POLICY BRIEF

Transparency for Trust: Bridging the Global Divide in AI Governance

2025

Prof. Cristina Godoy Bernardo Oliveira, University of São Paulo, Executive Committee, Center for AI and Machine Learning (CIAAM) (Brazil)
Prof. João Paulo Candia Veiga, University of São Paulo, Executive Committee, CIAAM, (Brazil)
Prof. Glauco Arbix, University of São Paulo, PI, C4AI USP-IBM-FAPESP, CIAAM (Brazil)
Prof. Christoph Burchard, Goethe University, Frankfurt am Main; Founding Speaker, Center for Critical Computational Studies (C3S) (Germany)
Prof. Juliane Engel, Goethe University, Frankfurt am Main; Founding Director for Transfer, C3S (Germany)
João Victor Gianecchini, Scholarship Holder, CIAAM (Germany)
Prof. Eduardo Saad-Diniz, University of São Paulo (Brazil)

02

Digital Transformation

Abstract

Research and development cooperation is crucial for advancing global AI governance. AI policy has shifted from broad principles to sector-specific regulations, necessitating new regulatory experimentation and accountability frameworks. This policy brief builds on the T20 Brazil 2024 Communiqué and addresses two critical dimensions: the Global Governance Monitoring System (GGMS) and the Transparency and Inclusivity Index for AI Systems (TIAIS). These mechanisms operationalise key T20 recommendations by fostering accountability, reducing inequalities, and enhancing public trust in AI. The GGMS, modelled on the Intergovernmental Panel on Climate Change, will be an independent multilateral body monitoring AI governance frameworks, societal impacts, and regulatory compliance. This system will leverage multi-stakeholder expertise, ensuring representation from underrepresented regions, particularly the Global South, as the São Luís Declaration highlighted. The TIAIS will create a standard to evaluate AI system transparency, inclusivity, and fairness, focusing on bias mitigation and cultural diversity in large language models.

Introduction

Global governance of artificial intelligence (AI) is predominantly shaped by a small group of countries in the Global North, whose institutions and private actors determine the terms of debate, set regulatory priorities, and define technical standards. As AI adoption accelerates across sectors and geographies – from health and education to justice and public services – many countries in the Global South remain structurally excluded from global norm-setting processes.[1] This exclusion is reflected in their limited representation in international governance forums and in the frequent transplantation of regulatory models that disregard local infrastructure, cultural norms, and socio-technical realities. These imported frameworks often fail to align with lower-income countries' legal systems, languages, and digital capacities, reinforcing epistemic asymmetries and governance dependence.[2]

Stark disparities in digital infrastructure, data access, and human capital exacerbate the unequal distribution of AI's benefits. As the G20 AI Readiness and Capacity Assessment Toolkit (2024) and UN E-Government Survey (2024)[3] highlight, regions such as Africa, Oceania, and small island developing states (SIDS) remain critically under-resourced. They are unlikely to close this gap by 2030 without targeted investments. Meanwhile, the demographic size and lower labour costs in the Global South make these regions vulnerable to data extraction

[1] Chinmayi Arun, "AI and the Global South: Designing for Other Worlds," in *The Oxford Handbook of Ethics of AI*, ed. Markus D. Dubber, Frank Pasquale, and Sunit Das (Oxford: Oxford University Press, 2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3403010.

[2] Igarapé Institute, *Global Futures Bulletin: Responsible Artificial Intelligence Efforts in the Global South* (Rio de Janeiro: Igarapé Institute, 2024),

https://igarape.org.br/wp-content/uploads/2024/01/Global-Futures-Bulletin-Responsible-Artificial-Intelligence-Efforts-In-the-Global-South.pdf.

[3] United Nations Department of Economic and Social Affairs. *E-Government Survey 2024: Accelerating Digital Transformation for Sustainable Development, With the Addendum on Artificial Intelligence*. New York: United Nations, 2024. https://publicadministration.un.org/en/egovkb/en-us/.

and experimental AI deployment, with little say in the design or regulation of these technologies. This imbalance creates a moral dilemma and a strategic threat to the legitimacy, sustainability, and global accountability of AI governance.[4]

A growing body of evidence points to the absence of standardised, transparent mechanisms for evaluating fairness, accountability, and inclusivity in AI systems, particularly in high-impact domains such as generative AI (GenAI).[5] According to UNESCO's 2023 global guidance on GenAI in education,[6] the release of ChatGPT exposed the regulatory vacuum surrounding these technologies: by mid-2023, only one country had enacted specific rules addressing GenAI, leaving public institutions, including schools, exposed to ethical, legal, and data privacy risks. Likewise, the OECD's 2023 Principles implementation report notes that most governance tools, including bias audits and ethical checklists, lack independent validation, risking "ethics washing" and undermining public trust.[7]

Regulatory fragmentation is also evident within multilateral institutions. The 2024 report on using AI across the UN system identifies a lack of inter-agency coordination and limited capacity for knowledge-sharing, resulting in siloed experimentation and missed opportunities for scaling responsible AI to advance the Sustainable Development Goals (SDGs).[8] Furthermore, the absence of shared

[4] Luke Munn, "The Uselessness of AI Ethics," *AI and Ethics* 3 (2023): 869–877, https://doi.org/10.1007/s43681-022-00209-w.

[5] Royal Society Open Science, "Article rsos.231994" (2024), https://royalsocietypublishing.org/doi/pdf/10.1098/rsos.231994.

[6] Fengchun Miao and Wayne Holmes, *Guidance for Generative AI in Education and Research* (Paris: UNESCO, 2023), https://unesdoc.unesco.org/ark:/48223/pf0000386693.

[7] UNESCO, *Global Education Monitoring Report 2023: Technology in Education – A Tool on Whose Terms?* (Paris: UNESCO, 2023), https://unesdoc.unesco.org/ark:/48223/pf0000388089. See also OECD, *The State of Implementation of the OECD AI Principles: Four Years On* (Paris: OECD, 2023), https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/10/the-state-of-implementation-of-the-oecd-ai-principles-four-years-on_b9f13b5c/835641c9-en.pdf.

[8] United Nations System, *Report on the Operational Use of AI in the UN System*. New York: Chief Executives Board for Coordination, High-Level Committee on Management, Task Force on the Use of Artificial Intelligence in the United Nations System, 2024. https://unsceb.org/sites/default/files/2024-11/Report%20on%20the%20Operational%20Use%20of%20AI%20in%20the%20UN%20System.pdf.

transparency metrics and monitoring mechanisms has led to inconsistent adoption, particularly in lower-income settings.

In parallel, the capability gap between high-income and low-income countries widens. As noted by the World Bank (2022)[9] and OECD GSG Report (2024),[10] this gap limits the Global South's participation in innovation and research and deepens dependencies on foreign technologies and regulatory infrastructures. Without structural interventions and inclusive governance instruments, the digital divide may evolve into a governance divide, preventing billions of people from shaping AI's ethical and developmental trajectories.

Furthermore, less than 10% of global AI governance frameworks originate in the Global South, highlighting an imbalance in defining global AI ethics. These issues necessitate globally coordinated, locally adaptable frameworks for monitoring AI deployment. Without standardised metrics or global monitoring systems, disparities will continue undermining legitimacy and equitable distribution.

Recommendations

This policy brief proposes five strategic actions to strengthen global AI research and development (R&D) governance. These recommendations address the need for regulatory innovation, responsible data sharing, and privacy-by-design approaches to digital transformation.

[9] Sharmista Appaya and Jeremy Ng, *Global Trends in AI Governance: Evolving Country Approaches* (Washington, DC: World Bank Group, 2024),

https://documents1.worldbank.org/curated/en/099120224205026271/pdf/P1786161ad76ca0ae1ba3b1558ca4ff88ba.pdf

[10] Organisation for Economic Co-operation and Development (OECD), *Futures of Global AI Governance: Co-Creating an Approach for Transforming Economies and Societies. GSG Background Note GSG(2024)1en* (Paris: OECD Global Strategy Group, 2024), https://www.oecd.org/about/secretary-general/global-strategy-group/.

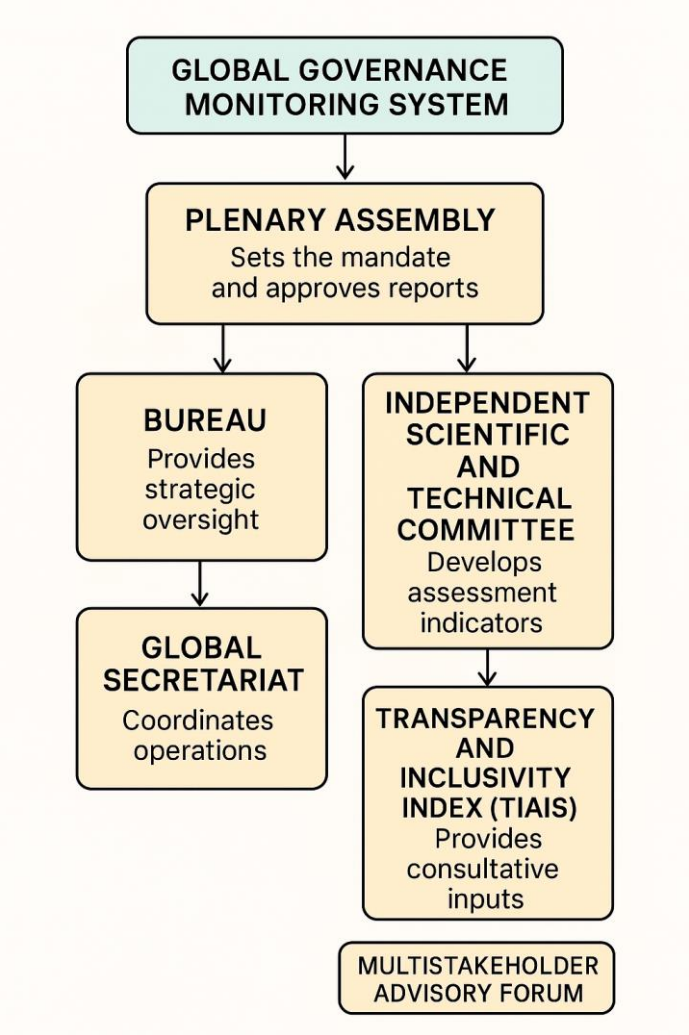
1. Establish a Global Governance Monitoring System

The Global Governance Monitoring System (GGMS) is a multilateral mechanism to assess AI governance frameworks, impacts, and compliance, inspired by the Intergovernmental Panel on Climate Change model. It would be an independent observatory to evaluate national and regional AI governance systems' maturity, effectiveness, and inclusiveness, ensuring multistakeholder and Global South representation.

Implementation:

- Convene a G20 expert working group to define the GGMS mandate, structure, and funding model.
- Engage UN agencies, academic networks, and regional digital coalitions to develop participation criteria through inter-agency coordination and capacity pooling, minimising duplication and maximising social benefit.
- Develop common indicators and scorecards for AI regulatory maturity and compliance.
- Publish annual monitoring reports and thematic alerts (e.g., on algorithmic harms, bias, etc.).

Figure 1. GGMS governance structure



Source: Compiled by author

2. Create a transparency and inclusivity index for AI systems

A Transparency and Inclusivity Index (TIAIS) is a standardised global metric to evaluate AI system fairness, bias mitigation, and cultural responsiveness. The TIAIS would be co-developed with stakeholders from the Global South to ensure relevance and legitimacy across contexts.

Implementation:

- Launch a public–private–academic task force to design and pilot the TIAIS, which will be integrated by experts from the Global North and South equally.
- Test the metric on large language models (LLMs) and other foundational models.
- Incorporate multilingual and intersectional dimensions (eg, gender, race, disability)
- Align the index with existing tools (eg, OECD AI Tools, NIST RMF) to ensure interoperability.

3. Foster regulatory experimentation and joint sandboxes

Accelerate responsible AI innovation through collaborative, cross-jurisdictional regulatory learning. Regulatory sandboxes allow safe testing of AI models and applications under monitored conditions, generating evidence for future rules and standards.

Implementation:

- Establish a G20 platform for sharing sandbox protocols, lessons learned, and outcomes.
- Fund regional sandboxes in the Global South with technical assistance and risk mitigation support.
- Include underrepresented stakeholders (civil society, micro, small and medium-sized enterprises [MSMEs], academia) in sandbox design.
- Use sandbox results to inform adaptive, risk-based regulation.

4. Advance privacy-preserving and responsible data governance

Support R&D and policy frameworks that balance data access with human rights. Data sharing and AI development must respect privacy, consent, and local data governance traditions, especially in sensitive domains like health, education, and financial services.

Implementation:

- Promote global cooperation in developing privacy-enhancing technologies (PETs).
- Build local capacity in data stewardship, anonymisation, and secure data infrastructure.
- Harmonise cross-border data standards aligned with G20 and UN digital compacts.
- Encourage data commons initiatives to improve Global South participation in AI R&D.

5. Institutionalise AI governance in global development agendas

Ensure that AI governance becomes a pillar of sustainable development and inclusion.

By mainstreaming AI governance into agendas like the SDGs, Global Digital Compact, and climate action, the G20 can anchor its leadership in long-term global goals.

Implementation:

- Embed AI governance indicators in the Voluntary National Reviews of SDG progress.

- Align GGMS and TIAIS metrics with UN human rights instruments and development priorities.
- Fund South–South and triangular cooperation on AI governance.
- Encourage multilateral development banks to invest in governance infrastructure.

6. Domestic action agenda for the Global South

To bridge the AI governance divide meaningfully, countries in the Global South must complement global cooperation mechanisms with ambitious and context-specific domestic strategies. These include structural, regulatory, and educational reforms designed to foster innovation, safeguard rights, and reduce dependencies on dominant AI powers.

- **Build inclusive and contextual AI strategies:** National AI plans must reflect local socio-economic priorities and avoid one-size-fits-all regulatory imports. Lessons from Brazil, Peru, Chile, Kenya, and India show the value of risk-based and human rights-centred AI frameworks tailored to each country's institutional maturity and strategic sectors.
- **Institutionalise regulatory experimentation via sandboxes:** Governments should implement regulatory sandboxes and regional testbeds, enabling safe experimentation with AI applications while collecting evidence on their social and ethical impact. These controlled environments are advantageous where formal regulatory systems are still developing. A South–South Sandbox Forum could allow countries to coordinate methodologies, share lessons, and reduce duplicative efforts.
- **Strengthen data infrastructure and sovereignty:** AI development depends on access to reliable, secure data and computing power. Investments in local data centres, public digital infrastructure, and open data frameworks are critical. Countries should also promote data stewardship models that ensure community participation, privacy, and equitable benefit-sharing.

- **Foster local talent and prevent brain drain:** AI ecosystems require more than foreign investment. Governments must support the training and retention of researchers, engineers, and policy professionals by integrating digital skills into primary and secondary education, offering scholarships, and creating local R&D incentives.
- **Promote participatory and inclusive governance:** National AI councils or advisory bodies should include government, industry, academia, civil society, and marginalised communities. This inclusive design helps counter "algorithmic colonisation" and ensures AI systems reflect local norms and needs.
- **Expand strategic cooperation and knowledge access:** To close the AI capability gap effectively, Global South countries should adopt a dual approach combining horizontal cooperation with assertive engagement in global knowledge-sharing. Strengthening South–South collaboration – through joint regulatory experimentation, shared datasets, and co-developed governance tools – can promote contextually appropriate and regionally driven solutions. At the same time, it is essential to push for greater openness from advanced AI powers, including China, the US, and the EU, by advocating for responsible technology transfer, transparency in safety research, and inclusive access to infrastructure and frontier models. This hybrid strategy fosters mutual learning and distributed innovation, enabling Global South countries to become adopters and co-creators of safe and trustworthy AI systems.

Table 1 summarises the main recommendations, positive outcomes, and possible trade-offs of the proposals suggested in this brief.

Table 1. Summary of recommendations, outcomes, and trade-offs

Recommendation	Positive outcome	Contradiction/trade-off
GGMS – Global Governance Monitoring System	- Creates a multilateral, evidence-based accountability mechanism - Elevates Global South voices in AI governance debates - Enables early detection of governance gaps and algorithmic risks	- Risk of political resistance to monitoring by sovereign states - Disputes over indicators and evaluation criteria may politicise outcomes - May require long time horizons for credibility and adoption
TIAIS – Transparency and Inclusivity Index	- Provides global benchmarking for fairness and cultural responsiveness - Incentivises developers to mitigate bias in LLMs - Promotes inclusive AI across languages, identities, and contexts	- Standardisation may overlook context-specific values and uses - Risks being adopted as a PR tool without real reform ("transparency washing") - Tensions between transparency and proprietary models
Regulatory experimentation and joint sandboxes	- Accelerates learning on safe AI deployment across jurisdictions - Enables flexible regulation based on real-world evidence - Strengthens South-South and cross-border regulatory collaboration	- Sandboxes may become 'innovation privilege zones' without inclusive design - Results may not scale to formal regulation if not well integrated - Risk of overreliance on untested models for high-stakes AI use
Privacy-preserving and responsible data governance	- Empowers data sharing with trust and safeguards - Advances global equity in AI research using sensitive data - Builds foundational infrastructure for ethical AI innovation	- PETs are technically complex and costly; LMICs may struggle to adopt them - Privacy vs utility trade-offs: stricter controls may hinder innovation - Inconsistent global norms may delay cross-border adoption
Institutionalise AI governance in global development agendas	- Aligns AI governance with SDGs, climate action, and digital equity - Encourages donor alignment and multilateral funding for inclusive tech - Elevates AI governance in national planning and reporting	- Risk of AI governance being diluted in overly broad agendas - National development plans may lack capacity to implement indicators - Possible duplication with existing international frameworks

Conclusion

AI has emerged as a transformative force but, without inclusive, transparent, and accountable governance, it risks deepening existing inequalities. The G20 has a unique opportunity to lead in shaping a global governance model that balances innovation with rights and sustainability.

By implementing the GGMS and TIAIS mechanisms and enabling regulatory learning, data governance, and development integration, the G20 can shift from fragmented, principle-based AI policies to operational frameworks rooted in fairness, adaptability, and global solidarity. These recommendations are not merely technical but also political, strategic, and urgent.

The time to act is now. Delaying global coordination will only widen the AI divide, erode public trust, and entrench digital neocolonialism. The G20 must seize this moment to lead responsibly – and inclusively – into the AI future.

T20 South Africa Convenors

The Institute for Global Dialogue (IGD)

The South African Institute of International Affairs (SAIIA)

The Institute for Pan-African Thought and Conversation (IPATC)

For publication enquiries, please contact t20@t20southafrica.org

Website: www.t20southafrica.org