and informed public oversight. Collectively, these mechanisms strengthen transparency and accountability across Small AI systems deployed in public service delivery.

## B. Financing Mechanisms

This paper proposes three complementary financing mechanisms to support Small AI as Digital Public Infrastructure across developing economies. The first is a Global Small AI Infrastructure Fund (GSAIF), modelled on the Green Climate Fund and capitalized at an initial USD 3-5 billion from Organization for Economic Co-operation and Development (OECD) member governments, multilateral development banks and private foundations (authors' proposal). The fund would support national Small AI compute utilities and regional model repository infrastructure. This proposal aligns with the Center for Strategic and International Studies (2025), which calls for coordinated investment from the United States private sector and multilateral organizations into artificial intelligence infrastructure and innovation projects in the Global South, reinforcing the case for a dedicated multilateral financing vehicle.

The second mechanism is Digital Public Infrastructure-linked sovereign bonds, enabling developing economies to raise commercial capital for artificial intelligence infrastructure by leveraging demonstrated returns from existing Digital Public Infrastructure systems as a credit enhancement base. In India, systems such as Aadhaar, Unified Payments Interface and FASTag generated an estimated USD 31.8 billion in 2022, contributing 0.9 per cent to gross domestic product, with projections of 2.9 to 4.2 per cent by 2030 as adoption expands (Nasscom & Arthur, 2024). These documented returns provide a fiscal foundation for sovereign bond instruments backed by Digital Public Infrastructure performance and scalability.

The third mechanism is public-private co-investment frameworks that mobilize private capital for application-layer development while maintaining public ownership of compute and model repository layers. This approach mirrors the layered structure that enabled rapid ecosystem expansion in India's Digital Public Infrastructure model. As documented by the G20 Task Force on Digital Public Infrastructure (2024), the India Stack combines digital identity,

real-time payments through Unified Payments Interface and secure data-sharing infrastructure to reduce transaction costs, maintain interoperability and stimulate private-sector innovation at the application layer, while preserving public governance of foundational infrastructure. The G20 Task Force further identifies this layered architecture as a defining feature of the Digital Public Infrastructure model, demonstrating how public infrastructure ownership can coexist with vibrant private innovation ecosystems.

### C. Scalability Pathways and the Four Cs

The scalability of Small AI–DPI is driven by three enabling conditions that closely align with the World Bank’s “Four Cs” framework: connectivity, compute, context and competency (World Bank, 2025). The first is *technical modularity*, which allows edge inference nodes to be deployed incrementally at low marginal cost, while open-source models can be continuously fine-tuned and added to repositories without disrupting core infrastructure. Similarly, small language model architectures are inherently composable, enabling flexible expansion of capabilities. The World Bank also emphasizes that affordable Small AI applications can help developing economies leapfrog traditional infrastructure constraints, with lightweight tools already demonstrating impact in areas such as clinical data analysis and small business expansion.

The second enabling condition is *institutional portability*, illustrated by the cross-country deployment of the Modular Open-Source Identity Platform (MOSIP). The Carnegie Endowment for International Peace (2025) notes that MOSIP, an Aadhaar-informed open-source identity system, is now being adopted across multiple countries, with nine of eleven implementing countries located in Africa. This demonstrates that Digital Public Infrastructure governance models can be transferred across diverse institutional contexts without requiring identical preconditions.

The third enabling condition is *ecosystem network effects*, where system value increases as participation expands across models, applications and users, generating self-reinforcing growth dynamics that can attract private investment at the application layer without sustained public subsidy. Rochet and Tirole’s (2003) theory of two-sided markets explains that platforms must simultaneously attract participants on both sides, with cross-side network externalities producing increasing returns to scale. In the context of Small AI, this implies that once a critical mass is reached, open infrastructure ecosystems can sustain private innovation. Regional cooperation frameworks such as the African Union Digital Transformation Strategy, ASEAN artificial intelligence governance frameworks and Caribbean Community digital economy initiatives further reinforce this dynamic by enabling compute pooling and cross-border model sharing, thereby lowering costs and expanding participation in Small AI ecosystems (African Union Commission, 2020; ASEAN Secretariat, 2024; CARICOM Secretariat, 2024).

Building on these foundations, scaling Small AI from pilot systems to population-level infrastructure requires coordinated expansion across four dimensions: geographic reach, sectoral scope, user depth and institutional integration. Geographic scaling follows a hub-and-spoke model extending from urban centres to district and village nodes and ultimately to edge devices, supported by universal service obligations and last-mile connectivity. Sectoral scaling progresses across domains such as agriculture, health, education and finance, driven by domain-specific data readiness and interoperability across applications. User depth expands through stages of awareness, registration, initial use, routine engagement and advanced utilization, supported by digital literacy, trust-building and user-centred design. Institutional integration advances from pilots to departmental adoption, ministry-wide implementation and ultimately whole-of-government deployment. Across all dimensions, human capital development, local innovation ecosystems, mobile-first deployment strategies and regional

coordination are critical. Energy-efficient and edge-optimized system design further enhances resilience, ensuring service continuity in resource-constrained environments and supporting scalable and sustainable deployment across developing economies.

**Table 2:** Multi-Tier Governance Architecture for Small AI-DPI

Governance Tier	Key Actors	Core Functions
<b>Global Tier</b>	UN agencies (ITU, UNDP, UNESCO); G20; World Bank; IMF; Bilateral donor coalitions	Set open standards; coordinate multilateral funding (GSAIF); maintain AI DPG registry; run SAIRI benchmarking
<b>Regional Tier</b>	African Union; ASEAN; EU; CARICOM; Pacific Island Forum; Regional Development Banks	Harmonize cross-border interoperability; pool compute resources; coordinate regional data sovereignty frameworks
<b>National Tier</b>	National AI authorities; DPI regulators; Central banks; Sector ministries; Civil society	License and regulate Small AI utilities; enforce open-access mandates; procure and deploy application ecosystems
<b>Community Tier</b>	Local governments; NGOs; Community health workers; Farmer cooperatives; Academic institutions	Co-design context-specific solutions; provide local data for fine-tuning; ensure last-mile service delivery

*Source: Authors' own construction*

## VII. POLICY RECOMMENDATIONS AND DEVELOPMENT IMPLICATIONS

### A. Establish the Global Small AI Infrastructure Fund (GSAIF)

The most urgent institutional recommendation of this paper is the creation of a *Global Small AI Infrastructure Fund (GSAIF)*, a multi-donor concessional trust fund capitalized at an initial USD 3–5 billion to finance national Small AI compute utilities and regional model repository infrastructure across developing economies. Existing multilateral financing mechanisms offer relevant governance precedents. The Green Climate Fund (GCF), established under the United Nations Framework Convention on Climate Change and the Paris Agreement, operates through a 24-member Board with equal representation from developed and developing countries, while deploying grants, concessional loans, guarantees and equity through accredited implementing entities. Similarly, the Global Fund for HIV/AIDS, Tuberculosis and Malaria use a country-owned, results-based financing model supported by governments, foundations and private sector contributions. Together, these institutions demonstrate that global public goods financing mechanisms can be sustained when anchored in shared governance structures, broad coalitions and clearly defined development objectives.

Experience from these institutions also highlights key feasibility and implementation constraints relevant to GSAIF. First, capitalization requires sustained political commitment; the GCF, for example, required approximately five years between establishment in 2010 and initial capitalization in 2015, suggesting that a realistic GSAIF timeline should anticipate a five- to seven-year formation period (Green Climate Fund, 2023). Second, absorptive capacity constraints often limit disbursement more than funding availability, as many recipient countries lack sufficient project preparation pipelines. Accordingly, GSAIF should integrate dedicated technical assistance for project design, implementation and institutional capacity building alongside financing. Third, governance legitimacy depends on meaningful country ownership; evidence from the GCF indicates that representation alone is insufficient unless recipient countries also shape funding priorities (Ciplet & Roberts, 2017). For this reason, GSAIF

disbursement mechanisms should adopt a country-driven model similar to that of the Global Fund, ensuring alignment with nationally defined priorities rather than donor preferences.

Additional safeguards are required given the distinct risks associated with artificial intelligence infrastructure. In particular, GSAIF must address vendor capture, where procurement processes become dominated by firms from high-income countries, and model dependency, where recipient states become locked into proprietary compute or model architectures. To mitigate these risks, GSAIF should prioritize open standards, open-source solutions and interoperability requirements across compute and model repository layers, while excluding proprietary closed-source systems from eligibility. Governance should be jointly shared among recipient countries, multilateral institutions, civil society and technical experts, with financing targeted primarily toward low-income countries through grants and highly concessional instruments. Accountability can be strengthened through performance-based allocations linked to the Small AI Readiness Index. In parallel, a Small AI Digital Public Goods Council, jointly convened by the United Nations Development Programme and the International Telecommunication Union, could coordinate global standards, open licensing frameworks and algorithmic accountability norms, building on their existing mandates in digital development coordination.

### **B. Adopt the Small AI Readiness Index (SAIRI)**

This paper also recommends the annual publication of a *Small AI Readiness Index (SAIRI)*, a composite indicator designed to measure developing economies' preparedness to build, govern and benefit from Small AI-Digital Public Infrastructure. Its feasibility is supported by established precedents in global development measurement. On the technological side, the International Telecommunication Union's ICT Development Index tracked and compared information and communication technology progress across countries from 2009 to 2018 before being relaunched with a revised methodology in 2023. On the governance side, the United Nations Development Programme Artificial Intelligence Landscape Assessment, launched in 2024 and implemented in more than fifteen countries by 2025, evaluates national readiness for responsible artificial intelligence development using a multi-stakeholder framework. SAIRI builds on these foundations by providing a dedicated benchmark for Small AI-DPI readiness, extending the World Bank's Four Cs framework of connectivity, compute, context and competency with a fifth dimension capturing governance readiness, including regulatory maturity, data sovereignty and alignment with international digital governance norms.

However, the implementation of composite indices presents well-known challenges. The Human Development Index has been criticized for aggregation bias, cross-country data inconsistencies and the risk of masking uneven performance across dimensions (Ravallion, 2012). Similarly, the temporary suspension of the ICT Development Index between 2018 and 2023 underscores the importance of stable data infrastructure and harmonized measurement systems (ITU, 2023). To address these limitations, SAIRI should report both aggregate scores and disaggregated dimension-level indicators, ensuring that strong performance in one area does not obscure weaknesses in others. It should also rely on harmonized international datasets, independent verification mechanisms and standardized reporting protocols rather than self-reported national data alone, thereby strengthening credibility and comparability.

Operationally, SAIRI would serve both as a diagnostic and a governance instrument, informing resource allocation under the proposed Global Small AI Infrastructure Fund while guiding national policy priorities and investment strategies. Countries would be assessed across five interconnected dimensions: compute capacity, open model ecosystem development, connectivity, governance readiness and skills and institutional capacity (Table 3). These

dimensions are intended to function as an integrated system rather than independent indicators. For instance, strong compute capacity without corresponding governance readiness may generate accountability risks, while strong skills without an open model ecosystem may constrain innovation. By linking infrastructure readiness to financing and reform priorities, SAIRI provides a practical mechanism for coordinating investment across the full Small AI development stack rather than serving as a purely descriptive measurement tool.

**Table 3:** Small AI Readiness Index (SAIRI) Dimensional Framework

SAIRI Dimension	Indicator Description	Measurement Approach	Policy Lever
<b>1. Compute Capacity</b>	Availability of national/regional shared AI compute; GPU-equivalents per capita; proportion at subsidized or open-access rates	ITU compute infrastructure surveys; national AI mission reports; energy intensity benchmarks (kWh per inference unit)	GSAIF capitalization; national compute utility investment; MDB concessional lending
<b>2. Open Model Ecosystem</b>	Local-language development-relevant open-source models available; quality, task coverage, and model card completeness	Digital public goods registries; model performance benchmarks on local-language NLP tasks; language diversity coverage	Open-source model funding; government open-licensing mandates; Development Dividend Framework
<b>3. Connectivity Adequacy</b>	Fixed and mobile broadband penetration; offline AI deployment feasibility; effective bandwidth for inference workloads	ITU/World Bank 4Cs connectivity data; offline inference feasibility testing; A4AI affordability index	Broadband investment; spectrum policy reform; last-mile connectivity programs
<b>4. Governance Readiness</b>	AI regulatory maturity; data protection legislation; algorithmic accountability; alignment with Global Digital Compact	OECD/ITU AI governance assessment; national legal review; data protection law coverage; G20 DPI compliance	National AI legislation; DPI governance architecture; open licensing mandates for publicly funded AI
<b>5. Skills and Institutional Capacity</b>	AI workforce depth; public sector AI literacy; university AI density; research output; government AI adoption rates	UNDP HDI; World Bank DPTR skills data; Stanford AI Index national scores; government digitization scorecards	AI education investment; university-industry partnerships; public sector reskilling programs

*Source:* Authors' own construction. SAIRI extends the World Bank (2025a) Four Cs framework (connectivity, compute, context, competency) with a governance dimension aligned with the UN Global Digital Compact (2024). NLP = Natural Language Processing; HDI = Human Development Index; DPTR = Digital Progress and Trends Report.

### C. Institutional Strategies for SAI-DPIF Adoption

Multilateral development institutions including the World Bank, International Monetary Fund, African Development Bank, Asian Development Bank and Inter-American Development Bank should formally recognize Small AI infrastructure as a distinct category of digital infrastructure eligible for concessional financing within existing digital economy lending windows. This aligns with two converging mandates. The Global Digital Compact, adopted on 22 September 2024, calls for increased investment in digital public goods and Digital Public Infrastructure, particularly in developing countries, through multi-stakeholder partnerships spanning governments, private sector actors and technical and academic communities to advance Sustainable Development Goals (United Nations, 2024). In parallel, the United Nations Conference on Trade and Development (2025) World Investment Report urges multilateral development banks to strengthen digital infrastructure through blended

finance, deepen global rules for digital investment and support innovation ecosystems ahead of the Fourth International Conference on Financing for Development. Within this context, priority actions include capitalizing the Global Small AI Infrastructure Fund (GSAIF), integrating Small AI compute and model repository investments into existing lending instruments and establishing AI knowledge-sharing platforms to enable South–South cooperation and reduce duplication across development programmes.

At the national level, governments in developing economies should integrate Small AI infrastructure into national digital economy strategies by treating it as critical infrastructure with dedicated budget allocations and institutional mandates. This includes establishing regulatory frameworks for Small AI compute utilities with open-access obligations and fair pricing, enforcing open licensing for publicly funded models under permissive regimes or Creative Commons and adopting interoperability standards to prevent vendor lock-in across application ecosystems. Governments should also formally designate Small AI systems as critical national infrastructure aligned with the World Bank Four Cs framework, with particular emphasis on compute access and competency development through national or regional facilities and large-scale artificial intelligence literacy programmes. India’s IndiaAI Mission provides an operational model that can be adapted across contexts, combining shared compute infrastructure with open model repositories. Across all strategies, governance frameworks must ensure transparency, auditability, human oversight and grievance redress mechanisms to maintain accountability and public trust in artificial intelligence-enabled services.

Open-source artificial intelligence communities, including Hugging Face, EleutherAI, Mozilla Foundation and Linux Foundation AI, should adopt a Development Dividend Framework that commits a defined share of capacity to training, fine-tuning and maintaining Small AI models for development-relevant tasks in low-resource languages. Evidence from UNESCO and University College London (2025) shows that smaller, task-specific models are more cost- and resource-efficient and more feasible in low-income and low-infrastructure contexts, where only 5 per cent of African artificial intelligence talent currently has access to sufficient compute resources. This reinforces the case for compact, efficient models as a pathway to democratize access under constraints of energy, water and connectivity. To accelerate adoption, governments and multilateral institutions should support this shift through targeted grant mechanisms and GovTech procurement preferences for open-source systems, building on proven Digital Public Infrastructure adoption pathways such as MOSIP and the Mojaloop Foundation, which demonstrate how open-source public goods can scale systemic digital transformation when aligned with public-sector demand and global development priorities.

#### **D. Development Implications**

The development implications of Small Artificial Intelligence extend beyond technological innovation to the broader challenge of inclusive economic transformation. Unlike frontier Artificial Intelligence systems that are concentrated in a few countries and firms, Small Artificial Intelligence lowers the barriers to participation by reducing infrastructure, energy and financial requirements. This enables governments, small businesses and local innovators in resource-constrained environments to develop and deploy Artificial Intelligence solutions tailored to local needs, languages and institutional contexts. By expanding access to productive digital technologies, Small Artificial Intelligence can strengthen human capital development, improve public service delivery and increase productivity across agriculture, healthcare, education and small enterprise sectors. In this sense, Small Artificial Intelligence functions as a developmental enabler that broadens the distribution of technological capabilities and reduces dependence on externally controlled Artificial Intelligence ecosystems.

From a development economics perspective, the Small Artificial Intelligence Digital Public Infrastructure Framework (SAI-DPIF) creates conditions for digital leapfrogging and structural transformation. As a shared infrastructure, it generates forward linkages by enabling new applications, services and entrepreneurial activity, while also creating backward linkages through demand for local skills, data stewardship, research and digital infrastructure investments (Hirschman, 1958). These multiplier effects can stimulate job creation and innovation ecosystems that are more aligned with local development priorities than capital-intensive frontier Artificial Intelligence models. Consequently, development policy should treat Artificial Intelligence not merely as a technology adoption challenge but as an infrastructure-building agenda. Investments in shared compute resources, open model repositories and interoperable application ecosystems can help countries build indigenous Artificial Intelligence capabilities, strengthen digital sovereignty and ensure that the economic benefits of Artificial Intelligence contribute to broad-based and sustainable development rather than further concentrating technological power.

### **E. Limitations and Future Research Directions**

The Small Artificial Intelligence Digital Public Infrastructure Framework (SAI-DPIF) introduced in this paper is a conceptual model rather than an empirically validated one. Although informed by evidence from Digital Public Infrastructure initiatives, Artificial Intelligence deployments and institutional practices, the relationships between Small Artificial Intelligence infrastructure and development outcomes remain theoretical. Accordingly, the framework should be viewed as a foundation for future empirical investigation rather than proof of demonstrated effectiveness.

Several limitations follow from this conceptual orientation. First, the framework requires empirical validation through natural experiments and large-scale implementations of Small Artificial Intelligence infrastructure. Current evidence remains limited, with most studies reporting pilot projects rather than sustained outcomes at scale. Second, the proposed governance architecture including the Global Small Artificial Intelligence Infrastructure Fund (GSAIF), Small Artificial Intelligence Readiness Index (SAIRI) and Small Artificial Intelligence Digital Public Goods Council requires significant institutional development before operationalization. The design of financing mechanisms, particularly for recurrent operational costs, also remains underdeveloped. Third, the political economy of adoption, including coalitional dynamics, vested interests and institutional conditions for implementation across diverse governance contexts, warrants further examination.

Future research should focus on evaluating the validity, feasibility and development impacts of the proposed framework. A key priority is longitudinal and econometric studies of large-scale Small Artificial Intelligence deployments to determine their effects on gross domestic product growth, productivity and poverty reduction. Research is also needed to compare alternative governance models and assess their cost-effectiveness relative to existing service-delivery mechanisms. In addition, scholars should examine the labour market implications of Small AI, particularly its effects on routine information-processing occupations. Finally, future research should examine algorithmic bias, data exclusion risks and the conditions under which Small AI-DPI can genuinely empower marginalized communities rather than replicating existing social exclusions in algorithmic form.

## VIII. CONCLUSION

This paper advances the argument that Small Artificial Intelligence constitutes an emerging layer of Digital Public Infrastructure and offers developing economies a distinct pathway to Artificial Intelligence-enabled development. Unlike frontier Artificial Intelligence systems, which are characterized by high compute requirements, capital concentration and energy intensity, Small Artificial Intelligence provides a more accessible and scalable model for resource-constrained environments. Drawing on public goods theory, platform economics, development linkage theory and leapfrogging theory, the paper introduces the Small Artificial Intelligence Digital Public Infrastructure Framework (SAI-DPIF) as a tripartite architecture comprising shared compute utilities, open model repositories and interoperable application ecosystems. It also demonstrates how governance, financing and scalability mechanisms can support practical implementation. The paper makes three principal contributions. First, it reframes the Artificial Intelligence and development challenge from one of technology adoption to one of infrastructure provision, with important implications for financing, regulation and capacity building. Second, it provides a theoretically grounded framework informed by evidence from institutions such as the World Bank, UNESCO, the International Telecommunication Union and the Center for Strategic and International Studies. Third, it introduces two policy instruments: the Global Small Artificial Intelligence Infrastructure Fund (GSAIF) and the Small Artificial Intelligence Readiness Index (SAIRI) to translate the framework into actionable institutional mechanisms. These contributions respond to growing concerns that Artificial Intelligence could widen global inequalities because of persistent disparities in compute capacity, data access and technical skills. As highlighted by the United Nations Development Programme's report *The Next Great Divergence: Why AI May Widen Inequality Between Countries* and the *World Development Report 2026*, high-income countries continue to dominate Artificial Intelligence investment and innovation, creating the risk of increasing developmental divergence. In this context, Small Artificial Intelligence should be viewed not merely as a technological alternative but as an infrastructure and governance strategy for inclusive development. Through coordinated multilateral investment, national policy alignment and open-source ecosystem participation, it offers a pathway for broadening access to Artificial Intelligence capabilities and their associated economic benefits. Future research should focus on empirically validating the framework through cross-country studies, refining tools such as the SAIRI, examining the political economy of adoption and generating evidence on impacts on productivity, economic growth and poverty reduction.

## REFERENCES

- Aayog, N. (2020). *Data empowerment and protection architecture: A secure consent-based data sharing framework* [Policy Document]. Government of India. <https://www.niti.gov.in/sites/default/files/2023-03/Data-Empowerment-and-Protection-Architecture-A-Secure-Consent-Based.pdf>
- Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *Journal of Economic History*, 46(2), 385–406. <https://doi.org/10.1017/S0022050700046209>
- AfricaNenda Foundation, World Bank, & United Nations Economic Commission for Africa. (2022). *The state of instant and inclusive payment systems (SIIPS) in Africa 2022*. AfricaNenda Foundation. [https://www.africanenda.org/uploads/files/EN\\_SIIPS\\_Report\\_web.pdf](https://www.africanenda.org/uploads/files/EN_SIIPS_Report_web.pdf)
- African Union Commission (2020). *The Digital Transformation Strategy for Africa (2020–2030)*. African Union. <https://au.int/sites/default/files/documents/38507-doc-dts-english.pdf>
- Agénor, P. R. (2010). A theory of infrastructure-led development. *Journal of Economic Dynamics and Control*, 34(5), 932-950.

- Alonso, C. T. Bhojwani, E. Hanedar, D. Prihardini, G. Una and Zhabska, K. (2023). *Stacking up the benefits: Lessons from India's digital journey* (IMF Working Paper No. 2023/078). International Monetary Fund, Volume 2023: Issue 078, <https://doi.org/10.5089/9798400240416.001>
- Alper, A., Pinar, O., Rabadan, G., Alfonso, G., Georgina, M., Harish, N., Piveteau, T. Y., Ramteke, N., & Sarkar, A. (2023). *G20 Policy Recommendations for Advancing Financial Inclusion and Productivity Gains through Digital Public Infrastructure (English)*. Payment Systems Washington, D.C.: World Bank Group. <http://documents.worldbank.org/curated/en/099092023121016458>
- ASEAN Secretariat (2024). *ASEAN Guide on AI Governance and Ethics*. Association of Southeast Asian Nations. [https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics-beautified\\_201223\\_v2.pdf](https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics-beautified_201223_v2.pdf)
- Bakhtina, M., Leung, K. L., Matulevičius, R., Awad, A., & Švenda, P. (2023, August). A decentralised public key infrastructure for X-Road. In *Proceedings of the 18th International Conference on Availability, Reliability and Security* (pp. 1-8). <https://doi.org/10.1145/3600160.3605092>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- ByteIota (2026). *Small language models: Gartner's 3x edge AI prediction*. <https://byteiota.com/small-language-models-gartners-3x-edge-ai-prediction/> (citing Gartner, April 2025)
- Calderón, C., & Servén, L. (2008). Infrastructure and economic development in Sub-Saharan Africa (Policy Research Working Paper No. 4712). *World Bank*. <https://doi.org/10.1596/1813-9450-4712>
- Carnegie Endowment for International Peace (2025). *Digital public infrastructure: A practical approach for Africa*. <https://carnegieendowment.org/research/2025/02/digital-public-infrastructure-a-practical-approach-for-africa>
- CARICOM Secretariat (2024). *CARICOM Heads of Government endorse digital resilience strategy, skills fund creation*. <https://caricom.org/caricom-heads-of-government-endorse-digital-resilience-strategy-skills-fund-creation/>
- Ciplet, D., & Roberts, J. T. (2017). Climate change and the transition to neoliberal environmental governance. *Global Environmental Change*, 46, 148–156. <https://doi.org/10.1016/j.gloenvcha.2017.09.003>
- Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). SAGE Publications
- Corradini, F., Leonesi, M., & Piangerelli, M. (2025). State of the Art and Future Directions of Small Language Models: A Systematic Review. *Big Data and Cognitive Computing*, 9(7), 189. <https://doi.org/10.3390/bdcc9070189>
- Cottier, B., et al. (2024). *The rising costs of training frontier AI models*. arXiv:2405.21015. <https://arxiv.org/abs/2405.21015>
- CSIS (2025). *The AI power surge: Growth scenarios for GenAI datacenters through 2030*. <https://www.csis.org/analysis/ai-power-surge-growth-scenarios-genai-datacenters-through-2030>
- Dawa Health. (2026). *DawaMom: AI-powered maternal health platform* [Platform website]. <https://dawa-health.com/>
- Dell Technologies. (2026, January 7). *The power of small: Edge AI predictions for 2026* [Blog post]. <https://www.dell.com/en-us/blog/the-power-of-small-edge-ai-predictions-for-2026/>
- Demirgüç-Kunt, A., Klapper, L., Singer, D., Ansar, S., & Hess, J. (2018). The Global Findex database 2017: Measuring financial inclusion and the fintech revolution. World Bank Publications. <https://doi.org/10.1596/978-1-4648-1259-0>, <https://documents1.worldbank.org/curated/en/332881525873182837/pdf/126033-PUB-PUBLIC-pubdate-4-19-2018.pdf>

- Dencik, L., Hintz, A., Redden, J., & Treré, E. (2019). Exploring data justice: Concepts, applications and directions. *Information, Communication & Society*, 22(7), 873–881. <https://doi.org/10.1080/1369118X.2019.1606268>
- Digital Public Goods Alliance. (2023). *DPG standard*. <https://new.digitalpublicgoods.net/standard>
- Digital India BHASHINI Division. (2026, March). BHASHINI: Bridging 1.4 billion citizens to services and opportunities via language AI. *eGov Magazine*. <https://egov.eletsonline.com/2026/03/bhashini-bridging-1-4-billion-citizens-to-services-and-opportunities-via-language-ai/>
- Dixit, A., Jakhar, S. K., & Kumar, P. (2025). Assessing Unified Payments Interface (UPI) adoption and usage through the interplay of UTAUT factors. *Humanities and Social Sciences Communications*, 12, Article 1065. <https://doi.org/10.1057/s41599-025-05313-w>
- Eisenmann, T. R., Parker, G., & Van Alstyne, M. W. (2006). Strategies for two sided markets. *Harvard Business Review*, Vol. October
- Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532–550. <https://doi.org/10.5465/amr.1989.4308385>
- Estache, A. (2010). Infrastructure finance in developing countries: An overview. *EIB Papers*, 15(2), 60-88. EIB Papers
- Evans, D. S., & Gawer, A. (2016). *The rise of the platform enterprise: A global survey*. The Center for Global Enterprise.
- Frischmann, B. M. (2009). Infrastructure commons in economic perspective. In *Internet Policy and Economics: Challenges and Perspectives* (pp. 29-55). Boston, MA: Springer US. [https://doi.org/10.1007/b104899\\_3](https://doi.org/10.1007/b104899_3)
- Frischmann, B. M. (2006). An economic theory of infrastructure and commons management. *Minnesota Law Review*, 89(4), 917–1030
- Gartner. (2025, April 9). Gartner predicts by 2027, organizations will use small, task-specific AI models three times more than general-purpose large language models [Press release]. <https://www.gartner.com/en/newsroom/press-releases/2025-04-09-gartner-predicts-by-2027-organizations-will-use-small-task-specific-ai-models-three-times-more-than-general-purpose-large-language-models>
- G20 Digital Economy Working Group. (2023a). G20 Digital Economy Ministers' Meeting outcome document and chair's summary. Government of India. <https://g7g20-documents.org/database/document/2023-g20-india-sherpa-track-digital-economy-ministers-ministers-language-g20-digital-economy-ministers-meeting-outcome-document-and-chair-summary#section-3>
- G20 Task Force on Digital Public Infrastructure. (2023b). Report of India's G20 Task Force on DPI for economic transformation, financial inclusion and development. Government of India. <https://www.pib.gov.in/PressReleaseIframePage.aspx?PRID=2033389&reg=3&lang=2>
- Government of India, Press Information Bureau. (2022, May 24). *MeitY holds brainstorming with researchers and startups to shape strategy for Digital India Bhashini* [Press release]. <https://www.pib.gov.in/PressReleaseIframePage.aspx?PRID=1827997>
- Government of India, Press Information Bureau. (2024, March 7). *Cabinet approves ambitious IndiaAI Mission to strengthen the AI innovation ecosystem* [Press release]. <https://www.pib.gov.in/PressReleaseIframePage.aspx?PRID=2012355>
- Government of India, Press Information Bureau. (2025, March 6). *MeitY launches AIKosha: IndiaAI datasets platform* [Press release]. <https://www.pib.gov.in/PressReleasePage.aspx?PRID=2108961>
- Government of India, Press Information Bureau. (2026a, March 25). *IndiaAI Mission: Statement of Union Minister of State for Electronics and IT Jitin Prasada in the Lok Sabha* [Parliamentary record]. <https://www.pib.gov.in/PressReleasePage.aspx?PRID=2209737>

- Government of India, Press Information Bureau. (2026b, March 13). *AI models developed under IndiaAI Mission* [Press release]. <https://www.pib.gov.in/PressReleasePage.aspx?PRID=2239614>
- Green Climate Fund. (2023). *GCF at a glance*. <https://www.greenclimate.fund/about>
- Hazari, B., & Mohan, V. (2024). Exclusion and the growth of AI technology: A trade-theoretic analysis. *Frontiers in Human Dynamics*, 6, Article 1203664. <https://doi.org/10.3389/fhumd.2024.1203664>
- Heeks, R. (2022). Digital inequality beyond the digital divide: Conceptualizing adverse digital incorporation in the global South. *Information Technology for Development*, 28(4), 688–704. <https://doi.org/10.1080/02681102.2022.2068492>
- Hirschman, A. O. (1958). *The strategy of economic development*. Yale University Press
- Ibrahimov, O. (2025). AI as public infrastructure: A critical review of the transition from tool to societal necessity. *Current Trends in Computing*, 3(2), 40-61. <http://dx.doi.org/10.2139/ssrn.5930164>
- IndiaAI Mission, Ministry of Electronics and Information Technology. (2025). *AIKosh: National AI repository* [Platform]. <https://aikosh.indiaai.gov.in>
- IEA (2025). *Energy and AI*. International Energy Agency. <https://www.iea.org/reports/energy-and-ai>
- IMF- International Monetary Fund. (2023). Pix: Brazil's successful instant payment system. *IMF Staff Country Reports*, 2023(289). <https://doi.org/10.5089/9798400249266.002>
- ITU - International Telecommunication Union (2024). Defining and building digital public infrastructure for all. <https://www.itu.int/hub/2024/10/defining-and-building-digital-public-infrastructure-for-all/>
- ITU (2024a). *The future of AI, regulation and industry development: Report of the AI for Good Global Summit 2024*. International Telecommunication Union. [https://www.itu.int/dms\\_pub/itu-t/opb/ai4g/T-AI4G-AI4GOOD-2024-5-PDF-E.pdf](https://www.itu.int/dms_pub/itu-t/opb/ai4g/T-AI4G-AI4GOOD-2024-5-PDF-E.pdf)
- Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, 10(1–2), 18–26. <https://doi.org/10.1007/s13162-020-00161-0>
- Jain, A. S. (2025). Decoding consent managers under the Digital Personal Data Protection Act, 2023: Empowerment architecture, business models and incentive alignment. *Journal of Data Protection & Privacy*, 7(4), 406–420. <https://doi.org/10.69554/MHLP2916>
- Kaur, A. (2025, October 14). *From divide to delivery: How AI can serve the Global South* [Commentary]. Center for Strategic and International Studies. <https://www.csis.org/analysis/divide-delivery-how-ai-can-serve-global-south>
- Kaul, I., Grunberg, I., & Stern, M. A. (1999). Defining global public goods. *Global public goods: International cooperation in the 21st century*, 2-19. Oxford University Press
- Khan, M. S., Umer, H., & Faruq, F. (2024). Artificial intelligence for low income countries. *Humanities and Social Sciences Communications*, 11(1), 1-13. <https://doi.org/10.1057/s41599-024-03947-w>
- Kim, J., et al. (2026). Reviewing the socio-technical dynamics of AI, data centers and digitalization on energy and the environment. *Renewable and Sustainable Energy Reviews*, 234, 116861. <https://doi.org/10.1016/j.rser.2026.116861>
- Lorenzo, M. F., Joia, L. A., & Stoll, F. C. (2026). Mobile Instant Payment Systems and Financial Inclusion: A Delphi-Based Analysis of Brazil's Pix. *The Electronic Journal of Information Systems in Developing Countries*, 92(3), e70064. <https://doi.org/10.1002/isd2.70064>
- Leroux, S., Bohez, S., De Coninck, E., Verbelen, T., Vankeirsbilck, B., Simoens, P., & Dhoedt, B. (2022). Energy consumption of on-device machine learning models for IoT intrusion detection. *Internet of Things*, 21, 100670. <https://doi.org/10.1016/j.iot.2022.100670>
- Lin, L., Wijayawardana, R., Rao, V., Nguyen, H., Gribga, W. E., & Chien, A. A. (2024). Exploding AI power use: An opportunity to rethink grid planning and management. In *Proceedings of*

- the 15th ACM International Conference on Future and Sustainable Energy Systems (e-Energy '24). pp. 434-441. <https://doi.org/10.1145/3632775.3661959>
- Microsoft Research. (2025). *Project Gecko* [Project page]. <https://www.microsoft.com/en-us/research/project/project-gecko/>
- Ministry of Electronics and Information Technology, Government of India. (2022, July). *Digital India Bhashini: India's AI-led language translation platform* [Policy document]. <https://static.pib.gov.in/WriteReadData/specificdocs/documents/2022/aug/doc202282696201.pdf>
- MIT Solve. (2023). *DawaMom — The Horizon Prize 2023* [Challenge solution profile]. <https://solve.mit.edu/challenges/the-horizon-prize-2023/solutions/73663>
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT\* '19)*, 220–229. <https://doi.org/10.1145/3287560.3287596>
- Mrisho, L. M., Mbilinyi, N. A., Ndalawa, M., Ramcharan, A. M., Kehs, A. K., McCloskey, P. C., Murithi, H., Hughes, D. P., & Legg, J. P. (2020). Accuracy of a smartphone-based object detection model, PlantVillage Nuru, in identifying the foliar symptoms of the viral diseases of cassava — CMD and CBSD. *Frontiers in Plant Science*, 11, Article 590889. <https://doi.org/10.3389/fpls.2020.590889>
- Nasscom & Arthur D. Little (2024). *Digital Public Infrastructure of India — Accelerating India's digital inclusion*. <https://nasscom.in/knowledge-center/publications/nasscom-arthur-d-little-indias-digital-public-infrastructure>
- North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511808678>
- Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.
- Ostrom, E. (2010). Beyond markets and states: Polycentric governance of complex economic systems. *American Economic Review*, 100(3), 641–672. <https://doi.org/10.1257/aer.100.3.641>
- Parker, G. G., Van Alstyne, M. W., & Choudary, S. P. (2016). *Platform revolution: How networked markets are transforming the economy and how to make them work for you*. W. W. Norton
- Parsheera, S. (2024). Stack is the new black?: Evolution and outcomes of the 'India-Stackification' process. *Computer Law & Security Review*, 52, 105947. <https://doi.org/10.1016/j.clsr.2024.105947>
- Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, Cal.: Sage Publications, 4(10.4135), 9781452226058
- Perez, C. (2002). Technological revolutions and financial capital: The dynamics of bubbles and golden ages. In *Technological revolutions and financial capital*. Edward Elgar Publishing.
- Pilz, K. F., Mahmood, Y., & Heim, L. (2025). *AI's power requirements under exponential growth: Extrapolating AI data center power demand and assessing its potential impact on U.S. competitiveness* (Report No. RRA3572-1). RAND Corporation. <http://www.rand.org/t/RRA3572-1>
- Ravallion, M. (2012). Troublesome trade-offs in human development index. *Journal of Development Economics*, 99(2), 201–209. <https://doi.org/10.1016/j.jdeveco.2012.01.003>
- Rochet, J. C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the european economic association*, 1(4), 990-1029.
- Rocco, T. S., & Plakhotnik, M. S. (2009). Literature reviews, conceptual frameworks, and theoretical frameworks: Terms, functions, and distinctions. *Human Resource Development Review*, 8(1), 120–130. <https://doi.org/10.1177/1534484309332617>
- Reboot Democracy AI. (2026, January 21). *Voices in every language: How India is building more inclusive AI*. <https://rebootdemocracy.ai/blog/voices-in-every-language-bhashadaan-india>

- Rochet, J.-C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4), 990–1029. <https://doi.org/10.1162/154247603322493212>
- Royal Academy of Engineering. (2024). *Esther Kimani: Early crop pest and disease detection device — Africa Prize for Engineering Innovation 2024*. <https://africaprize.raeng.org.uk/2024-cohort/esther-kimani/>
- Sajjadi Mohammadabadi, S. M., Kara, B. C., Eyupoglu, C., Uzay, C., Tosun, M. S., & Karakuş, O. (2025). A Survey of Large Language Models: Evolution, Architectures, Adaptation, Benchmarking, Applications, Challenges, and Societal Implications. *Electronics*, 14(18), 3580. <https://doi.org/10.3390/electronics14183580>
- Samuelson, P. A. (1954). The pure theory of public expenditure. *Review of Economics and Statistics*, 36(4), 387–389. <https://doi.org/10.2307/1925895>
- Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). *DistilBERT, a distilled version of BERT: Smaller, faster, cheaper and lighter* [NeurIPS Workshop Paper]. <https://doi.org/10.48550/arXiv.1910.01108>
- Sengupta A., Barbosa, A.C., & Samdub, M.T. (2025). Understanding interrelationships between AI and digital public infrastructure (DPI) in India and Brazil. *The African Journal of Information and Communication (AJIC)*, 35, 1–11. <https://doi.org/10.23962/ajic.i35.20800>
- Sheik, T., Maple, C., & Epiphaniou, G. (2022, June). Considerations for secure mosip deployment. In *Competitive Advantage in the Digital Economy (CADE 2022)* (Vol. 2022, pp. 135-143). IET
- Stiglitz, J. E. (1999). Knowledge as a global public good. *Global public goods: International cooperation in the 21st century*, 308(1999), 308-325
- Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645–3650. <https://doi.org/10.18653/v1/P19-1355>
- Thukral, A., & Lakshit, D. (2025). India's digital public infrastructure: Analyzing UPI and Aadhaar in GDP growth and cost optimization. *Journal of Emerging Investigators*. <https://doi.org/10.59720/25-046>
- UNCTAD (2025). *World Investment Report 2025*. <https://unctad.org/publication/world-investment-report-2025>
- UNESCO (2021). *Recommendation on the Ethics of Artificial Intelligence*. Adopted 24 November 2021 at the 41st Session of the UNESCO General Conference. <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>
- UNESCO & University College London. (2025). *Smarter, smaller, stronger: Resource-efficient generative AI and the future of digital transformation*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000394521>
- United Nations Development Programme. (2023, September 18). *UN digital event mobilizes global leadership and US\$400 million to support digital public infrastructure*. UNDP. <https://www.undp.org/press-releases/un-digital-event-mobilizes-global-leadership-and-us400-million-support-digital-public-infrastructure>
- United Nations Development Programme. (2025a). *A matter of choice: People and possibilities in the age of AI* [Human Development Report 2025]. UNDP. <https://hdr.undp.org/content/human-development-report-2025>
- United Nations Development Programme. (2025b). *The next great divergence: Why AI may widen inequality between countries*. UNDP Asia-Pacific. <https://www.undp.org/asia-pacific/press-releases/ai-risks-sparking-new-era-divergence-development-gaps-between-countries-widen-undp-report-finds>
- Wang, X., Xu, Z., & Sui, X. (2025). Intelligent data analysis in edge computing with large language models: applications, challenges, and future directions. *Frontiers in Computer Science*, 7, 1538277. <https://doi.org/10.3389/fcomp.2025.1538277>

- Wilkinson, B., et al. (2025). *The case for expanding digital public infrastructure*. Chatham House. <https://doi.org/10.55317/9781784136604>
- World Bank. (2023). *A digital stack for transforming service delivery*. World Bank Group. <https://documents1.worldbank.org/curated/en/099755004072288910/pdf/P1715920edb5990d60b83e037f756213782.pdf>
- World Bank. (2025a). Digital progress and trends report 2025: Strengthening AI foundations. *World Bank*. <https://doi.org/10.1596/978-1-4648-2264-3>
- World Bank. (2025b). *Small AI, big impact: Harnessing artificial intelligence for development*. World Bank Group. <https://blogs.worldbank.org/en/voices/small-ai-big-impact-harnessing-artificial-intelligence-for-development>
- World Bank (2026). *World Development Report 2026: Decoding AI for development*. World Bank Group. <https://www.worldbank.org/en/publication/wdr2026>
- Zhang, Y., & Zhang, T. (2025). Beyond Private or Public: Large Language Models as Quasi-Public Goods in the AI Economy. *arXiv preprint arXiv:2509.13265*