Task Force 05
INCLUSIVE DIGITAL TRANSFORMATION

Bridging the AI Governance Divide

Gordon LaForge, Senior Policy Analyst, New America (USA)

Robert Muggah, Co-Founder, Igarapé Institute (Brazil)

Gabriella Seiler, Consultant, Igarapé Institute (Brazil)

TF05

Abstract

This policy brief examines the global gap in responsible AI frameworks and the risks that emerge when AI policies and practices developed primarily for the Global North are exported to the Global South, where socioeconomic context is different and regulation and infrastructure tend to be less advanced. This is based on a review of around 500 AI safety regulations, and guidelines formulated by regional organizations, governments, private companies, and civil society. The brief assesses how asymmetries in this global ecosystem affect and will affect developing countries, both in terms of potential harms and opportunity costs of failing to enable the development of homegrown AI capabilities. Finally, the brief recommends three practical steps the G20 can take to narrow the global AI governance divide and bring about a safer, more equitable AI ecosystem.

Keywords: AI gap; AI governance; equitable AI; global South in AI

Introduction

At a time of deepening global competition over AI and the microchips, data centers, and the critical raw minerals that power it, the world experienced a rare glimmer of multilateral cooperation. On 24 March 2024, the UN General Assembly adopted an unprecedented resolution to promote safe, secure and trustworthy AI. Backed by the US and China and co-sponsored by 120 other states, the resolution underlined the importance of respecting, protecting and promoting human rights in the design, development, and deployment of AI. The resolution also stressed how AI must be human-centric and be applied in a way that is inclusive and advances sustainable development. Though non-binding, the resolution sent a signal that the international community recognizes the urgency of ensuring safe and secure AI.

While promising, the high-level statement of universal principles masks deep and consequential divisions in how different nations, companies, and non-governmental organizations are approaching the governance of AI. Despite incipient national efforts to develop cutting-edge AI systems, AI design and development are overwhelmingly concentrated in the US, China, and a handful of Western European countries. In its 2024 Artificial Intelligence Index Report, the Stanford Institute for Human-Centered Artificial Intelligence found that of the 109 "notable" machine learning models that have been most influential in the AI ecosystem, 61 were created in the United States, 25 in Western Europe, and 15 in China. Only two were created in a developing country (Egypt). Similarly, over the last five years not a single AI foundation model has originated from a developing economy. The concentration of this emerging technology that some say will be more transformative than electricity (and even fire) in the rich world stands to widen global inequality. In a survey of 60-plus chief economists conducted by the World

Economic Forum at the end of 2023, 94% said AI would improve productivity in high-income countries in the next five years, while only 53% said the same would be true in poorer countries.

Just as there is a vast disparity in AI development, so too is there a sharp divide in AI governance between the rich world and the developing world. The US, China, and the European Union have emerged as the leading standard-setters for AI governance, including as it relates to safety and security. Whether it is China's model of binding state oversight of AI development, the US laissez-faire approach that allows technology companies to self-govern, or the European Union's comprehensive AI Act based on mitigating user risk, developing countries are apt to adopt regulatory frameworks from the rich world. The EU is a particularly influential inspiration for lawmakers and regulators abroad, a phenomenon Anu Bradford has called the "Brussels Effect."

High-income countries are also leading and setting the agenda for international AI governance efforts, such as the May 2024 AI safety summit in Seoul hosted by the UK and South Korea. Various subnational governments, regional organizations, multilateral bodies, civil society networks, and corporate groupings are developing proposals and policies for safe and responsible AI governance. These too are overwhelmingly concentrated in China, Europe, and the US. New research from the Igarape Institute and New America has identified over 470 sets of AI principles promulgated by governments, companies, and civil society organizations across more than 60 countries from 2011 to 2023. The vast majority of these rules and standards were produced in wealthy countries where AI use is more prevalent and the debate and alarm over AI has crescendoed. Nearly two-thirds of these documents were created in Europe, the US, and China. Only 5 percent were developed in Latin America and the Caribbean. Just 2 percent came from Africa.

Understanding the AI Governance Divide

While the digital divide is well understood, the widening digital governance divide, particularly as it applies to AI, is not. Part of the reason for this is the sheer speed at which AI technologies are advancing. While AI has been in use for decades, generative AI only went mainstream in the past two years. Policy makers are still learning about the technology and debating how best to govern it. But it is already clear that the rules that will shape the trajectory of a transformational technology are being set by a small number of nations. The disparity in rule-setting may prove as consequential as asymmetries in AI capabilities, for even those populations not creating or directly accessing AI tools will be affected by AI, and thus the governance frameworks that shape it. This has profound implications for which AI risks are addressed, what harms might occur, and whether the technology will be widely and safely accessible in a way that mitigates global inequalities.

Unless developing countries have power to shape AI governance, the technology will end up reflecting the national and commercial interests of wealthy nations, often to the detriment of countries and societies with less power and fewer resources. That could mean developing countries are less able to develop homegrown AI ecosystems that advance sustainable development, or that dislocations and harms ensue from adopting imported tools and policies ill-suited to local circumstances. The potential consequences are not confined to lower- and middle-income countries alone: they are global. Labor displacement, government abuses of human rights, widening economic inequality, political destabilization – these and other factors will drive conflicts and migration that will not stay neatly confined within national borders.

There are divergences in the AI governance priorities of the large, standard-setting jurisdiction. China, which was an early mover in establishing binding restrictions on AI

development and use and promulgated its latest rules for generative AI in August 2023, aims to support generative AI development while at the same time require companies producing advanced AI to obtain prior government approvals before they can be released to the wider public lest they threaten state control or undermine "core socialist values". Despite a 2023 White House Executive Order that imposes some, though at this stage vague, restrictions on advanced frontier models as well as safeguards to limit synthetic content, the US government has refrained from constraining the activities of AI companies, instead allowing leaders such as Amazon, Anthropic, Google, Meta, Microsoft, and OpenAI to commit to voluntary principles on AI safety, security, and trust.

But the largest, and perhaps most consequential gaps in priorities for AI governance are between these rich countries and the developing world. In general, the rich-world frameworks focus primarily on protecting against what they perceive to be misuses of AI or challenges to the political or socioeconomic status quo. These risks include violations of privacy, weaponization, loss of control, disinformation, and others. All are valid and important to address. But for the developing world, misuse of AI may be less of a risk than missed use. The digital divide is already a critical development issue and source of global inequality, with the 2.7 billion people who lack internet access overwhelmingly located in the developing world. According to the IMF, increasing internet penetration by 10 percent in sub-Saharan Africa could result in as much as a 4 percentage point increase in per capita GDP. Missing out on the adoption of AI applications in domains such as healthcare, education, agriculture, disaster response, and others will limit sustainable development and widen the gap between the upper-income and lower-income countries. The imperative for developing countries is to build out their AI ecosystems. But if the fear of misuse is the dominant precept in global AI governance, then policy frameworks

will emphasize control and containment, a paradigm that will favor large incumbents and hinder the adoption and development of AI systems and applications by smaller players.

At the same time, tools that are built in the rich world and exported to developing countries without regard for the local context and without local frameworks for safe and responsible use risk causing immediate harm to citizens. In Brazil, for instance, facial recognition technology and predictive policing tools developed in the US have been deployed by the city of São Paulo for public safety programs amid several concerns about whether the technology could exacerbate racial and social discrimination. So while the development of local AI companies and applications are critical, AI safety and AI innovation are not mutually exclusive. In fact, sustainable AI innovation depends on frameworks for safety and trust, otherwise economic displacement, political and social backlashes, and accidents could derail technological development.

While some general principles of safe and trustworthy AI are universal, such as human oversight, other specific standards and regulations developed in the rich world may be ill-suited for developing nations. To take but one example, in Kenya, lawmakers tabled a draft AI bill that included a requirement that AI companies would have to obtain a costly government license to release an AI model, a popular idea among US lawmakers and one promoted by large American AI companies, which would benefit from such a barrier to entry. The bill provoked a backlash among the Kenyan startup community, which rightly argued that the licensing requirement would stifle homegrown AI innovation and development.

Further, many AI-related dislocations are more urgent and consequential in developing countries. Labor market disruptions in upper-income countries are at this point speculative; in lower-income countries they have already arrived. Consider recent revelations about the mistreatment and abusive working conditions endured by Kenyan

data annotators — some of the millions of low-wage workers who perform the tedious and at times traumatic work of labeling the training data ingested into large AI models. Data labeling is a key component in the AI supply chain that is outsourced to Global South nations and that could potentially employ a workforce of more than a billion in the coming years. Yet principles for safe and fair AI labor outsourcing are scarce in the global AI policies and discussions underway.

Recommendations for Bridging the Global AI Governance Divide

The G20, especially with Brazil and soon South Africa holding the presidency, is uniquely positioned to help narrow the global AI governance divide. Its members are the most influential countries in the Global North and Global South, representing 85 percent of the world's economic output. As such, the G20 should do three things to support an AI global governance ecosystem that is equitable and inclusive.

- One, it should seek to negotiate a statement on AI governance goals that reflects the priorities of Global South members and stakeholder groups.

- Two, the G20 should convene national ICT and other ministers responsible for overseeing the digital economy to discuss strategies for building shared policy frameworks and increasing policymaking capacity.

- Three, the G20 should create an Equitable AI Development Forum that enables the joint development of AI capabilities among nations, knowledge-sharing, and the distribution of shared AI infrastructure such as datasets among countries.

First, the G20 can directly narrow the AI governance divide by putting forth a declaration on AI governance standards that represent the priorities of Global South members. These standards need not be binding to have an important normative and agenda-setting role. Part of the declaration could call on jurisdictions to develop AI policies, principles, and standards that reflect global best practice but are also tailored to local realities. It could also be the start of an informed discussion within the G20 and elsewhere about sovereign AI, including oversight of data pools, and the risks posed by

fragmented governance in a geopolitically volatile world. Importantly, the process within the G20 for developing these standards should be inclusive, involving private and civic stakeholders from within member-states.

Second, in the same way that finance ministers gathered under the G20 in the wake of the Global Financial Crisis in 2008, the digital-focused ministers of the G20 countries should gather to discuss how best to advance global cooperation on AI policymaking. A focal point of these discussions could be how to build public oversight, governance capacity, and literacy for responsible AI policymaking. AI is not a single technology, but rather a collection of several foundational ones that will affect portfolios ranging from national defense to healthcare to labor. That means single, centralized AI agencies within governments might be inadequate to properly oversee all aspects of the technology, especially in larger countries. Civil servants and officials across government will need to become versed in responsible AI. International organizations, think tanks, universities, and government-to-government learning exchanges can help design and deliver such public AI literacy efforts.

Finally, for developing countries to both avert some of the risks of AI and to have a bigger say in global AI governance, more needs to be done to develop homegrown AI capabilities and ecosystems. The G20 can help do this by proposing an Equitable AI Development Forum, or eventually an Observatory, that involves countries from the Global North and Global South. Such a forum could facilitate a range of activities. At a minimum, it could be a place where countries could band together to invest in shared datasets and compute capacity for training and running AI models.

Governments might also work together under the Forum to design regulatory sandboxes to encourage innovation and experimentation with new technologies. That might also include facilitating North-South and South-South university and researcher

networks – such as the Artificial Intelligence for Development Africa – that could help share knowledge, resources, and talent for AI research. Governments could also use the Forum to advance initiatives to tap into diaspora communities for AI talent, such as is being pursued by Nigeria's Ministry of Communications, Innovation, and Digital Economy. There are also considerable and as yet untapped resources for data-rich countries such as India, Indonesia, Nigeria, South Africa and others to leverage their data pools, though any of these efforts must be pursued with data protection and privacy top of mind.

The Forum should also involve private companies, leveraging both technical skills and resources. Given constraints on national budgets, public-private partnerships could help finance and realize AI development goals. Governments could consider removing barriers to foreign investment for AI development that fosters local technological capacity building and take steps to de-risk investments to crowd in blended financing to drive "safer" and "more secure" AI as a comparative advantage. Governments could partner with philanthropic foundations to pool AI resources in a manner similar to how Gavi, the Vaccine Alliance procures and distributes vaccines to poor populations.

Conclusions

Ultimately, a safe and beneficial AI future will depend on ensuring the proceeds of the technology are shared and the risks of it are mitigated in an equitable way. Achieving that will depend on governance and international cooperation. A single global AI agency or treaty would be both insufficient and politically impossible. Rather, it will take a messy, multi-institution "regime complex" that comprises different governance functions, collaborations, and agreements. The G20 has an essential role to play in this ecosystem. It should take a bold position on AI that focuses on closing the global AI governance and capabilities divide. Doing so will help advance sustainable development and reduce the risk of harms and dislocations in both the Global South and the Global North.

References

Anthony, Aubra; Sharma, Lakshmee; Noor, Elina. "Advancing a More Global Agenda for Trustworthy Artificial Intelligence.", *Carnegie Endowment for International Peace*, April 30, 2024.

https://carnegieendowment.org/research/2024/04/advancing-a-more-global-agenda-for-trustworthy-artificial-intelligence?lang=en

LaForge, Gordon; Gruver, Patricia. "Governing the Digital Future.", *New America*, October 4, 2023.

https://www.newamerica.org/planetary-politics/reports/governing-the-digital-future/

Muggah, Robert. "When it Comes to Responsible Safe AI, a Global Divide is Emerging.", *Global Observatory*, April 11, 2024.

https://theglobalobservatory.org/2024/04/when-it-comes-to-responsible-and-safe-ai-a-global-divide-is-emerging/.

Muggah, Robert; Seiler, Gabriella; Laforge, Gordon. "AI and the Global South." *Project Syndicate*, March 2, 2023.

https://www.project-syndicate.org/commentary/ai-governance-first-principles-must-include-global-south-by-robert-muggah-et-al-2023-03.

Muggah, Robert; Szabó, Ilona. "Artificial Intelligence will Entrench Global Inequality." *Foreign Policy*, May 29, 2023. https://foreignpolicy.com/2023/05/29/ai-regulation-global-south-artificial-intelligence/?t pcc=recirc_latest062921.

Neaher, Giulia; Laforge, Gordon; Muggah, Robert; Seiler, Gabriella, "Responsible and Safe AI: A Primer for Policymakers in the Global South.", *Igarapé Institute,* April 2024.

https://igarape.org.br/wp-content/uploads/2024/04/SD_EN_Responsible-and-Safe-AI.pdf

Perrault, Ray; Clark, Jack. "Artificial Intelligence Index Report 2024.", *Stanford Institute for Human-Centered Artificial Intelligence*, 2024. https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_2024_AI-Index-Report.p df

G2O
BRASIL 2024
BUILDING A JUST WORLD
AND A SUSTAINABLE PLANET

T2O Brasil 2024
Let's rethink the world

Let's
rethink
the world

CEBRI
Rethink Tank

FUNDAÇÃO
ALEXANDRE
DE GUSMÃO

ipea
Instituto de Pesquisa
Econômica Aplicada