30 September 2023

His Excellency Amandeep Singh Gill
Office of the Secretary-General's Envoy on Technology
United Nations Secretariat, 27th floor
New York, NY 10017
Via email: techenvoy@un.org

Re: Global AI Governance and the United Nations

Your Excellency,

My name is Ryan Nabil, and I serve as the Director of Technology Policy and Senior Fellow at the National Taxpayers Union Foundation, a think-tank in Washington, DC, where I research US and international approaches to AI governance, data privacy, and emerging technologies.

Prior to NTUF, I worked as a Research Fellow at the Competitive Enterprise Institute and conducted research as a Fox Fellow at the Institut d'Études Politique de Paris (Sciences Po).

I have attached my paper entitled 'Global AI Governance and the United Nations: The UN Should Update Its Existing Institutional Framework Instead of Creating a Global AI Agency'.

As the High-Level Advisory Board on AI undertakes preparations for its inaugural meeting, I hope that the essay will provide helpful insights on how the United Nations can play a more effective and constructive long-term role in global AI governance.

Best wishes,
Ryan

Ryan Nabil
Director of Technology Policy and Senior Fellow
National Taxpayers Union Foundation
122 C St NW, Ste 700
Washington, DC 20001
rnabil@ntu.org
+1-703-683-5700

Global AI Governance and the United Nations
The UN Should Update Its Existing Institutional Framework Instead of Creating a Global AI Agency

By Ryan Nabil[1]
National Taxpayers Union Foundation
30 September 2023

I. Introduction

As national governments and world leaders reflect on the best methods to regulate AI, the United Nations faces growing calls to create a global agency for artificial intelligence.[2] However, given the distinct challenges that AI poses to different areas of global governance, a single, global AI regulator is unlikely to be effective in addressing such a wide variety of challenges. Instead, such an overly broad organisation will likely struggle to manage the competing strategic priorities and divergent political values of various member states. Instead of creating an international AI organisation without a specific mandate, the United Nations should consider launching agency-specific AI initiatives within the framework of its existing institutional architecture.

Since the commercialisation of generative AI applications like ChatGPT and Google Bard took off last autumn, policymakers in multilateral institutions and national governments have put forward new suggestions for global AI governance. One such suggestion that has enjoyed growing popularity is the idea of creating an international organisation for AI. Recently, the UN Secretary General António Guterres and Tech Envoy Amandeep Gill have both supported calls for the creation of a global agency to address future AI safety risks and promote international cooperation.[3] Such support comes against the backdrop of similar calls by the United Kingdom Prime Minister Rishi Sunak and OpenAI CEO Sam Altman to create an international AI agency.[4]

While some experts might be quick to dismiss global governance of AI,[5] it is an increasingly important and complex issue that merits closer examination and a more analytical, deliberative approach. As a starting point, the United Nations leadership and the High-Level AI Advisory Board must consider several related questions. First, what type of international organisations do world leaders mean when they advocate the creation of a global AI agency? Second, since international organisations are no monolith, what is a useful taxonomy of such organisations and their

---

[1] Ryan Nabil is the Director of Technology Policy and Senior Fellow at the National Taxpayers Union Foundation (NTUF), a think-tank based in Washington, DC. NTUF is located at 122 C St NW, Ste 700, Washington, DC 20001, USA. The author can be contacted at rnabil@ntu.org and +1-703-683-5700.

[2] W. Henshall, 'How the UN Plans to Shape the Future of AI', *Time* (21 Sep. 2023), https://time.com/6316503/un-ai-governance-plan-gill/.

[3] W. Henshall, 'The UN Plans to Shape AI'. B. Fung, 'UN Secretary General Embraces Calls for a New UN Agency on AI in the face of "potentially catastrophic and existential risks"', *CNN* (18 July 2023), https://www.cnn.com/2023/07/18/tech/un-ai-agency/index.html

[4] D. Martin, 'The Five Key Points on Rishi Sunak's Agenda for His Meeting with Joe Biden', *The Telegraph* (8 June 2023), https://www.telegraph.co.uk/news/2023/06/06/rishi-sunak-joe-biden-washington-ai-london-watchdog/. S. Altman, G. Brockman, & I Sutskever, 'Governance of superintelligence', OpenAI (22 May 2023), https://openai.com/blog/governance-of-superintelligence#GregBrockman.

[5] Cf. A. Thierer, *Existential Risks & Global Governance Issues Around AI & Robotics* (Washington, DC: R Street, June 2023), http://dx.doi.org/10.2139/ssrn.4174399.

institutional design? Third, as AI regulation is frequently compared to the regulation of nuclear energy, how do AI and nuclear energy differ, and what do these differences mean for the effective regulation of AI at the international level?[6] Finally, what are the precise concerns and objectives of the United Nations in distinct domains of AI applications, and what institutional frameworks are best suited to address these concerns?

A detailed discussion of these questions goes beyond the purview of this brief essay, and this analysis by no means constitutes an exhaustive list of questions on which the High-Level AI Advisory Board should deliberate. Instead, this essay seeks to contribute to the growing body of scholarship that provides analytical frameworks for addressing AI-related global governance challenges in distinct domains, such as the regulation of autonomous weapons, the development of technical standards, and international development.

II. Global AI Governance and Taxonomy of International Organisations

The High-Level AI Advisory Board would benefit from a more thorough assessment of the institutional functions and features of existing international organisations and evaluating possible institutional models for AI governance. Although a growing number of international leaders advocate the creation of a global AI agency, there appears to be considerable variation in the type of organisation they support. For example, Prime Minister Sunak has called for the creation of an international organisation like the European Organisation for Nuclear Research (CERN), an intergovernmental centre for research and cooperation in particle physics.[7] Meanwhile, other leaders and executives, such as Mr Altman, support the founding of an organisation like the International Atomic Energy Agency (IAEA) for AI regulation.[8] However, a CERN-like research centre would have a fundamentally different institutional function and design than an organisation modelled after the IAEA. Given such substantial differences, the United Nations would benefit from a more systematic approach to analysing and evaluating possible institutional models for AI governance.

A recent paper by prominent AI and international relations researchers at Google DeepMind, Oxford, Stanford, and Université de Montreal provides a useful starting point for such a discussion.[9] More specifically, the authors recommend four possible models for an international AI organisation: i) the Commission on Frontier AI, ii) Advanced AI Governance Organisation, iii) Frontier AI Collaborative, and iv) AI Safety Project (Table 1). The main challenge with this taxonomy is that it does not consider distinct aspects of AI governance (e.g., arms control, human rights, and trade policy) and the disparate institutional frameworks required to address domain-specific policy concerns. However, within the context of a specific domain, such as laws of war, the proposed framework can inform the debate about institutional models best suited to address AI-related governance challenges.

Evaluating the strengths and weaknesses of possible models can also help clarify whether the United Nations would provide the most suitable platform for carrying out proposed activities. For example, a CERN-like research centre ('AI Safety Project') would ideally provide advanced computing

---

[6] Altman, Brockman, & Sutskever, 'Governance of Superintelligence'.

[7] D. Martin, 'Rishi Sunak's Agenda for Meeting with Biden'.

[8] Altman, Brockman, & Sutskever, 'Governance of Superintelligence'.

[9] L. Ho et al., 'International Institutions for Advanced AI', *Computers and Society* (10 July 2023), https://doi.org/10.48550/arXiv.2307.04699.

resources and cloud platforms for AI collaboration between leading technology companies, universities, and governments. However, given that AI capabilities tend to be concentrated in a handful of countries and technology companies, such efforts might be more effective as an intergovernmental project by like-minded countries or as a consortium of technology companies and research institutions with funding from several governments. Instead, the United Nations' efforts might be more effective in areas like arms control initiatives and the development of AI-related technical standards, where the UN enjoys a comparative advantage and deep institutional expertise.

Table 1. Taxonomy of Possible Institutional Models for Global AI Governance

Proposed Model	Description	Example Organisations
1) Commission on Frontier AI (Intergovernmental)	— Establish scientific positions on AI risks and possible mitigation strategies — Improve public awareness and understanding of AI and related risks — Help governments improve AI expertise	— Intergovernmental Panel on Climate Change (IPCC) — Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) — Scientific Assessment Panel of the United Nations Environment Programme (SAP)
2) Advanced AI Governance Organisation (Intergovernmental or multi-stakeholder)	— Contribute to AI standardisation and the alignment of AI governance norms and standards — Assist in implementing AI standards and norms — Potentially perform compliance monitoring function (independently or in collaboration with private sector or industry groups) for a global AI governance regime	— International Civil Aviation Organisation (ICAO) — International Atomic Energy Agency (IAEA) — Financial Action Task Force (FATF)
3) Frontier AI Collaborative (International public private partnership)	— Promote access to advanced AI systems by creating an international public-private partnership — Help underserved countries and societies benefit from AI — Promote global access to advanced AI models and computing for AI safety and governance research	— Gavi (the Vaccine Alliance) — Global Fund to Fight AIDS, Tuberculosis, and Malaria — Atomic Development Authority after WWII
4) AI Safety Project (Private or public platforms)	— Bring together leading researchers and engineers for AI safety research — Provide access to AI models and computing resources for researching AI risk mitigation strategies — Potentially promote collaboration between researchers and the private sector	Scientific collaborations, including: — European Organisation for Nuclear Research (CERN) — International Thermonuclear Experimental Reactor (ITER)

Source: L. Ho et al. (2023)[10]

---

[10] L. Ho et al., 'International Institutions for Advanced AI'.

III. Why the United Nations Should Exercise Caution Against Creating a Global AI Agency without a Specific Mandate

The United Nations should exercise caution against creating an overly broad global AI agency without a specific mandate because of the institutional challenges that designing and operating such an institution would pose. Among the four suggested models, variations of the second model — a global AI agency or "advanced AI governance organisation" — appear to be the most common proposal by international leaders (Table 1). The supporters of such a model argue that advanced AI systems will create existential and other forms of risks as they grow in prominence in the international society. Such concerns were similarly observed with the development and spread of nuclear technologies. Thus, as was the case with atomic energy, the creation of a similar global AI agency is needed to deal with AI safety, according to its supporters.[11]

Given certain similarities between AI and nuclear energy, it is all too understandable why world leaders draw a comparison between the two. As the case with electronics, automobiles, and modern medicine, AI and nuclear energy are general-purpose technologies in that they can promote growth and innovation in the entire economy. Likewise, nuclear and AI systems are also dual-purpose technologies. While AI and nuclear energy can be used to promote economic growth, they can also be used in developing offensive capabilities, with autonomous weapons and nuclear weapons being the two obvious examples.

Notwithstanding these general similarities, the differences between AI and nuclear technologies means that AI governance will require a fundamentally different approach. A major difference between nuclear and AI technologies is that the former is a physical technology, while the latter is primarily digital.[12] This difference means that the principal components for and means of developing and improving nuclear energy and AI capabilities are distinct. Because nuclear energy is a primarily physical entity, physical infrastructure and materials–such as nuclear reactors and fissile materials–are necessary to produce nuclear weapons.[13] Producing nuclear energy requires not only the knowledge of nuclear science and engineering itself, but also occupational privileges or high-level clearance from member states' governments. Meanwhile, although advanced AI systems do require advanced physical computing infrastructure, the constraining input factors are primarily digital: the availability of high-quality data sets, training models, and algorithms — along with AI expertise. As a result, the barriers to entry in AI are much lower than the case for nuclear technologies.[14]

That is one reason why, whereas state actors dominate the nuclear sector, private actors play a much more important role in the AI landscape than government entities. In the United States, while the defence establishment has played a critical role in developing certain advanced AI capabilities, recent AI innovation has been largely driven by the private sector and research institutions. Due to the

---

[11] I. Stewart, 'Why the IAEA Model May Not Be Best for Regulating Artificial Intelligence', *Bulletin of the Atomic Scientists* (9 June 2023), https://thebulletin.org/2023/06/why-the-iaea-model-may-not-be-best-for-regulating-artificial-intelligence/.

[12] The author thanks Ms Lindsey Carpenter, Attorney at the National Taxpayers Union Foundation, for her insights on the topic.

[13] I. Stewart, 'IAEA model for AI'.

[14] IY. Afina & P. Lewis, 'The Nuclear Governance Model Won't Work for AI', Royal Institute of International Affairs (28 June 2023), https://www.chathamhouse.org/2023/06/nuclear-governance-model-wont-work-ai. Stewart, 'IAEA model for AI'.

multiplicity of non-state actors, AI governance is fundamentally more multifaceted than nuclear governance and requires a different approach.[15] Consequently, an IAEA-like organisation designed to monitor compliance of foreign governments and nuclear plants with the relevant regulations is much less relevant in the context of AI governance.

Finally, and most importantly in the context of international governance, the risks associated with nuclear technologies can be defined more concretely than AI safety risks. Although the Chinese, Russian, and US governments might differ in their conceptions of international law and global governance, they are nevertheless likely to agree that nuclear proliferation is generally harmful to international security. In contrast, long-run existential risks from advanced AI remains a subject of intense scholarly and policy debates. The resulting lack of scientific consensus means that any international organisation will have a significantly difficult time agreeing on which AI risks to prioritise and developing mechanisms to address those risks.[16]

This potential for disagreement between member states poses a serious challenge to any future AI governance initiatives, albeit potentially less so if such efforts target specific, well-defined areas such as lethal autonomous weapons systems (LAWS). Even in the context of well-defined risks, AI governance initiatives can struggle to find a meaningful consensus among member states, as evidenced by the recent difficulties in the UN efforts to regulate LAWS.[17]

Likewise, short and medium-term AI risks vary widely depending on the precise context in which AI applications are used. For example, one can consider three areas in which AI application poses important but different types of risks. First, as mentioned, a major risk that AI poses in the context of international security is the development of AI-enabled LAWS, which the UN has sought to regulate in the past.[18] Second, in the context of fundamental rights, facial recognition and surveillance technologies already pose significant challenges to civil liberties, especially in countries with poor human rights records.[19] Third, since AI capabilities tend to be concentrated in a handful of countries, the question of how developing and emerging-market countries can develop AI capabilities and leverage artificial intelligence to promote economic growth is becoming an increasingly crucial one for global governance.[20]

These examples are distinct issues in the fields of arms control, human rights, and international development, respectively, which require different policy approaches and institutional frameworks

---

[15] R. Nabil, 'Comments to the White House: The Need for A Flexible and Innovative AI Framework', National Taxpayers Union Foundation (7 July 2023), https://www.ntu.org/foundation/detail/letter-to-the-white-house-the-need-for-a-flexible-and-innovative-ai-framework. R. Iyengar, 'Who's Winning the AI Race? It's Not That Simple', *Foreign Policy* (27 Mar. 2023), https://foreignpolicy.com/2023/03/27/us-china-ai-competition-cooperation/.

[16] Е. Постникова [E. Postnikova], 'Совбез ООН призвали не допустить создание «машин-убийц» на основе ИИ' ['The UN Security Council urged to prevent the creation of "killer machines" based on AI'], *РБК* (18 July 2023), https://www.rbc.ru/politics/18/07/2023/64b6c01f9a79477796c5b78c.

[17] S. Reeves, R. Alcala, & A. McCarthy, 'Challenges in Regulating Lethal Autonomous Weapons Under International Law', *Southwestern Journal of International Law* 28 (2020), 101–118, https://ssrn.com/abstract=3775864.

[18] J. Lewis, 'The Case for Regulating Fully Autonomous Weapons', *Yale Law Journal* 124/4 (2015), 9, https://www.yalelawjournal.org/comment/the-case-for-regulating-fully-autonomous-weapons.

[19] Cf. Lewis, The Case for Regulating Fully Autonomous Weapons'.

[20] See 'Chapter 4: The Economy and Education' in Stanford Institute for Human-Centered Intelligence (HAI), *Artificial Intelligence Index Report* (Stanford, 2022), 172–195, https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf.

specific to those domains. The variety of contexts in which AI poses distinct challenges means that a single global agency will be much less appropriate for AI governance than for nuclear regulation and other areas of UN competence. Instead, the United Nations's efforts might be better spent by identifying domain-specific AI policy challenges, analysing the strengths and weaknesses of the UN's institutional framework in those domains, and developing AI-related initiatives within the context of existing UN institutions.

To that end, the United Nations could create working groups with AI experts and interested parties that help identify and evaluate areas of governance with an AI nexus where the UN can play an active role. Once such areas are identified, the UN should then identify and evaluate potential AI-related risks and set its policy objectives within the context of those domains. The institutional models described earlier could help the UN evaluate which frameworks are best suited for the UN's policy objectives in these policy areas. This approach would also help the UN leadership evaluate the extent to which collaboration with other multilateral organisations will be helpful, or whether certain complex issues might eventually require the creation of a new UN institution.

IV. Areas Where the United Nations Could Play an Effective Role in Global AI Governance

Given the UN's unique role as a platform that brings together developed and developing countries and countries of varying political persuasions, it could provide an important venue for discussing international AI norms and principles. However, normative disagreements between major powers (e.g., China, Russia, and the US) and even like-minded jurisdictions (e.g., the US and EU) on AI governance will nonetheless pose a major challenge – as became evident in the recent UN Security Council debate on AI.[21]

As a result, the UN's efforts are more likely to be successful if it focuses on developing *voluntary* AI principles, frameworks, and agreements — similar to the Organisation of Economic Development and Cooperation's (OECD) approach to AI.[22] These voluntary AI governance initiatives could be especially helpful for some developing economies that might lack the resources and expertise to develop national AI policies and look to multilateral institutions for guidance on such issues.

The UN could also play an active role in helping nations identify and mitigate current and future AI safety risks. To that end, the United Nations could set up multidisciplinary working groups that evaluate AI risks in different domains and recommend possible mitigation strategies in consultation with national governments, the private sector, academic institutions, and the civil society. The UK government has proposed a similar risk assessment framework as part of its recent AI White Paper — a framework that we have recommended for the US government in our recently submitted comments to the White House.[23]

---

[21] E. Postnikova, 'The UN Security Council'.

[22] La Cour des comptes [The Court of Auditors], *Comparaison de 10 stratégies nationales sur l'intelligence artificielle* [Comparison of 10 National AI Strategies] (2023), https://www.ccomptes.fr/system/files/2023-04/20230403-comparaison-strategies-nationales-strategie-nationale-recherche-intelligence-artificielle.pdf.

[23] UK Department for Science, Technology, and Innovation & the Office for Artificial Intelligence, *A Pro-Innovation Approach to AI Regulation* (White Paper, CP 815, 2023), https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper. Nabil, 'US AI Governance'.

While these national-level mechanisms are designed to help individual countries mitigate potential AI risks, the UN can play an important role in evaluating potential risks at the global level. Such evidence-based, impartial analysis could also inform the national strategies of member states, particularly from the developing world, as they design national and regional strategies to mitigate AI safety risks. Likewise, the UN could draw upon its deep expertise in technical standards and enable the International Telecommunications Union (ITU) and its working groups to play a more active role in AI standardisation.

V. Conclusion

To paraphrase the German Chancellor Olaf Scholz, the international order does appear to be in the middle of a *Zeitenwende*, especially as emerging technologies and rising powers create new challenges for global governance.[24] Against this backdrop, the United Nations can play a constructive role by providing a platform for intergovernmental dialogue, helping national governments craft better AI policies, and promoting international cooperation in developing AI norms and standards. To that end, the UN needs a flexible, pragmatic approach that considers the context-specific nature of AI and the distinct challenges it poses in different areas of global governance.

In the long term, it is possible that solving some future AI safety risks might require the creation of new institution(s) within the current UN architecture. However, at a time when the precise long-term risks are unclear and subject to debate, the UN should instead focus on clearer, pressing challenges where the UN can play a constructive role. To do so, the UN should first identify domain-specific AI risks and policy objectives and whether the UN's institutional design and comparative advantages are well-suited for policy efforts in those domains. Likewise, UN agencies with an AI nexus should also develop the required expertise to help address AI-related challenges within their remit. Instead of prematurely creating a global AI agency without a specific mandate, a more flexible, well-calibrated, and iterative approach would allow the UN to play a more effective role in the emerging international institutional architecture for AI governance.

---

[24] O. Scholz, 'Die globale Zeitenwende: Wie ein neuer Kalter Krieg in einer multipolaren Ära vermieden werden kann' ['The Global Zeitenwende: How to Avoid a New Cold War in a Multipolar Era'], *Foreign Affair*s (5 Dec. 2022), https://www.foreignaffairs.com/germany/die-globale-zeitenwende.