Background paper

Digital and AI sovereignty: Four evidence-informed cases

This paper was commissioned by UNESCO and is part of UNESCO's work on AI governance in education. The views and opinions expressed in this paper are those of the author and should not be attributed to UNESCO.

This paper can be cited with the following reference: Benotti, L., Busaniche, B., Gómez, M. J. and Trovato, M. 2026. *Digital and AI sovereignty: Four evidence-informed cases*. Paris, UNESCO. Paper commissioned to support the development of policy briefs on AI governance in education.

For further information, please contact: luciana.benotti@unc.edu.ar

Luciana Benotti, Beatriz Busaniche, Marcos Javier Gómez and Margarita Trovato

Abstract

This paper examines the growing relevance of AI sovereignty, and what the concept means when applied to education. As achieving AI sovereignty relies on capacity building, educational priorities become central. To counter the risks of rapid obsolescence, unpredictable costs and cultural homogenization, the paper argues for building AI sovereignty in education by prioritizing foundational AI knowledge and local capabilities.

AI sovereignty is not a single idea. The paper distinguishes five dimensions of AI sovereignty in education: epistemic and pedagogical sovereignty; linguistic and cultural sovereignty; data sovereignty; infrastructural and fiscal sovereignty; and jurisdictional and technological sovereignty. and illustrates each through four cases.

- **Argentina, Córdoba:** A professional development course taught teachers the fundamentals of how AI works and how to evaluate it, without relying on commercial AI, transforming teachers into critical AI evaluators rather than passive users. It illustrates epistemic and pedagogical sovereignty.
- **Pan-African network:** The Masakhane grassroots research network combined hardware-efficient small language models with a community-led data licensing framework (the Esethu Framework) that protects and reinvests in local languages. It illustrates linguistic and cultural data sovereignty.
- **Brazil, PBIA:** A national AI plan that names sovereignty as a guiding principle, pairing investment in sovereign compute, the domestic computing infrastructure needed to train and run AI, with the development of Portuguese-language models grounded in national data, to reduce dependence on foreign infrastructure and models. It illustrates infrastructural, fiscal and linguistic sovereignty.
- **Kingdom of the Netherlands, GPT-NL**. A sovereign national language model built by public research and education bodies to reduce dependence on foreign providers and keep data within the national and European legal jurisdiction. It illustrates jurisdictional and technological sovereignty.

The paper closes with suggestions and questions for further deliberation, organized across national, regional and institutional levels of governance.

Table of contents

Context, relevance and purpose

Different framings of AI sovereignty in education

AI sovereignty is a recent concept. The concept draws on debates about digital and data sovereignty, which have gained prominence over the past decade, and on older concerns about technological dependency. But as a distinct concept focused on AI, it has only emerged in the 2020s, becoming sharpened by the concentration of AI models, data and compute in a small number of firms and countries.

Sovereignty was originally a concept of the nation state, and the digital environment unsettles that framework. Digital sovereignty has been defined as the control of data, software, standards and protocols, processes, hardware, services and infrastructure (Floridi, 2020). Building on this, Belli (2023) proposed the Key AI Sovereignty Enablers (KASE) framework, which treats AI sovereignty as the capacity of a country to understand, develop and govern AI systems, and to retain control, agency and self-determination over them. More recently, Jiang and Belli (2025) define digital sovereignty as the exercise of agency, power and control in shaping digital infrastructure, data, services and protocols.

From a cultural perspective, UNESCO frames the question in terms of cultural data sovereignty, noting that cultural data is a key source for training AI models. However, it remains largely unrecognized in policy, and its governance raises distinct questions about rights, representation and the protection of cultural and linguistic diversity (UNESCO, 2025*a*). Education is central to building and maintaining that capacity, which means AI sovereignty in education must be developed across several domains at once.

While these framings all share a concern with the principles of control and self-determination, they emphasize different aspects, such as infrastructure and data, the capacity to build and govern models, and cultural representation. It is precisely because AI sovereignty spans these domains that the term is invoked in very different ways, depending on who is using it and what they are trying to protect. For some, it means keeping data within a national legal jurisdiction; for others it means the capacity to build and run models instead of renting them. For others still, it is about language and culture, where knowledge and ways of speaking are represented in the AI systems that learners and teachers use. These framings overlap, but they are not the same and a policy that secures one, may leave the others untouched. A country can host its data locally and still depend entirely on a foreign model. It can build a national model and still teach learners only to use it instead of understanding it.

For education, it is useful to distinguish the dimensions of AI sovereignty explicitly. They are not mutually exclusive, and most systems will need to attend to more than one, but naming them clarifies what a given policy secures and what it leaves exposed. This paper identifies five and illustrates each with a case.

Epistemic and pedagogical sovereignty: the capacity of teachers and learners to understand how AI works, to evaluate its outputs, and to judge when and whether to continue to use it or rely on it. This is sovereignty held not by the state but by the people in the classroom, and it is the focus of the Córdoba case in Argentina.

Linguistic and cultural sovereignty: control over how a community's language, knowledge and ways of knowing are represented in AI systems, and protection against the homogenization that follows when systems are trained overwhelmingly on other languages and worldviews. UNESCO has developed this dimension in its work on cultural data sovereignty (UNESCO, 2025*a*). It is central to the pan-Africa Masakhane.

Data sovereignty: governance over how data, especially the data of the learners and communities, is collected, licensed, stored and used, and who benefits from it. UNESCO's report on Indigenous People-Centred AI in Latin America and the Caribbean calls for community autonomy over data and participation across all phases of AI development (González Zepeda and Martínez Pinto, 2023). This dimension is also illustrated by the Masakhane case, through the Esethu Framework.

Infrastructural and fiscal sovereignty: the capacity to invest in and control the compute, models and infrastructure a system depends on, rather than renting them from an external provider. This aims to mitigate the risks of loss of control, recurring costs and dependence that renting entails. This is illustrated in the Brazil case.

Jurisdictional and technological sovereignty: the capacity to build and operate AI under a system's own legal and institutional control, keeping data and models within its own jurisdiction rather than being dependent on infrastructure governed by others' laws. This is illustrated in the Kingdom of the Netherlands case.

Different education systems face different constraints, so they prioritize different dimensions. A system with scarce compute and many under-resourced languages may begin with linguistic and data sovereignty; a well-resourced system concerned with legal control may invest in jurisdictional sovereignty; and a system exposed to volatile procurement costs may focus on infrastructural independence.

The cases that follow do not represent a single model to replicate. Each shows a different dimension of sovereignty being built in a unique context and under different constraints. Together they map the range of what AI sovereignty in education can mean.

Why this matters now

Across the Global South, universities and education systems increasingly rely on commercial AI-enabled platforms. These systems are developed within specific epistemic, linguistic and regulatory contexts, embedding assumptions about originality, relevance, quality and authority (Ahmed, 2026).

> *'The new AI race has revitalized the discourse on digital sovereignty. AI education is now a core feature of national security, workforce competitiveness and cultural sovereignty. This framing elevates AI from a tool of innovation to an instrument of geopolitical power, and places education, skills and capacity building at the heart of this strategic landscape'*
> - Hamadeh and Amin (2025)

UNESCO's *Recommendation on the Ethics of Artificial Intelligence* reaffirms that integrating AI into education must respect international law and national sovereignty regarding data governance. Consequently, Member States hold a distinct mandate to regulate the protection of personal data and enforce the right to privacy according to established human rights standards (UNESCO, 2022).

The question then is no longer only whether a system can access AI, but whether it can cultivate the local and regional capacity to develop, deploy and govern AI. This places weight on sovereign infrastructure, on the capacity to build, attract and retain talent, while strengthening foundational AI education by teaching the core principles of how it functions. As the United Nations (2026) Independent International Scientific Panel on AI observes: 'Countries that rely on foreign models, cloud infrastructure and data pipelines may gain access to AI while losing practical control over its standards, safeguards and local fit.' The challenge is multifaceted, ranging from near-immediate obsolescence to the delegation of cognitive and human capabilities.

Against this background, the paper reviews how educational experiences build the foundations of sovereignty within the learning environment. This includes:

- capacity-building in model evaluation
- appropriate infrastructure
- the development of the local ecosystem
- independence from any single external provider.

This analysis is indispensable because major commercial AI providers are rapidly gaining ground on state capacity in education, proposing the integration of their models directly into the classroom.

How different systems are building AI sovereignty

Argentina: Commercial AI and highly digitalized languages

Local realities: In 2024, a professional development course on AI engaged more than 300 secondary school teachers (Ivetta et al., 2025). The initiative emerged in response to teachers' concerns about students' use of generative AI. Using the EDIA platform, participants explored conceptual foundations of AI models and audited large language models (LLMs) in the local sociolect of Spanish spoken in the province of Córdoba. Spanish is a highly digitalized language, but the Córdoban sociolect has lexical, morphological and syntactical markers that make it recognizable. The course fostered a critical and participatory understanding of AI stereotypes among teachers, contributing to more reflective perceptions of AI (Gómez et al., 2025).

A collaborative benchmark emerged as a byproduct of the course, called HESEIA, to explore bias and stereotypes in artificial intelligence (Ivetta et al., 2025). The benchmark, which encodes 40,000 sentences in Córdoban Spanish, contains intersectional stereotypes that commercial AI fails to reject and self-debias — thus exposing the limits of current commercial AI safety mechanisms for underrepresented sociolects of highly digitalized languages. A similar finding emerged when the benchmark was extended to 18 Latin American countries (Ivetta et al., 2026). This is not an isolated finding but another example of previously documented cases of commercial AI's deeply embedded biases against the sociolects of historically marginalized language communities (Hofmann et al., 2024).

AI infrastructure used: The initiative relies on EDIA, an open platform designed for participatory AI auditing, collaborative dataset creation and educational activities. EDIA is hosted on local infrastructure provided by Córdoba's High Performance Computing Center (CCAD). As a result, model interactions, educational activities and generated data remain on local servers until they are validated by peer teachers.

Sovereignty insights: This experience illustrates how AI sovereignty in education can start by developing the capacity of teachers to understand, evaluate and question AI systems. It also shows that meaningful AI education is not about learning to use commercial AI. In this case, teachers were able to critically evaluate model responses, recognize AI biases and contribute to evaluation benchmarks. Rather than centring curricula around commercial tools, education should foreground concepts of computer science and AI that provide the conceptual basis for understanding current and future AI systems. These include concepts such as algorithms, data, machine learning, digital networks and the internet.

International and regional initiatives have already begun defining curricula and activities grounded in these concepts. These include UNESCO's *AI competency framework for students* (Miao et al., 2024), UNESCO's *AI competency framework for teachers* (Miao and Cukurova, 2024), Program.AR's curricular proposal (Dabbah et al., 2024), and Ceibal's AI proposal for teachers (Kolezar et al., 2026). These initiatives prioritize foundational knowledge of AI systems, critical thinking, human agency, ethics and competencies that remain relevant as technologies change.

Pan-African network: Commercial AI and minimally digitalized languages

Local realities: Marivate et al. (2025) highlight that commercial AI severely underperforms when evaluated on low-resource African languages, due to massive data imbalances in global internet corpora. In their words, '[commercial] AI systems fail to support the many African learners who use indigenous languages daily.' Low-resource languages lead to critical gaps in AI safety guardrails (Hu et al., 2025) and increased costs (Ahia et al., 2023) compared to highly digitalized languages, like English and Spanish. They also point out that 'In South African classrooms, learning is multilingual, cultural and interactive. Teachers draw on songs, play and oral traditions to nurture curiosity … Ignoring these expressions marginalizes non-Western ways of knowing.' As Jiang et al. (2025) argue, commercial AI begins to dictate the linguistic, cultural and historical frameworks, and risks cultural homogenization.

AI infrastructure used: To counteract this epistemic exclusion, the pan-African Masakhane grassroots research network[1] operationalized a localized pipeline that aligns data governance directly with model architecture. A key technical milestone is InkubaLM, an open-source, decoder-only Small Language Model (SLM). Containing 400 million parameters and trained on 2.4 billion tokens, InkubaLM is engineered specifically to execute multilingual tasks and handle natural code-switching within constrained, local server or laptop environments.

Sovereignty insights: This hardware-efficient architecture is supported by the Esethu Framework, a community-centric data curation and licensing model (Rajab et al., 2025). The Esethu Framework replaces extractive data-mining with transparent documentation, local linguistic stewardship and a legal mechanism that transforms commercial licensing fees into direct financial reinvestment for community data creators. In Africa, indigenous languages are spoken by most of the population while colonial languages are second languages. The situation is different in Latin America, where, over five centuries, Spanish and Portuguese colonizers imposed European languages across administrative, religious and educational systems. The digitalization of Latin American indigenous languages is even more complex than in Africa, due to the low percentage of the population that speaks them (González Zepeda and Martínez Pinto, 2023).

Brazil: Sovereign infrastructure and national language models

Local realities: Brazil has made AI sovereignty an explicit goal of national policy. The Brazilian Artificial Intelligence Plan (PBIA) 2024–2028, coordinated by the Ministry of Science, Technology and Innovation, names sovereignty as one of its three guiding principles, alongside inclusion and sustainability. It commits around US$4 billion to AI sovereignty (MCTI, 2025), most of which comes from state sources (UNESCO, 2025*b*). Education and workforce development sit at the centre of the plan, which funds AI training and scholarships.

Infrastructure used: Sovereignty as defined in the PBIA has two linked components. The first is infrastructural: a national supercomputer and sovereign public computing capacity. This reduces reliance on a foreign cloud infrastructure and mitigates the risks of loss of control, recurring costs and

[1] See https://www.masakhane.io

overall dependence that renting it would entail (UNCTAD, 2024). The second is linguistic: the development of large-scale language models in Portuguese, trained on national datasets that reflect Brazil's cultural, social and linguistic diversity. The explicit aim is to reduce dependence on foreign data and models (UNESCO, 2025*b*; MCTI, 2025). A wider Brazilian ecosystem of Portuguese-language models, from research efforts to national initiatives, has grown since 2023 for the same reason: global multilingual commercial models currently underserve Brazilian Portuguese, and dependence on foreign foundations carries cost and licensing constraints that limit local use (Pires et al., 2023; Freitag, 2025).

Sovereignty insights: Brazil connects to two dimensions of sovereignty that are often treated separately. Its infrastructural and fiscal investment, in sovereign compute, is paired with a linguistic cultural one, in Portuguese-language models grounded in national data. The reasoning is that controlling the models a system depends on, while ensuring they represent the national language and its varieties, is part of the same sovereignty. The scale of the investment also carries a life cycle obligation: sovereign infrastructure requires sustained maintenance, updating and cybersecurity if it is not to become underused or outdated. The emphasis in this paper is different: Brazil illustrates how a large system can pursue sovereignty not only over its infrastructure but over the language its models speak and the culture they reproduce.

Kingdom of the Netherlands: sovereign public infrastructure and legal jurisdiction

Local realities. Sovereignty concerns are not confined to systems with few resources. Well-resourced systems face a different version of the same dependency: reliance on a small number of foreign commercial providers for the models, compute and cloud infrastructure on which public services increasingly run. Crucially, the data and legal control happen outside the national jurisdiction. The Kingdom of the Netherlands has responded to this by investing in sovereign public AI capacity rather than relying solely on external providers.

Infrastructure used. GPT-NL[2] is a sovereign Dutch large language model, built by public research and education bodies, the applied-research organization TNO, the research-and-education IT cooperative SURF, and the Netherlands Forensic Institute, with funding from the Ministry of Economic Affairs. The model is trained on legally sourced data on national computing infrastructure and is designed to meet European data-protection and AI-regulation requirements in-house, rather than through a commercial vendor (Barbereau and Dom, 2024). Differently from other so called open GPTs, the training data is open and redistributed under a CC-BY license (Van Oort et al., 2026).

Sovereignty insights. The explicit purpose of this project is to reduce dependence on non-European providers and to keep data and models within the Dutch and European legal jurisdictions.[2] The public investment is modest by commercial standards (Loohuouis, 2026), and that is part of the point: for a well-resourced system, the question is less about whether it can afford commercial AI than whether it wishes to retain sovereign control over the infrastructure its education and public services depend on. Building this capacity and meeting regulatory obligations directly rather than through a vendor is a way of paying now to preserve the ability to decide later.

[2] See https://www.tno.nl/en/digital/artificial-intelligence/gpt-nl

Suggestions for policy action

These suggestions are recommendations for how to operationalize the key insights emerging from the cases across three levels of governance: national, regional and institutional. Each has a different scope of action.

The national level

Build sovereign infrastructure and the capacity to use it, at a realistic scale
The concentration of computing power is clear: in 2025, an estimated 75 per cent of the compute across 500 of the largest AI clusters was in the United States of America, 15 per cent in the People's Republic of China and 10 per cent in the rest of the world (United Nations, 2026). Faced with this imbalance, different education systems have a choice between renting infrastructure from external providers, which carries the risk of the loss of control, recurring cost and dependence, and investing in sovereign capacity of their own. Brazil's plan shows the second path at national scale, pairing a national supercomputer with Portuguese-language models. The Kingdom of the Netherlands approach shows this path for a smaller system building a sovereign public model to keep data and models under its own jurisdiction.

The lesson is not that every system should build frontier infrastructure, which only a few are able to. It is more that sovereign capacity should be pursued at a realistic scale: compute suited to local needs, the human expertise to build and adapt smaller models, and regional cooperation and shared compute, where national capacity is out of reach. This is illustrated in the Argentina case by the country's CCAD hosting small AI within a Latin American network. Recent analysis suggests that training usable, auditable and locally aligned models is fiscally feasible for middle-income countries when timelines and hardware are chosen pragmatically, rather than to compete at the global frontier (Malagon et al., 2025).

Favour open and small models where resources are constrained
Open and small models present two feasible strategies, particularly in low-resource settings, as the Masakhane research shows for African languages and the Córdoba case in Argentina shows for teacher training and the Kingdom of the Netherlands is an example of not only opening weights but also data. Their adaptation for low-resource languages remains challenging, but their development is a mechanism for learning how they are developed in order to improve epistemic sovereignty.

Require vendor-agnostic procurement and the right to exit
Procurement policy should treat AI models as interchangeable components rather than permanent commitments. This is to ensure that education budgets are not exposed to a single provider's price changes or market shifts, and that a system has the built-in ability to change course. It is thus necessary to specify data portability, interoperability and the terms of exit at the point of purchase, and not after dependence has set in. Guidance resources can draw on the distinctions made by the Independent International Scientific Panel on AI (United Nations, 2026) between proprietary, open-weight, open-source and open-development models — with each carrying different degrees of third-party control and capacities to mitigate misuse.

The regional level

Sovereignty at the regional level sometimes comes from pooling capacity across borders, instead of relying only on in-country resources particularly where the challenges are regional in nature. Examples are language variation, system capacity or evaluation infrastructure.

Align data governance at the community, national and regional levels
Meaningful sovereignty is achievable when data governance aligns with community stewardship. Frameworks that replace extractive data practices with transparent documentation, linguistic stewardship and direct reinvestment in data creators, such as the African Esethu Framework case, offer a replicable model for localized ecosystems that reduce high-compute dependencies, consistent with states' obligations to protect personal data and privacy, and to preserve cultures underrepresented in commercial systems.

Build shared benchmarks and auditing infrastructure regionally
Evaluation infrastructure can be built outside large commercial labs, as the teacher-student HESEIA benchmark shows. But the underlying challenge is regional: a language such as Spanish varies across countries and no single national benchmark covers it. With 118 countries absent from major AI governance discussions (and mostly in the Global South) (United Nations, 2026), pooling resources across a linguistic region is often more realistic than each country working alone. Ministries could support extending initiatives like HESEIA into shared, cross-country benchmarks.

The institutional level

Adopting educational policies aligned with the principles of AI sovereignty requires multi-level coordination. The professional development of teachers and researchers is essential, as they bring AI literacy into the classroom. However, operational training alone is not enough. It is imperative to design benchmark curricula that integrate and prioritize an understanding of how AI works, its conceptual foundations and the diverse interdisciplinary fields that intersect with it.

Design curricula for deep understanding of AI and not only platform use
AI curricula should focus on durable foundational concepts. This foundational understanding should not depend on any specific commercial AI product (because those products change, and teaching to them ties learning to a vendor and dates quickly). The Córdoba case in Argentina shows the competencies that emerge when learning addresses the actual mechanics of AI systems: how models are pre-trained; how they represent meaning and why this produces hallucinations and stereotypes; how alignment can produce sycophancy; and how settings such as decoding temperature make outputs non-deterministic. Understanding these mechanisms enables learners to evaluate outputs, recognize limitations, and decide on task delegation. An example curriculum was implemented in Latin America and is freely available.[3]

Ground teachers' professional development in AI foundations
Teachers' professional development is essential for scaling AI literacy in classrooms and the development of AI competencies among both teachers and students, as analysed and recommended in UNESCO's twin *AI competency framework for teachers* (2024) and *AI competency framework for students* (2024). Priority should be given to conceptual understanding, critical evaluation and pedagogical strategies over platform-specific skills. Open tools running on local infrastructure make these approaches accessible without dependence on commercial ecosystems. As the Córdoba case showed, empowering teachers has a multiplier effect, extending critical AI literacy to thousands of students through everyday classroom practice.

[3] See https://ia.vialibre.org.ar/en/course-2026

Questions for further exploration

These are the key follow-up questions emerging from this paper:

- How should education systems weigh investment in domestic compute against regional cooperation and shared infrastructure, given the steep concentration of computer power?
- Are open and small models the most realistic path for systems with limited computational resources and low-resource languages, and what would it take to make them viable at scale?
- How can community, national and regional data governance be aligned, so that local linguistic and cultural stewardship is preserved instead of being overridden?
- How can regions pool resources to build shared benchmarks and auditing infrastructure, rather than each country attempting this alone?
- How can curricula and teacher development shift from teaching the use of particular tools towards a durable understanding of how AI works?

Conclusion

Sovereignty in education is less about owning the technology than it is about building the capacity to understand, evaluate and govern it. This is particularly in relation to: teachers and learners; data governance; infrastructure; and procurement. AI sovereignty is not one thing, and no system will secure every dimension at once. Indeed, the cases in this paper show different dimensions of sovereignty being built in different conditions and often outside the large commercial labs: teacher-led AI auditing in Argentina; community-governed language models across Africa; sovereign infrastructure and national language models in Brazil; and sovereign public capacity in the Kingdom of the Netherlands. What they share is a recognition that access to AI is not the same as control over it, and that the difference, for an education system is the difference between a genuine choice and structural dependency.

Further exploration: Links to practical tools and resources

1. https://ia.vialibre.org.ar/en/home/: AI Literacy Initiative by Fundación Vía Libre, including educational resources, courses, and activities for critical AI literacy.
2. https://edia.ngrok.app/: EDIA platform for participatory AI auditing, collaborative benchmark creation, and educational activities.
3. https://cursos.vialibre.org.ar/: Virtual learning environment of Fundación Vía Libre, including self-paced versions of teacher professional development courses and educational materials in Spanish.

Educational resources (in Spanish)

1. *Diez preguntas frecuentes y urgentes sobre Inteligencia Artificial* [Ten frequent and urgent questions about artificial intelligence]: https://program.ar/wp-content/uploads/2024/08/Diez-preguntas-frecuentes-y-urgentes-sobre-Inteligencia-Artificial.pdf
2. *¿Qué necesita la IA para ser IA?* [What does AI need to be AI?] - Educational activity exploring the foundations of AI systems: https://repositorio.curriculum.program.ar/repositorio-diseno-curricular/que-necesita-la-ia-para-ser-ia
3. *Un modelo de aprendizaje automático por dentro: entre datos y opinions* [A machine learning model from the inside: between data and opinions] - Educational resource explaining how machine learning models are built and how data shape their outputs: https://repositorio.curriculum.program.ar/repositorio-diseno-curricular/un-modelo-de-aprendizaje-automatico-por-dentro-entre-datos-y-opiniones/
4. *Hay una vaca en mi departamento: ¿entrené bien mi modelo de IA? Errores y sesgos* [There is a cow in my apartment: did I train my AI model well? Errors and biases] - Educational activity on training data, model errors, and bias in AI systems. https://repositorio.curriculum.program.ar/repositorio-diseno-curricular/hay-una-vaca-en-mi-departamento-entrene-bien-mi-modelo-de-ia-errores-y-sesgos.

References

Ahia, O., Kumar, S., Gonen, H., Kasai, J., Mortensen, D. R., Smith, N. A. and Tsvetkov, Y. 2023. Do all languages cost the same? tokenization in the era of commercial language models. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. Singapore, Association for Computational Linguistics, pp. 9904-9923. Available at: https://doi.org/10.18653/v1/2023.emnlp-main.614 (Accessed 18 August 2026.)

Ahmed, S.A. 2026. Algorithmic dependence and digital colonialism: a conceptual framework for artificial intelligence in education and knowledge systems of the Global South. *Frontiers in Education*, Vol 11. Lausanne, Frontiers Media SA. Available at: https://doi.org/10.3389/feduc.2026.1720563 (Accessed 17 August 2026.)

Barbereau, T. and Dom, L. 2024. *GPT-NL: Towards a Public Interest Large Language Model*. The Hague, TNO. Available at: https://resolver.tno.nl/uuid:0c3607e0-14fc-4ed6-abc2-731e07debb2b (Accessed 18 August 2026.)

Belli, L. 2023. Exploring the Key AI Sovereignty Enablers (KASE) of Brazil, to Build an AI Sovereignty Stack. L. Belli, and W.B. Gaspar (eds), *The Quest for AI Sovereignty, Transparency and Accountability*, Vol. 8. Cham, Springer.

Dabbah J., Fleitas D., Garzón M., Gómez M., Martínez M.C., Martínez López P. E. and Peretti, G. 2024. *Propuesta curricular para la inclusión de las Ciencias de la Computación en la educación obligatoria de la Argentina* [Curricular proposal for the inclusion of Computer Science in compulsory education in Argentina]. Buenos Aires, Fundación Sadosky. (In Spanish.) Available at: https://curriculum.program.ar/wp-content/uploads/2023/01/Program.ar_Propuesta-Curricular-para-la-inclusion-de-las-Ciencias-de-la-Computacion.pdf (Accessed 17 August 2026.)

Floridi, L. 2020. The Fight for Digital Sovereignty: What it is, and Why It Matters, Especially for the EU. *Philosophy & technology*, Vol. 33, No. 3. Berlin, Springer Nature, pp. 369-378. Available at: https://doi.org/10.1007/s13347-020-00423-6 (Accessed 17 August 2026.)

Freitag, R. 2025. *Linguistics for a Brazilian Artificial Intelligence (AI)*. São Paulo, SciELO - Scientific Electronic Library Online. Available at: https://blog.scielo.org/en/2025/07/18/linguistics-for-a-brazilian-artificial-intelligence-ai/ (Accessed 18 August 2026.)

Gómez, M. J., Dabbah, J. and Benotti, L. 2025. A workshop on artificial intelligence biases and its effect on high school students' perceptions. *International Journal of Child-Computer Interaction*, Vol. 43. Amsterdam, Elsevier.

González Zepeda, L. E. and Martínez Pinto, C. E. 2023. *Inteligencia artificial centrada en los pueblos indígenas: perspectivas desde América Latina y el Caribe* [Indigenous People-Centered Artificial Intelligence: Perspectives from Latin America and the Caribbean]. Paris, UNESCO. (In Spanish.) Available at: https://unesdoc.unesco.org/ark:/48223/pf0000387814 (Accessed 17 August 2026.)

Hamadeh, S. and Amin, H. 2025. AI, education and digital sovereignty. *Frontiers in Education*, Vol. 10. Lausanne, Frontiers Media SA, p. 1677727. Available at: https://doi.org/10.3389/feduc.2025.1677727 (Accessed 17 August 2026.)

Hofmann, V., Kalluri, P. R., Jurafsky, D., and King, S. 2024. AI generates covertly racist decisions about people based on their dialect. *Nature*, Vol. 633. Berlin, Springer Nature, pp. 147-154. Available at: https://doi.org/10.1038/s41586-024-07856-5 (Accessed 17 August 2026.)

Hu, Y., Hee, M. S., Nakov, P. and Lee, R. K.-W. 2025. *Toxicity Red-Teaming: Benchmarking LLM Safety in Singapore's Low-Resource Languages*. Stroudsburg, Association for Computational Linguistics. Available at: https://aclanthology.org/2025.emnlp-main.612 (Accessed 27 August 2026.)

Ivetta, G., Gomez, M. J., Martinelli, S., Palombini, P., Echeveste, M. E., Mazzeo, N. C., Busaniche, B. and Benotti, L. 2025. HESEIA: A community-based dataset for evaluating social biases in large language models, co-designed in real school settings in Latin America. *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*. Suzhou, Association for Computational Linguistics, pp. 25106-25128. Available at: https://doi.org/10.18653/v1/2025.emnlp-main.1275 (Accessed 17 August 2026.)

Ivetta, G., Palombini, P., Martinelli, S., Gomez, M. J., Echeveste, M. M., Dev, S., Prabhakaran, V. and Benotti, L. 2026. *Adaptive Data Collection for Latin-American Community-sourced Evaluation of Stereotypes (LACES)*. Stroudsburg, Association for Computational Linguistics. Available at: https://aclanthology.org/2026.findings-acl.203/ (Accesssed 17 August 2026.)

Jiang, M. and Belli, L. 2025. Introduction: Contesting Digital Sovereignty: Untangling a Complex and Multifaceted Concept. M. Jiang and L. Belli (eds), *Digital Sovereignty in the BRICS Countries: How the Global South and Emerging Power Alliances Are Reshaping Digital Governance*. Cambridge, Cambridge University Press, pp. 1–38. Available at: https://doi.org/10.1017/9781009531085 (Accessed 17 August 2026.)

Jiang, L., Chai, Y., Li, M., Liu, M., Fok, R., Dziri, N., Tsvetkov, Y., Sap, M. and Choi, Y. 2025. Artificial Hivemind: The Open-Ended Homogeneity of Language Models (and Beyond). *Advances in Neural Information Processing Systems 38*. San Diego, Neural Information Processing Systems Foundation, Inc. Available at: https://doi.org/10.52202/085713-2732 (Accessed 18 August 2026.)

Koleszar, V., Capdehourat, G., Curi, M. E., Valente, J., Paolillo, S., Barrios, I., Pereiro, E., Seo, J., Yoon, S., Park, J., Shin, T. S., Park, J., Kwon, K., Hwang, J. W., Oh, E. and Moon, K. 2026. *IA para enseñar, enseñar IA. Marco Referencial Docente sobre Inteligencia Artificial* [AI for Teaching, Teaching AI: A Teacher's Framework on Artificial Intelligence]. Montevideo, Ceibal. (In Spanish.) Available at: https://documentos.ceibal.edu.uy/portal/2025/11/Ceibal-Marco-IApara-ensenar.pdf (Accessed 17 August 2026.)

Loohuois, K. 2026. *Netherlands moves GPT-NL from lab to live: first pilots under way*. London, Computer Weekly. Available at: https://www.computerweekly.com/news/366642524/Netherlands-moves-GPT-NL-from-lab-to-live-first-pilots-under-way (Accessed 18 August 2026.)

Malagon, S., Ulloa Ruiz, M. A., Sandoval Plaza, T. E., Bolívar, G. R. R., García Mesa, V. and Alvarado Morales, I. 2025. *The Feasibility of Training Sovereign Language Models in the Global South: A Study of Brazil and Mexico*. Ithaca, arXiv. Available at: https://doi.org/10.48550/arXiv.2510.19801 (Accessed 18 August 2026.)

Marivate, V., Zondi, N.C. and Masikisiki, B. 2025. Ensuring inclusive, contextualized AI in education: Considerations towards a roadmap. *AI and the future of education: Disruptions, dilemmas and directions*. Paris, UNESCO, pp. 121-124. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000395236 (Accessed 18 August 2026.)

MCTI. 2025. *IA para o Bem de Todos: Plano Brasileiro de Inteligência Artificial* [AI for the Common Good: Brazilian Artificial Intelligence Plan]. Brasilia, Ministério da Ciência, Tecnologia e Inovação (MCTI). (In Portuguese.) Available at: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/plano-brasileiro-de-inteligencia-artificial-pbia-_vf.pdf (Accessed 18 August 2026.)

Miao, F. and Cukurova, M. 2024. AI competency framework for teachers. Paris, UNESCO. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000391104 (Accessed 17 August 2026.)

Miao, F., Shiohira, K. and Lao, N. 2024. *AI competency framework for students*. Paris, UNESCO. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000391105 (Accessed 17 August 2026.)

Pires, R., Abonizio, H., Almeida, T. S. and Nogueira, R. 2023. *Sabiá: Portuguese large language models*. Ithaca, arXiv. Available at: https://doi.org/10.48550/arXiv.2304.07880 (Accessed 18 August 2026.)

Rajab, J., Aremu, A., Chimoto, E. A., Dunbar, D., Morrissey, G., Thior, F., Potgieter, L., Ojo, J., Tonja, A. L., Chetty, M., Nekoto, W. N., Moiloa, P., Abbott, J., Marivate, V. and Rosman, B. 2025. The Esethu Framework: Reimagining Sustainable Dataset Governance and Curation for Low-Resource Languages. *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics*, Vol. 1. Vienna, Association for Computational Linguistics, pp. 30763–30776. Available at: https://doi.org/10.18653/v1/2025.acl-long.1487 (Accessed 18 August 2026.)

UNCTAD. 2024. *Brazil launches the Brazilian Artificial Intelligence Plan 2024-2028*. Geneva, UNCTAD. Available at: https://investmentpolicy.unctad.org/investment-policy-monitor/measures/4930/launches-the-brazilian-artificial-intelligence-plan-2024-2028 (Accessed 18 August 2026.)

UNESCO. 2022. *Recommendation on the Ethics of Artificial Intelligence*. Paris, UNESCO. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000381137 (Accessed 17 August 2026.)

—. 2025*a*. *UNESCO global report on cultural policies, Culture : the missing SDG*. Paris, UNESCO. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000395504 (Accessed 17 August 2026.)

—. 2025*b*. *Brazil: readiness assessment report on artificial intelligence*. Paris, UNESCO. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000393091 (Accessed 18 August 2026.)

United Nations. 2026. *Preliminary Report of the Independent International Scientific Panel on AI: Evidence-based assessment of opportunities, risks and impacts of artificial intelligence*. New York, United Nations Publications. Available at: https://www.un.org/independent-international-scientific-panel-ai/en/preliminary-report (Accessed 17 August 2026.)

Van Oort, J. J., Brinkkemper, F., de Graaf, E., Vanroy, B. and Lensink, S. 2026. GPT-NL public corpus: A permissively licensed, Dutch-first dataset for LLM pre-training. *Proceedings of the Fifteenth Language Resources and Evaluation Conference*. Palma de Mallorca, ELRA Language Resource Association, pp. 6757–6768. Available at: https://doi.org/10.63317/5fbtc336wwx2 (Accessed 18 August 2026.)

UNESCO – a global leader in education

Education is UNESCO's top priority because it is a basic human right and the foundation for peace and sustainable development. UNESCO is the United Nations' specialized agency for education, providing global and regional leadership to drive progress, strengthening the resilience and capacity of national systems to serve all learners. UNESCO also leads efforts to respond to contemporary global challenges through transformative learning, with special focus on gender equality and Africa across all actions.

The Global Education 2030 Agenda

UNESCO, as the United Nations' specialized agency for education, is entrusted to lead and coordinate the Education 2030 Agenda, which is part of a global movement to eradicate poverty through 17 Sustainable Development Goals by 2030. Education, essential to achieve all of these goals, has its own dedicated Goal 4, which aims to ***"ensure inclusive and equitable quality education and promote lifelong learning opportunities for all."*** The Education 2030 Framework for Action provides guidance for the implementation of this ambitious goal and commitments.

unesco
United Nations
Educational, Scientific
and Cultural Organization

Education 2030

Stay in touch

aied@unesco.org

https://www.unesco.org/en/digital-education

UNESCO
7, place de Fontenoy
75007 Paris
France

Follow @UNESCO on social media

ED/FLI/ICT/2026/05

4 QUALITY EDUCATION

Sustainable Development Goals