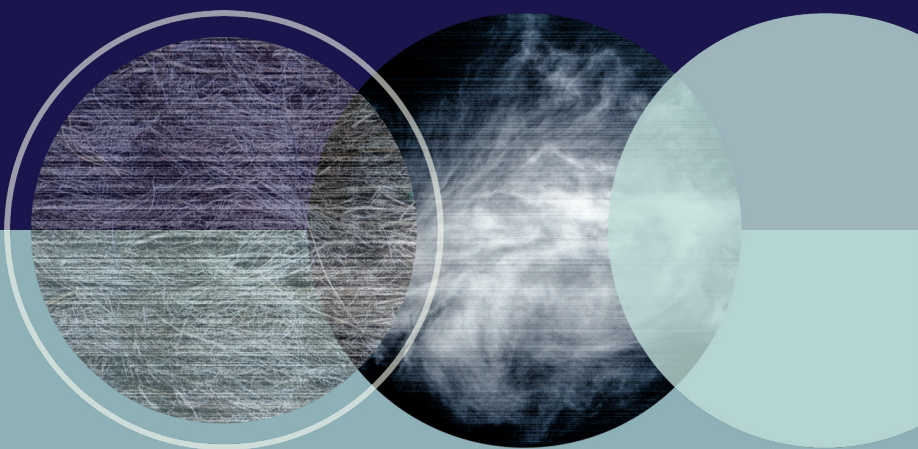




**Policy and Governance
Working Group**

September
2023

Recommendations to the United Nations on Global AI Governance



**SCIENCE, TECHNOLOGY,
AND SOCIAL VALUES LAB**
INSTITUTE FOR ADVANCED STUDY

30 September 2023

Mr. Amandeep Singh Gill
Office of the Secretary-General's Envoy on Technology
United Nations Headquarters
405 East 42nd Street
New York, NY, 10017
United States

submitted via email to: techenvoy@un.org

Response of the AI Policy and Governance Working Group to the Envoy on Technology Call for Papers on Global AI Governance

The [AI Policy and Governance Working Group](#), convened by the Institute for Advanced Study, submits this white paper in response to the United Nations Secretary-General's Envoy on Technology's Call for Papers on Global AI Governance.

Members of the AI Policy and Governance Working Group represent a mix of sectors, disciplines, perspectives, and approaches. Despite these differences, we agree that it is not only possible but necessary to address the multitude of concerns raised by the expanding use of artificial intelligence (AI) systems and tools and their increasing capabilities. We also agree that both present-day harms and neglected risks on the horizon warrant urgent attention in order to fulfill the public's legitimate expectation of safety and respect for their rights. We share the belief that these issues require immediate and ongoing action from industry, governments, academia, and civil society to meet public expectations.

The rapid worldwide deployment of advanced AI tools, systems, and applications underscores the need to establish agile, effective, global governance mechanisms and policies. Yet the specific parameters and features of any such global AI governance regime depend upon the specific outcomes sought. These include investing in AI research and development to deliver prosperity for people and the planet, especially with regard to progress toward the Sustainable Development Goals (SDGs); promoting monitoring and accountability in the use of AI to ensure national and international security; and developing accountability mechanisms and oversight of AI tools and systems to mitigate risk to public safety.

For nearly a decade, multilateral bodies have proposed principles and guidelines to steer the responsible use of AI. Individual countries, regions, and regulatory bodies are taking steps to create regulatory and legislative guardrails to guide and govern the development of AI. There is ample opportunity to make these processes more responsive to the harms and challenges facing the majority of the world's population that lives in the Global South. We remain optimistic that

global and domestic AI policy have the potential to be more democratic and inclusive than they have been so far, and emphasize that the needs of these communities should be foregrounded in policymaking.

The UN Secretary-General has an opportunity for bold leadership and to build upon these previously established guidelines and in-process legal developments, carrying forward existing work by pursuing deliberate action and robust policies and practices — and, especially, ones informed by meaningful input from those most affected by the development and deployment of AI tools and systems.

Although many singular solutions have been proposed, there is no singular institution that can “solve” the challenge of global AI governance. The governance of AI requires a multipronged approach addressing the broad potential application of advanced artificial intelligence, its acute risk to international and economic security, and the concrete effects and potential externalities of specific use cases. Therefore, this response addresses high-level considerations the Secretary-General and the Office of the Envoy on Technology should consider when weighing a global governance framework for advanced AI tools, systems, and applications. A fit-for-purpose global AI governance framework must 1) be nimble and quickly deployable by smartly leveraging the breadth of existing UN institutions, agencies, and initiatives, 2) coordinate and incentivize *policy interoperability* between countries, regions, and regulatory bodies, and 3) develop and utilize gap analyses when necessary.

Towards a Global AI Governance Framework

Leverage Existing Institutions, Agencies, and Initiatives Towards Actionable Guidelines and Concrete Outcomes. Any global AI governance framework must advance with prudent dispatch through a wide range of approaches. The governance challenge of requiring both agility and breadth means that it is incumbent upon the stakeholder community to accelerate existing approaches and leverage established institutions. The proliferation of disparate AI principles and frameworks risks fragmenting global governance efforts without establishing the necessary enforcement mechanisms. The primary challenge is not a deficit of principles, but a need for concrete steps that demonstrate progress toward operationalizing existing ones.

Instead of reinventing the wheel, we encourage the UN Tech Envoy and others to build upon the example of organizations like the OECD and Partnership on AI (PAI), which have made considerable strides in synthesizing AI principles into actionable guidelines toward common outcomes, and to build on existing initiatives such as the US-EU Trade and Technology Council, the Global Partnership on AI, the G7 and OECD processes, and the African Union AI Continental Strategy. This approach makes the most of established frameworks and will allow for a more coherent and effective international governance landscape for AI. A similar approach must be taken in deploying current multilateral institutions and UN agencies to maximize the potential benefits of AI technologies for all people and for the planet, and to address current and future risks and harms.

Building new institutions of global scale is a complex and time-intensive endeavor. While doing so may be warranted in some instances, this UN global AI governance process is also an opportunity for the international community to move with the urgency required of the moment. An immediate milestone should be a comprehensive review of existing international organizations that assesses their capacity to expand upon their current mandates for AI governance. Undue duplication of effort will make governance efforts less effective, and should be avoided, while existing entities with capacity to take on key challenges should be identified, resourced, and empowered.

Incentivize Policy Interoperability. The inherently global scope of the use of AI demands a governance strategy built on the principle of *policy interoperability*. This is not to advocate for globally uniform regulations, but rather for a coherent, mutually recognizable system akin to “APIs” in the digital realm, the technical pathways that allow different applications and systems to communicate and interoperate—an analogy to suggest effective interactions among diverse policy frameworks. Real-world examples include mutual recognition clauses in Free Trade Agreements; these create a pathway for reciprocity in technical assessment that can be adopted in global AI governance and also expanded from shared technical assessment to shared policy priorities. Existing standards-setting organizations can also play a role in facilitating technical aspects of policy interoperability.

To be clear, we are advocating for a governance structure that translates effectively across varying jurisdictional approaches. An interoperable governance framework would allow countries, regions, and regulatory bodies to work effectively together. Policy interoperability also enables jurisdictions to set their own policy priorities—in-line with local needs and the specific context relevant for determining thresholds for fairness, responsibility, and safety—while still aligning with a globally recognized set of core commitments and accountability and safety mechanisms. Such an approach could foster both technical and policy coherence—fulfilling the aspirations of the Global Digital Compact—and may promote effective, inclusive global governance for AI. These are adaptable to global AI governance and are also expandable from shared technical assessment to shared policy priorities.

Existing mechanisms, such as OECD's Environmental Performance Reviews, Due Diligence Guidance for Responsible Business Conduct, and Trade Facilitation Agreements, along with UN Periodic Reports, may provide actionable blueprints for this interoperable governance vision. We recommend exploring an approach modeled after the International Civilian Aviation Organization (ICAO). In much the same way as the ICAO establishes a global framework of minimum aviation safety standards and audits state compliance, a formally recognized Policy Interoperability Framework could serve a similar function for AI governance and outline mechanisms for mutual recognition of AI policies across jurisdictions, providing a high-level schema for the interface of diverse regulatory systems.

Augment Existing Efforts Where Possible; Use Gap Analyses to Identify Where New Efforts are Needed. Different AI issues necessitate distinct governance models, and many issues fall under the jurisdiction of existing bodies. While national bodies are best equipped for overseeing AI in national infrastructure, international oversight is imperative for areas such as security,

climate, pandemics, humanitarian aid, and equitable AI benefits. In line with our initial recommendation to leverage existing initiatives, principles, and institutions, we propose bringing a gap analysis mindset to any proposal to create new structures or initiatives. If the UN is considering establishing a new entity for global AI governance, it should have a theory of what gap it is filling. Should a systematic gap analysis be required—and in many instances, it may not be—the UN should catalog thematic gaps and overlaps in AI governance. Second, it should survey existing international and UN-affiliated organizations' initiatives within these thematic domains. Lastly, an assessment should follow to determine where the UN's unique expertise and mandates can offer the most novel, actionable contributions. This mindset can strategically guide the UN to areas where it can add the most value, complementing rather than duplicating existing efforts. We contend that the UN is particularly well-suited to address matters such as:

- international security, cybersecurity, and biosecurity;
- information integrity and threats to democratic processes;
- human rights protections, consumer protection, and privacy;
- economic development and security;
- climate, energy, and environmental sustainability; and ○
global health and pandemics.

This targeted focus not only refines the scope of global AI governance, but also necessitates the inclusion of specialized policymaking and field expert voices and professionals, thereby creating new constellations of expertise essential for nuanced, domain-specific governance.

At the same time, the needs of Global South countries must be prioritized and realized through leapfrogging technological bottle-necks, upskilling, and infrastructure development for collaboration with industry, governments, civil society, and academia. This can be accelerated through foreign direct investment, public-private partnerships, and levers such as advanced market commitments and responsible use of open-source resources.

No matter which particular institutional models for global AI governance gather momentum and find support, the conditions that surround these institutions will determine their ultimate efficacy. We offer these recommendations in that spirit.

Thank you for your consideration. Please contact aipolicy@ias.edu with any comments or questions.

Sincerely,

AI Policy and Governance Working Group*

Alondra Nelson

Dorothy Chou

Helen Toner

Iason Gabriel
Inioluwa Deborah Raji
Irene Solaiman
Jaan Tallinn
Janet Haven
Karine Perset
Marc Aidinoff
Marc-Etienne Ouimette
Miranda Bogen
Rumman Chowdhury
Sébastien Krier
Stephanie Ifayemi
Suresh Venkatasubramanian
William S. Isaac

*Members of the working group are participating in their personal capacities and these recommendations do not reflect the perspective of any of the organizations with which they are affiliated. Bios for the signatories can be found [here](#).

We thank Tatiana Carayannis, Karen Kornbluh, Alex Pascal, Marie-Therese Png, Maria Ressa, Olatunbosun Tijani, and Tim Wu for sharing their insights.