

October 2026

Global Web

The State of Digital Access, 2026

A Baseline Assessment of Linguistic Diversity in the Global Digital Ecosystem



Coalition on
Digital Impact

Table of Contents

EXECUTIVE SUMMARY.....3

PART I: VOICES FROM THE LAST MILE..... 5

 Survey Overview.....5

 The Needs of Last-Mile Organizations..... 5

 The Needs of Communities Served..... 7

PART II: THE GLOBAL DIGITAL DEVELOPMENT LANDSCAPE..... 11

 Pillar 1: Infrastructure and Access..... 12

 Pillar 2: Digital Skills and Literacy..... 12

 Pillar 3: Digital Governance and Responsibility..... 13

 Pillar 4: Innovation and Transformation..... 14

PART III: CONNECTING DEMAND AND SUPPLY..... 16

CONCLUSION AND NEXT STEPS..... 17

 Next Steps..... 17

APPENDICES..... 18

 A. Glossary of Key Terms..... 18

 B. Program and Technology Reference..... 19

 C. Survey Respondents List..... 21

 D. Survey Data / Charts..... 22

Executive Summary

The Coalition on Digital Impact ([CODI](#)) was formed with the vision that everyone can navigate the Internet in their own language. While the Internet has become essential infrastructure for modern life, linguistic diversity online remains limited with only approximately ten languages accounting for most content despite more than 7,000 languages being spoken worldwide. As artificial intelligence (AI) accelerates the pace of digital interaction, this concentration threatens to deepen linguistic homogenization at an unprecedented scale, quietly excluding billions of people from the opportunities the Internet provides.

This report presents CODI's baseline assessment, bringing together two parallel lines of inquiry: what do last-mile organizations working directly with underserved language communities (the demand side) need to do their work, and what resources are available to support the advancement of digital representation (the supply side)?

Key Findings

- **Demand is real and persistent.** More than half of the last-mile organizations surveyed report that the communities they serve cannot effectively navigate the Internet in their native language. The barriers are not peripheral. They are structural, touching every layer of the digital experience.
- **Supply is substantial but linguistically blind.** A robust ecosystem of global programs, policies, and technologies exists to advance digital development. However, linguistic diversity is rarely an explicit design requirement. Most supply-side initiatives treat language as a secondary concern, addressed through add-on translation rather than foundational design.
- **A coordination gap exists.** Last-mile organizations, global program funders, and technology providers operate largely in parallel. A structured bridge between the demand and supply sides does not yet exist.
- **Universal Acceptance is necessary but not sufficient.** Universal Acceptance (the technical standard ensuring all domain names, scripts, and email formats are accepted across systems) is a critical foundation. But

communities cannot participate online if their scripts are readable but not searchable, their content exists but cannot be found, or their language disappears precisely when they need it most.

This report is offered as a shared baseline for CODI, our partners, funders, and the broader community. It is the foundation for future phases of deeper engagement, collaborative resource development, and policy advocacy aimed at ensuring that every person can participate in the Internet in their own language.

Part I: Voices from the Last Mile

The demand side of CODI's baseline inquiry captures the perspectives of last-mile organizations: civil society groups, community initiatives, research institutes, infrastructure providers, technology organizations, and policy advocates who work directly with language communities underserved by the current digital ecosystem. Their voices reveal not only the immediate needs of the organizations themselves but the lived realities of the tens of millions of community members they serve.

Survey Overview

CODI conducted an online survey titled "The Global Web: Understanding Your Needs" from December 2025 to February 2026. Respondent profiles spanned six civil society and advocacy organizations, three infrastructure providers, two research institutes, two technology organizations, two community initiatives, one policy organization, and one consulting organization. Geographically, respondents operated across Africa (Ghana, Kenya, Uganda, Ethiopia), South Asia (India, Nepal), Southeast Asia, East Asia, Australia, North America (Canada, United States), and Europe (Netherlands, Slovenia, EU). The full list of responding organizations can be found in [Appendix C](#).

The Needs of Last-Mile Organizations

Respondents are organizations that actively innovate by creating language datasets, developing digital tools, conducting policy advocacy, and building community capacity. But across regions and organization types, four structural constraints consistently limit their impact.

1. Funding and Sustainability

Short-term, project-based funding are the single most frequently cited barriers. Organizations working on language preservation, localization, and digital literacy often operate on grant cycles that cannot sustain the long-term investment

required for meaningful impact. Respondents called for diversified, sustained funding mechanisms, particularly in support of foundational language technologies and community-led initiatives. Many emphasized the need for long-term partnerships rather than discrete project grants.

For example, CLEAR Global noted the need for "data that drives funding to foundational technology." Rise Afrika Consulting identified "limited funding and short-term project models" as the primary constraint on sustained language data development and community validation. Internet Society Uganda highlighted the need for a "diversified funding strategy" to ensure financial and project sustainability.

2. Data, Evidence, and Advocacy Tools

Organizations reported needs for stronger data, statistics, key performance indicators, and case studies to demonstrate the extent of linguistic exclusion and make the case for investment. Multiple respondents noted the critical importance of challenging the dominant narrative that language access has effectively been solved. This narrative crowds out funding and policy attention at precisely the moment when the gaps are most acute.

For example, the .au Domain Administration (auDA) highlighted the need for "data regarding need and value" alongside trustworthy translation tools. Hybrid Lynx requested "case studies, data and KPIs, community level engagement and reports." GNKS Consult BV emphasized the importance of "success stories and case studies that help share lessons learned."

3. Multilingual Technology and Infrastructure

There is widespread demand for improved multilingual user interfaces, reliable and accurate translation tools, language data resources and corpora for underrepresented languages, and user-friendly content management systems capable of supporting multilingual workflows. Technical gaps, including limited developer expertise in local-language tools, insufficient support for non-Latin scripts, and lack of standardized terminology remain significant constraints.

For example, the Digital Empowerment Foundation called for "user friendly tools to be used by rural communities across all age groups" along with technical training. HakiDigital Initiative requested multilingual user interfaces alongside narrative and communication materials. Rise Afrika identified the specific need for "multilingual data resources and corpora specific to East African languages" to support better language models and localization quality.

4. Capacity Building and Policy Guidance

Organizations seek technical training in localization, digital tool development, and language data generation for their staff and the communities they work with. Equally important is the need for practical policy frameworks grounded in inclusive digital language standards, which organizations can use to influence governments and platforms. Many respondents operate at the intersection of community service and policy advocacy, requiring both operational tools and policy leverage.

For example, the Endangered Languages Project noted that community partners need "technical training and capacity building, as well as access to technical tools and resources." GNKS Consult BV emphasized the need to ensure "policy frameworks are well informed" by standards and success stories.

The Needs of Communities Served

The communities represented in this survey collectively speak hundreds of languages, including Indigenous, African, South Asian, and European languages. Many use non-Latin scripts or languages without standardized orthographies. Collectively, these communities number in the tens of millions across rural and urban contexts on every inhabited continent. Their challenges are foundational and immediate.

1. Access and Affordability

Mobile phones are the primary and often only point of Internet access for most communities represented in this survey. Connectivity is frequently intermittent, costly relative to local incomes, and unreliable, particularly in rural and border

regions. Some communities served have no Internet access at all. In other cases, the choice is between connectivity and food.

These constraints are not merely inconvenient: they limit the types of multilingual services that can realistically be used. Data-intensive applications, including AI-based translation services and online learning platforms are effectively unavailable to communities with low-bandwidth or intermittent connections. Lightweight, offline-capable, and localized solutions are essential, not optional.

" The cost of Internet can equal the price of a meal. "

Internet Society Uganda

2. Inability to Navigate the Internet in One's Native Language

More than half of survey respondents indicated that the communities they serve cannot effectively navigate the Internet in their native language.

This is among the most striking findings of the survey, that **linguistic presence across the full digital experience is the exception, not the norm, for the majority of the world's languages.**

"Most East African language speakers access only news and community information in their native languages; activities such as online shopping, e-learning, or professional work are conducted in English or French due to the scarcity of localized platforms."

Rise Afrika Consulting

Where local-language content exists, it is typically limited to news and social media. Professional, educational, commercial, and government services are largely available only in dominant languages.

3. Lack of Localized Content and Tools

Communities face a scarcity of educational, governmental, and commercial resources in their own languages. Reported technical barriers include poor support for non-Latin scripts, online forms that reject Internationalized Domain Name (IDN)

email addresses, absence of language-specific keyboards and input tools, and limited indexing of local-language content in search engines. EGIGFA specifically identified IDN email rejection as "a huge challenge." The Endangered Languages Project reported that community partners face "all of the above: lack of language-specific keyboards, lack of access to other foundational tech requirements, online forms not accepting IDN email addresses, lack of content in their specific language, poor translation of content."

4. Translation Quality and Trust

Automated translations are frequently inaccurate or culturally inappropriate, reducing trust in digital services and increasing the risk of misinformation. Respondents noted that platform-generated translations often rely on generic language models rather than community-based speakers, producing content that ranges from misleading to harmful. The absence of standardized terminology complicates localization efforts and erodes confidence in digital tools. Rise Afrika specifically called out "localization and representation gaps, where global platforms rely on heritage speakers or generic language models rather than indigenous and community-based speakers, leading to inaccuracies, tokenism, and reduced trust."

5. Digital Literacy and Social Barriers

Low digital literacy, limited awareness of available tools, unreliable electricity, and gender-based access barriers further restrict meaningful participation. Internet Society Uganda noted that women in rural areas, in particular, face "significantly less meaningful connectivity due to factors like reliance on men for device use and limited time due to care work." The Indigenous Connectivity Institute raised a deeper structural question: whether digital technology, as currently constituted, can effectively carry the spirit and meaning of languages rooted in oral tradition.

What We Learned

Across geographies, organization types, and language communities, the demand-side assessment reveals a consistent picture. Organizations working closest to underserved language communities face persistent barriers that extend well beyond Internet connectivity. Their needs span sustainable funding,

multilingual technologies, localized content, policy guidance, and capacity building, while the communities they serve continue to struggle to participate online in their own languages.

Key finding: Demand for multilingual digital participation is real, widespread, and persistent.

Part II: The Global Digital Development Landscape

The supply side of CODI's baseline inquiry examines the global ecosystem of programs, policies, and technologies aimed at advancing digital development and representation. A broad review of sources from major multilateral organizations, including the World Bank, UNESCO, other UN agencies, the G20, the Internet Telecommunication Union (ITU), and the Digital Cooperation Organization (DCO), alongside national strategies from governments including the United States, United Kingdom, France, Sweden, Brazil, South Africa, and Nigeria, reveals a substantial and increasingly coordinated global effort.

In particular, UNESCO's [International Decade of Indigenous Languages](#) (2022-2032) has elevated linguistic diversity, language technologies, and multilingual digital participation as global priorities.

This landscape scan finds that supply-side programs can be organized according to a Four-Pillar Strategic Framework that has emerged as a convergence point across national strategies and multilateral bodies. These pillars move beyond simple connectivity to encompass the social, technical, and legal dimensions of digital development. Cutting across all four pillars is the [Digital Public Infrastructure](#) (DPI) approach. This approach, developed through collaboration between the World Bank, the Gates Foundation, and Co-Develop focuses on foundational, reusable "building blocks" that can underpin entire digital economies.

A Critical Observation

Despite the scale and sophistication of supply-side programs, the explicit design of digital systems, tools, and content to support underrepresented languages is rarely an articulated goal or design requirement. Most programs treat language as a secondary concern, addressed through post-hoc translation or optional localization rather than foundational design. This is the central misalignment between the demand and supply sides.

Pillar 1: Infrastructure and Access

The first pillar addresses the physical and technical foundations of digital participation: connectivity, devices, and affordability. Programs at this level include major national broadband initiatives such as the US [Broadband Equity, Access, and Deployment](#) (BEAD) Program, the Nigeria [National Broadband Plan](#) (2020–2025) targeting 70% broadband penetration, and Sweden's commitment to near-universal 1 Gbps access. Globally, the UNICEF/ITU [Giga initiative](#) works to connect every school in the world to the Internet, while China's [Digital Silk Road](#) provides state-funded telecommunications infrastructure to developing countries.

The Digital Public Infrastructure identity layer works to ensure that individuals have verifiable digital identities as the prerequisite for accessing digital services. These include open-source systems such as [MOSIP](#) (Modular Open Source Identity Platform), [OpenCRVS](#) (civil registration), and the World Bank's [ID4D](#) program. These systems are operational in dozens of countries, including Ethiopia's [Fayda national ID program](#), which provides services via low-tech USSD channels for citizens without smartphones.

From a linguistic inclusion perspective, infrastructure and connectivity are necessary but insufficient conditions. A community member with reliable Internet access still faces all of the barriers documented in the demand-side findings if the systems they connect to do not support their language. Universal Acceptance ensures that IDNs, internationalized email addresses, and non-Latin scripts are fully accepted across all systems, and is the linguistic dimension of this pillar that remains most underdeveloped.

Pillar 2: Digital Skills and Literacy

The second pillar focuses on human capital development, encompassing the skills and capabilities that enable people to participate meaningfully in digital life. Programs in this pillar include the UK [Digital Access Programme](#), the British Council's [Skills for Inclusive Digital Participation](#) (SIDP) initiative operating in Indonesia, Kenya, and Nigeria, and the ITU's [Digital Transformation Centres](#).

France's [Digital Mediators](#) program deploys trained human intermediaries in communities to help citizens navigate online public services, recognizing that access alone does not guarantee meaningful participation.

Major technology companies are increasingly active in this pillar. Google.org funds [digital skills and AI literacy](#) programs for underserved communities. Microsoft's [AI for Good initiative](#) supports global skilling efforts with emphasis on economic inclusion. Meta has pursued [connectivity and digital literacy initiatives](#), particularly in the Global South, including tools and AI capabilities designed for local languages and cultural contexts.

From a linguistic inclusion perspective, digital skills programs frequently assume that training can be delivered in dominant languages and that the underlying tools being taught are already available in community languages. The demand-side findings demonstrate that this assumption does not hold for most underrepresented language communities. Training programs that do not also address the linguistic barriers in the tools they teach produce incomplete results.

Pillar 3: Digital Governance and Responsibility

The third pillar encompasses the legal frameworks, safety mechanisms, and rights protections that constitute a trusted digital environment. Key policies include Brazil's General Personal Data Protection Law ([LGPD](#)), Nigeria's Data Protection Regulation ([NDPR](#)), and the [Budapest Convention on Cybercrime](#), which works to harmonize national cybersecurity laws. The UK's Cybersecurity Capacity Maturity Model ([CMM](#)) has been deployed in over 90 countries to assess national cyber resilience. The UN and United Nations Development Program (UNDP)'s Universal DPI [Safeguards Framework](#) provides a comprehensive set of principles for evaluating Digital Public Infrastructure implementations, including inclusion, privacy, and accountability.

The DPI payments and data exchange layers sit within this pillar. Open-source payment systems such as [Mojaloop](#), Brazil's [Pix](#) instant payment system (which

reached 76% of the adult population within two years of launch), and the [G2Px](#) World Bank initiative to digitize government-to-person transfers are transforming financial access at scale. Data exchange systems including [X-Road](#), [DHIS2](#), and [GovStack](#) are enabling governments to digitize services without vendor lock-in.

From a linguistic inclusion perspective, governance frameworks establish the rights and safeguards within which digital participation occurs. Yet most current frameworks are silent on linguistic rights in digital spaces. Data protection laws protect personal information but do not address the right to access services in one's own language. Cybersecurity programs build technical resilience but rarely address the linguistic dimension of online safety. The fact is that security warnings, fraud alerts, and protective notifications are largely unavailable in underrepresented languages.

Pillar 4: Innovation and Transformation

The fourth pillar focuses on applying digital tools across sectors to drive growth, solve public problems, and advance sustainable development. Programs in this pillar include the UK's [AI for Development initiative](#), which supports responsible AI research labs in African universities, France's [Citizen Initiatives Accelerator](#) connecting civil society projects with government, and the ITU's [AI for Good](#) platform. The Digital Public Goods Alliance ([DPGA](#)) catalogs and promotes open-source digital tools that can be freely adopted and adapted across countries.

Linguistic inclusion is most explicitly addressed in this pillar, primarily through AI research. Meta's [No Language Left Behind](#) project has developed multilingual AI models intended to address linguistic exclusion in AI systems. Google Research and Google.org invest in [low-resource language models](#) and culturally inclusive AI tools. Microsoft's [AI for Good labs](#) support applied research in health, agriculture, and humanitarian contexts. UNICEF's [RapidPro](#) platform allows organizations to build interactive messaging services via SMS and WhatsApp, enabling service delivery and data collection in low-connectivity, potentially multilingual contexts.

However, most AI language initiatives focus on a subset of major world languages, leaving the thousands of underrepresented languages that CODI's respondents work with largely unaddressed. Commercial AI language investment tends to follow market size, creating a structural misalignment with the communities most in need.

UNESCO's [Global Roadmap for Multilingualism in the Digital Era](#) provides an important policy framework. It calls for investment in language data, speech technologies, machine translation, digital content creation, and community participation to ensure that AI and language technologies support linguistic diversity rather than reinforce existing digital inequalities.

What We Learned

The supply-side assessment paints a very different picture. Governments, multilateral organizations, civil society, and the private sector have invested heavily in digital infrastructure, digital skills, governance, and innovation. Yet across these efforts, linguistic inclusion is rarely treated as a foundational design objective. Instead, language is typically addressed through translation or localization after systems have already been developed.

The assessment also highlights the role of Universal Acceptance. Ensuring that all domain names, email addresses, and scripts function correctly across digital systems is an essential prerequisite for multilingual participation. However, Universal Acceptance addresses only one aspect of a much broader challenge. Communities also need content, services, tools, and trustworthy information that are available in their own languages.

Key findings:

- The supply ecosystem is substantial, but it remains largely linguistically blind.
- Universal Acceptance is necessary, but it is not sufficient to achieve meaningful multilingual digital participation.

Part III: Connecting Demand and Supply

This baseline assessment presented in this report confirms that the scale of the challenge of multilingual digital participation is not a lack of interest or effort. Last-mile organizations working directly with underserved language communities have clearly articulated persistent needs for localized content, multilingual technologies, sustainable funding, stronger evidence, and practical policy guidance. At the same time, governments, multilateral organizations, philanthropic institutions, and technology companies are investing heavily in digital infrastructure, skills, governance, and innovation. Yet these two ecosystems rarely intersect. Organizations working directly with underserved language communities often lack visibility into available programs and resources, while global digital development initiatives rarely incorporate the needs of those communities into their design. This coordination gap helps explain why significant investment in digital development has not translated into equitable multilingual participation.

Key finding: Last-mile organizations, global program funders, and technology providers operate largely in parallel. A structured bridge between the demand and supply sides does not yet exist.

Conclusion and Next Steps

This baseline assessment provides a shared understanding of both the demand for multilingual digital participation and the current landscape of digital development initiatives. By bringing these perspectives together, it provides a common understanding of the current landscape and highlights opportunities for greater collaboration among community organizations, governments, funders, researchers, standards bodies, and technology providers.

Next Steps

This baseline report is the foundation for an annual program of inquiry, advocacy, and collaboration. Future work will broaden and deepen CODI's understanding of the digital landscape through crowdsourced contributions from the digital access community, structured research partnerships with university and NGO institutions, and a lightweight, web-based resource catalog making supply-side initiatives visible and searchable.

CODI invites last-mile organizations, funders, technology providers, and policymakers to engage with the findings of this report, work to close gaps, and join the growing coalition committed to building an Internet where every person can fully participate in their own language.

Appendices

A. Glossary of Key Terms

Term	Definition
Demand Side	The perspective of last-mile organizations and the communities they serve — what they need for meaningful linguistic participation online.
Supply Side	The ecosystem of global programs, policies, and technologies designed to advance digital development and inclusion.
Fallback Failure	A moment when a user is forced to abandon their own language and operate in another language because the digital system they are using does not support their language at that point of interaction.
Last-Mile Organizations	Civil society groups, community initiatives, research institutes, and technology organizations working directly with underserved language communities at the edge of the digital ecosystem.
Linguistic Operability	The condition in which a language is fully functional across the complete range of digital interactions — from script rendering to safety communications — so that users can complete any common digital task without being pushed into a language they did not choose.
Linguistic Presence	The measurable existence of a language across the layers of the Internet.
Underrepresented Languages	Languages that are functionally absent or marginally present in the global digital ecosystem, regardless of the number of speakers. Most of the world's approximately 7,000 languages are underrepresented online.
Universal Acceptance (UA)	The technical standard and practice of ensuring that all Internet domain names (including internationalized domain names), email addresses, and scripts are fully accepted, processed, and displayed correctly by all Internet-enabled systems.

B. Program and Technology Reference

The following table provides a summary reference for the supply-side programs and technologies discussed in this report, organized by the Four-Pillar Strategic Framework.

Program / Initiative	Owner / Source	Pillar
UK Digital Access Programme (DAP)	UK Government	1: Infrastructure & Access
Nigeria National Broadband Plan	Nigeria	1: Infrastructure & Access
BEAD Program	USA	1: Infrastructure & Access
Giga (School Connectivity)	UNICEF / ITU	1: Infrastructure & Access
MOSIP (Digital Identity)	Open-source	1: Infrastructure & Access (DPI)
World Bank ID4D	World Bank	1: Infrastructure & Access (DPI)
Universal Acceptance (UA)	UASG / ICANN	1: Infrastructure & Access
DigiGirls	CyberSafe Foundation / UK	2: Digital Skills & Literacy
Skills for Inclusive Digital Participation (SIDP)	British Council	2: Digital Skills & Literacy
Digital Mediators	France	2: Digital Skills & Literacy
ITU Digital Transformation Centres	ITU	2: Digital Skills & Literacy
Essential Digital Skills Framework	UK	2: Digital Skills & Literacy
Google.org Digital Skills & AI Literacy	Google.org	2: Digital Skills & Literacy
Microsoft AI for Good / Digital Skills	Microsoft	2: Digital Skills & Literacy / 4: Innovation
Meta Connectivity & Digital Literacy	Meta	2: Digital Skills & Literacy / 4: Innovation
Universal DPI Safeguards Framework	UN / UNDP	3: Digital Governance
Mojaloop (Payments)	Open-source	3: Digital Governance (DPI)
Pix (Brazil Payments)	Central Bank of Brazil	3: Digital Governance (DPI)
X-Road (Data Exchange)	Estonia / Open-source	3: Digital Governance (DPI)
GovStack	Global Initiative	3: Digital Governance

Program / Initiative	Owner / Source	Pillar
Cybersecurity Capacity Maturity Model	UK / Partners	3: Digital Governance
LGPD	Brazil	3: Digital Governance
NDPR	Nigeria	3: Digital Governance
AI for Development	UK Government	4: Innovation & Transformation
No Language Left Behind	Meta AI	4: Innovation & Transformation
Google Language AI / Low-resource Models	Google Research	4: Innovation & Transformation
RapidPro	UNICEF	4: Innovation & Transformation
AI for Good	ITU	4: Innovation & Transformation
OpenG2P	Open-source	4: Innovation & Transformation
Digital Public Goods Alliance (DPGA)	DPGA Secretariat	4: Innovation & Transformation

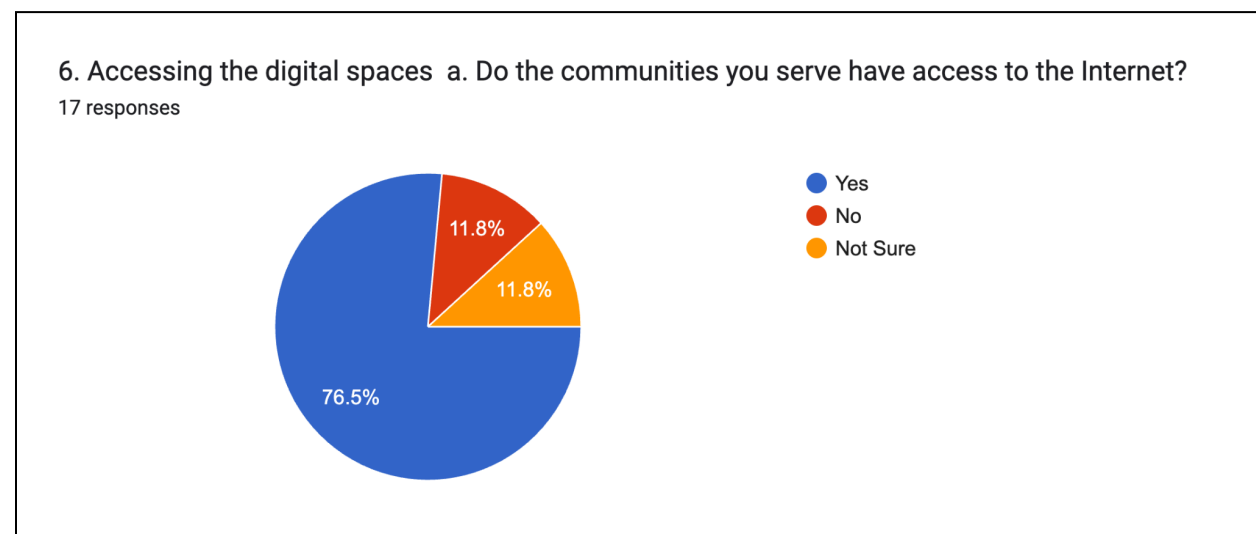
C. Survey Respondents List

Organization	Region	Type
Aenbee	India	Technology
Amitel India Pvt Ltd.	India (NCR)	Infrastructure Provider
auDA — .au Domain Administration	Australia	Infrastructure Provider
CLARIN.SI	Slovenia / Central & Eastern Europe	Infrastructure Provider
CLEAR Global	Global	Civil Society / Advocacy
Digital Empowerment Foundation	India	Civil Society / Advocacy
EGIGFA	Ghana	Civil Society / Advocacy
Endangered Languages Project	US / Canada / Belgium	Civil Society / Advocacy
GNKS Consult BV	Netherlands	Policy
HakiDigital Initiative	Kenya	Community Initiative
Centre for European Languages	EU	Academic / Research
Humanities & Social Sciences Commons	Canada	Academic / Research
Hybrid Lynx	Global (6 countries)	Technology
Indigenous Connectivity Institute	Canada	Community Initiative
Internet Society Uganda	Uganda	Civil Society / Advocacy
Raynews	Nepal	Civil Society / Advocacy
Rise Afrika Consulting Firm	Ethiopia	Consulting

D. Survey Data / Charts

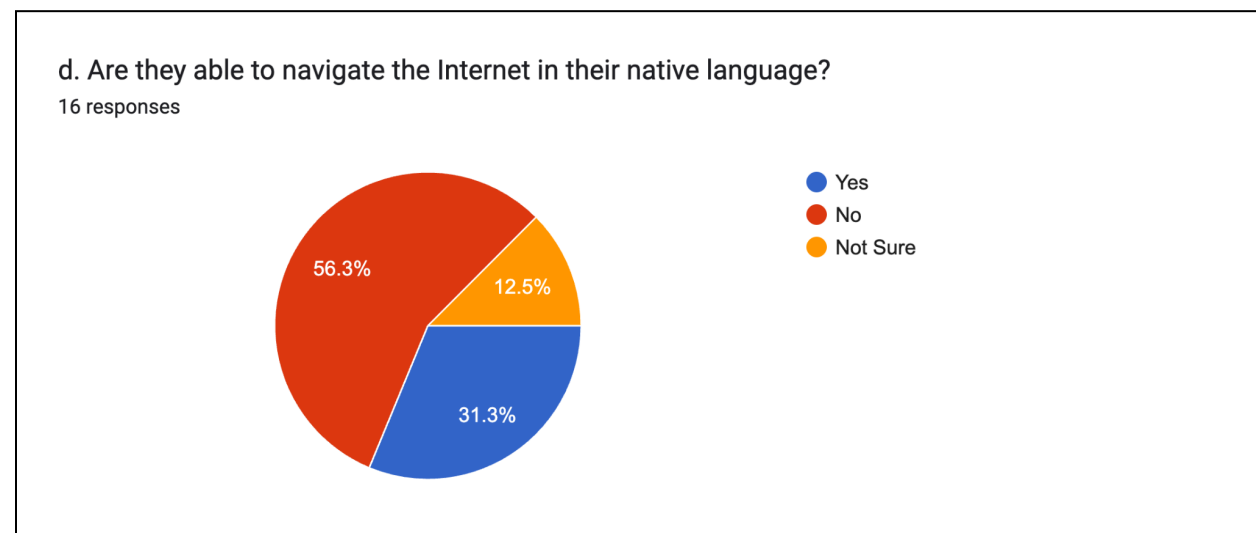
Accessing digital spaces

Nearly 25% of the respondents indicated that their communities had no access to the Internet or were unsure if they had access.



Navigating the Internet in one's native language

More than half of the respondents indicated that their communities couldn't navigate the Internet in their native language.

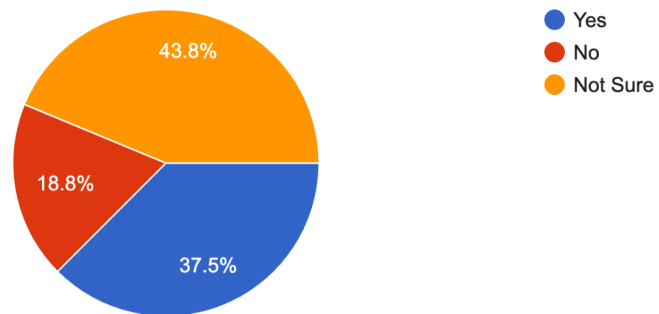


Using AI tools

Only a third of respondents thought their communities used AI tools.

f. Do they use artificial intelligence (AI) tools such as Google Gemini or ChatGPT?

16 responses



info@codi.global