UN REINVENTION RESEARCH GROUP

Toda Peace Institute

Dag Hammarskjöld Foundation



Governing What Sovereignty Cannot

The case of artificial intelligence (AI) governance

Renata Dwan

Reinventing the United Nations – Paper 5 · Policy Brief No. 335 – October 2026

A joint initiative of the Toda Peace Institute and the Dag Hammarskjöld Foundation

About the Author



RENATA DWAN

Renata Dwan has more than two decades' experience of multilateral policymaking. As Special Advisor to the UN Special Envoy on Digital and Emerging Technologies, she led support to the negotiation of the Global Digital Compact. She directed the United Nations Institute for Disarmament Research (UNIDIR), initiating new programmes on conflict and technology, and led major UN-wide reform initiatives as Chief of UN Peacekeeping Policy and Best Practices. She served in Afghanistan, DRC, Haiti, Mali and Syria and advised the EU on crisis management. Outside international organisations, Renata was Deputy Director of Chatham House, overseeing the institute's research agenda and programmes, and Director of the Armed Conflict programme at Stockholm International Peace Research Institute (SIPRI). She has a D.Phil, M.Phil and B.A. from the University of Oxford.

To cite this policy brief:
Dwan, Renata. 2026. Governing What Sovereignty Cannot: The Case of Artificial Intelligence (AI) Governance. Policy Brief No. 335, Toda Peace Institute. https://toda.org/publications/policy-briefs-and-reports/governing-what-sovereignty-cannot-the-case-of-artificial-intelligence-ai-governance/

Author's views: The views expressed are those of the author and do not necessarily reflect those of the Dag Hammarskjöld Foundation or Toda Peace Institute.

Declaration of AI use: The author did not use generative AI tools in preparing this paper. The author has reviewed the final text and takes full responsibility for its accuracy, sources and references, arguments, analysis and conclusions.

Toda Peace Institute
Samon Eleven Bldg. 5F,
3-1 Samon-cho, Shinjuku-ku
Tokyo 160-0017, Japan

Tel. +81-3-3356-5481
Fax. +81-3-3356-5482
Email: contact@toda.org

© 2026 Toda Peace Institute

Abstract

UN reinvention depends on delivering collective action where sovereignty cannot. This paper explores how the UN became a principal body for navigating global issues that go beyond the power, authority, and capabilities of any one state. AI governance is the foremost such challenge today and one that offers the UN scope for renewed relevance and impact. This paper examines how the Organisation might rise to this challenge. It highlights continuity in the way the UN has approached AI governance to date as a global issue and recommends four actions for the first 100 days of the next Secretary-General.

‘Harmonising the actions of nations’: the UN and global challenges

There is a range of challenges that do not fit neatly within any one pillar of the United Nations but whose impacts affect the peace and security, development, and rights of all. Climate change, the health of the oceans, the spread and eradication of disease, or the stable functioning of international financial systems are subjects of sovereign concern but not of sovereign disposal, even for the most powerful or capable of governments.

The UN Charter recognises the potential for such challenges and frames them as “international problems of an economic, social, cultural or humanitarian character” alongside its provisions for threats, acts of aggression and other breaches of the peace. International cooperation is required to solve them and to that end the United Nations is “to be a centre for harmonizing the actions of nations.”[1]

In the eight decades since, the Organisation and its specialised agencies have been a principal forum for tackling a broad range of international problems. In so doing, the UN system has shaped the norms, frameworks, and mechanisms for the governance of issues that go beyond the power and capabilities of any one state.

Governing the natural world

An early and formative case was the 1967 proposal, by Malta, that the seabed and ocean floor beyond national jurisdiction be declared the common heritage of humanity. Later codified in the 1982 United Nations Convention on the Law of the Sea (UNCLOS), the principle established the normative and legal foundations for the governance of designated natural domains beyond national jurisdictions: shared ownership and responsibility for their stewardship.

Peace and security concerns drove these early efforts to tackle ‘international problems’. Cold War competition between the Soviet Union and the United States raised the prospect of occupation or annexation of shared domains. The 1959 Antarctic Treaty, negotiated in Washington at the invitation of the US, froze territorial claims, reserved Antarctica for peaceful purposes and prohibited nuclear explosions and radioactive-waste

[1] *Charter of the United Nations*, Article 1.3 and 1.4.

disposal. International alarm over US–Soviet arms racing drove negotiation of the 1967 Outer Space Treaty that banned weapons of mass destruction in orbit and the appropriation of "celestial bodies."[2] Governing these shared domains was as much about mitigating the risks and impacts of great power competition as organising shared stewardship of the natural world. Where ownership remained sovereign, as in the case of rainforests, the principle of shared concern and legitimate interest in their management was enshrined in treaties such as the Convention on Biological Diversity.[3]

The positive benefits of global governance were also highlighted. Scientific exploration as the right of all states, the promotion of scientific collaboration, and shared findings were prominent in early initiatives, premised on their potential to support social and economic development.[4] Scientific evidence and cooperation remain a central tenet of the UN's approach to governing global goods. As the Intergovernmental Panel on Climate Change (IPCC) reflects, such collaboration is seen as both a means to collective action and an outcome of cooperation.

Shared revenue and benefit-sharing arrangements as an incentive for cooperation emerged later. UNCLOS established a central institution, the International Seabed Authority, mandated, among other functions, to share the financial benefits of seabed mining equitably. Other environmental governance mechanisms include transfers and subsidies to support compliance. The Montreal Protocol Multilateral Fund, for example, is financed by developed-country parties on the UN scale of assessments to meet the agreed incremental costs incurred by developing countries in phasing out ozone-depleting chemicals. Behind each of these initiatives was a coalition of engaged governments, diplomats, scientists, and UN offices driving the negotiation of complex legal frameworks, often over many years.

Governing the human world

The end of the Cold War created new opportunities—and challenges—for cooperation. Leaders at the first Summit of the Security Council in 1992 declared that "non-military sources of instability in the economic, social, humanitarian and ecological fields have become threats to peace and security. The United Nations membership as a whole, working through the appropriate bodies, needs to give the highest priority to the solution of these matters."[5]

Borrowing from economic theory, the UN Development Programme introduced the concept of Global Public Goods (GPG) to describe non-excludable and non-rival systems and resources that benefit multiple groups and do not discriminate against any population group or generation such as infectious disease eradication, climate adaptation, or stable global financial systems.[6] Shared concern was assumed: the focus was on how such outcomes could be collectively delivered.

---

[2] Full text at Antarctic Treaty - Wikisource; Outer Space Treaty at United Nations treaties and principles on outer space.

[3] At the 1992 Earth Summit negotiations, developing states rejected proposals for the internationalisation of rainforests and climate systems. The compromise was to declare them subjects of 'common concern', rather than 'common heritage', of humankind.

[4] The Outer Space Treaty declared astronauts to be 'envoys of mankind' who must be rendered all possible assistance. On the codification of scientific collaboration for environmental protection see the Rio Declaration on Environment and Development (1992), Vienna Convention for the Protection of the Ozone Layer (1985), UNCLOS (1982) and the Convention on Biological Diversity (1992).

[5] Note by the President of the Security Council, 31 January 1992, S/23500.

[6] The distinguishing characteristics of global public goods were identified as publicness and quasi-universal benefits. Inge Kaul, Isabelle Grunberg & Mark Stern (eds), *Global Public Goods: International Cooperation in the 21st Century* (New York: UNDP, Oxford: Oxford University Press, 1999).

GPG thinking expanded multilateral agendas. Attention and resources were directed at policy outcomes such as conservation or pandemic preparedness, with donors and institutions earmarking funds and measuring progress against delivery. Framing challenges such as organised crime or migration as collective action problems enabled international consideration of domestically sensitive issues through the negotiation of norms and principles, standards, and practical incentives.[7] The role of non-state actors was highlighted, especially in technology governance, where companies and experts were active in UN-convened World Summits on the Information Society (WSIS) and the Internet Governance Forum.[8] The UN's universality made it a central convenor while its practical role in facilitating cooperation through information, technical expertise and capacity development drove the growth, especially, of specialised UN agencies.

For all the emphasis on technical solutions, political consensus underpinned increased collective action. The breakdown of that consensus over the past 15 years and growing scepticism of the costs and benefits of cooperation have upended progress. The principle of shared concern, much less shared ownership, is today challenged. The US, China, Russia, and private companies are probing exploitation of space and the high seas for strategic and commercial gain. Burden-sharing and benefit distribution arrangements on issues such as carbon emissions and pandemic preparedness are contested by wealthy as well as developing states and companies.[9] Short-term interest-driven risk-taking is legitimated in nationalist and populist politics worldwide. As US withdrawal from the Paris Agreement demonstrates, exit from cooperation comes with few costs for a state but significant negative impact on the regime.

Collective action failures are intensifying criticism of the machinery of cooperation, from international law to multilateral institutions. The UN has been slow to adapt, relying on moral appeals and technical approaches to fundamentally political challenges. Yet no coherent alternative has emerged. Neither the US nor China is willing, or potentially able, to bear the disproportionate costs of supplying global goods. Regional cooperation, as Covid-19 demonstrated, has limited impact in filling the vacuum. Ad hoc coalitions struggle to sustain cooperation beyond specific issues. Meanwhile the scale, frequency and impact of global shocks and disruptions mount.

Governing the artificial world

Artificial intelligence (AI) is a particularly novel global disruption.[10] As a general-purpose technology with wide applicability and huge potential, AI is disseminating at unprecedented speed and scale. Its development relies on components that are globally supplied but privately owned. The more powerful an AI system, the more intensive those infrastructure and capital requirements, concentrating advanced (or frontier) AI capabilities in a handful of US and Chinese companies and compounding US–Chinese rivalry.

Concern at the destabilising potential of AI competition, misuse, or loss of control is mounting. A series of documented jailbreaks and unauthorised cyber hacks by frontier AI models in mid-2026 has led some tech leaders to call for greater oversight of the most advanced models. Both US and Chinese governments have resisted calls for slowing down advanced AI development, citing their competition, even as they explore bilateral dialogue on AI risks and, in China's case, introduce restrictions at home. The temporary withdrawal,

[7] See, for example, Jack Hirshleifer, "From Weakest Link to Best-shot: The Voluntary Provision of Public Goods", *Public Choice*, Vol 41 (1983), 371-386, and Ewa Latoszek and Agnieszka Klos (eds.), *Global Public Goods and Sustainable Development in the Practice of International Organizations* (Brill 2023).

[8] See WSIS Outcome documents - December 2005 and A/RES/60/252 (2006), establishing the Internet Governance Forum (IGF). A/RES/80/173 (2025) formalised IGF as a permanent UN forum.

[9] See ongoing negotiations around the Pathogen Access and Benefit Sharing (PABS) system annex to the WHO Pandemic Agreement adopted in May 2025.

[10] The widely used OECD definition of AI is a machine-based system that, for explicit or implicit objectives, infers from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.

by the US, of foreign nationals' access to two advanced models and the refusal by one of the leading AI labs to share its latest model with UK and other AI safety institutes for pre-release testing have highlighted global vulnerabilities to AI risks as well as protectionism.[11] It is leading some governments to pursue AI sovereignty as a national security imperative despite concerns over feasibility, cost, and control.

For many of the 80+ governments that have produced national AI strategies, AI's potential for service delivery and productivity makes adoption and diffusion a development priority notwithstanding AI's sensitivity to local context for effective application.[12] In developed countries, the focus is on how AI could upend work, especially white-collar jobs. Environmental concerns around data centres and energy dominate local politics. Papal and public attention is seized by the ethical and social impact of AI companions and AI-enabled disinformation for harm and trust.[13] No single technology is raising so many governance challenges across so many domains in such a short timeframe, even as scientists and economists debate its long-term trajectory and revenue potential.

The UN has been a comparatively early mover to explore AI opportunities and risks. The International Telecommunication Union (ITU) convened its first AI for Good Global Summit in 2017 to champion applications for development. The Human Rights Council adopted its first biennial resolution on new and emerging digital technologies and human rights in 2019 and has declared the right to digital privacy as well as the need for AI regulation to adhere to human rights law. The first global normative statement on AI was UNESCO's 2021 Recommendation on the Ethics of Artificial Intelligence.[14]

The UN's decentralised structure, in this instance, facilitated action. These initiatives were largely autonomous and siloed. They took place in parallel to discussions in the UN disarmament machinery, maintaining, at least formally, a distinction between military and civilian use. They were more agile than the General Assembly's ten-yearly reviews of digital transformation and Internet governance. Limited Security Council engagement on emerging technologies, reinforced by divisions among its five permanent members, enabled action outside it. [15]

Yet caution in moving beyond normative statements was evident and early initiatives had limited visibility or traction.[16] It took the leadership of the Secretary-General to put the issue of AI governance at the centre of the UN political agenda. Guterres used a range of methods to make the case for digital and AI cooperation between 2018 and 2024, including establishing two high-level panels to set out rationales and priorities for multistakeholder action, appointing an Envoy on Technology, setting out a roadmap for improved UN coordination, and proposing the negotiation of a Global Digital Compact (GDC). UN activity prompted the US, and subsequently China, to sponsor resolutions in the General Assembly calling for international cooperation

---

[11] Eleanor Olcott, Joe Leahy, Zijing Wu & Demetri Sevastopulo, "Why the US and China are at odds on AI", *Financial Times*, 15 September 2026; Mollie Feng, "What to Know About China's First AI Companion Rules", *Just Security*, 6 August 2026, Joao Marinotti, "Legally or not, the US government is controlling global access to the world's most powerful AI", *The Conversation*, 21 July 2026.

[12] See World Development Report 2026, *The Promise of Artificial Intelligence* (WBG: Washington, 2026) which declares that "Developing countries should leave it to the rich world to agonize over whether AI will one day doom humanity", xix.

[13] Tony Romm, "With AI riches at stake, pressures mount to share the wealth", *The New York Times*, 16 August 2026; Leo XIV, Encyclical, *Magnifica Humanitas*, 15 May 2026, "Encyclical Letter of His Holiness Leo XIV Magnifica Humanitas; Should AI labs be treated like the owners of dangerous animals?", *The Economist*, 6 August 2026; Shiona McCallum, "Safety of AI chatbots for children and teens faces US inquiry", *BBC News*, 12 September 2025.

[14] ITU AI for Good Global Summit 2017, A/HRC/Res/41/11, 11 July 2019, A/HRC/RES/48/4, 7 October 2021; UNESCO, Recommendation on the Ethics of Artificial Intelligence November 2021.

[15] A/RES/80/57, 1 December 2025 is the most recent mandate to the Convention on Certain Conventional Weapons to continue to address autonomous weapons systems; the most recent WSIS review is A/RES/80/173, 17 December 2025; the Security Council has held four ministerial-level meetings on AI convened by UK (18 July 2023), US (19 December 2024), Republic of Korea (24 September 2025) and French presidencies (23 September 2026).

[16] ITU Plenipotentiary Conference in 2018 refused to include AI regulatory guidance or governance in ITU's remit, agreeing to an explicit AI mandate only in 2022 ITU Resolution 214 (Bucharest 2022). Bucharest also saw the election of a US national to ITU leadership, succeeding a Chinese national.

on safe, secure, and trustworthy AI (US) and capacity building (China) in 2024.[17] The adoption of the GDC at the Summit of the Future, which committed states to set up an independent scientific panel on AI and an annual global dialogue on AI governance, was a reflection of the impact that sustained leadership by a Secretary-General can have on issues of collective action outside the agenda of the Security Council.[18]

The establishment of AI governance as an issue of shared concern built on past global governance approaches. Guterres called for action to "govern AI for humanity." The Independent International Scientific Panel on AI posits scientific understanding as a foundation for collective action. The UN's convening function is reflected in the Global Dialogue on AI Governance which frames cooperation technically in terms of interoperability. The outgoing Secretary-General proposed a range of measures to tackle AI access and benefits—an AI capacity development network to help developing countries access compute and infrastructure, a global fund for capacity building and a global data framework to pool data for machine learning—none of which have generated, to date, significant support.[19]

Perhaps the newest feature of the UN's approach is the lack of emphasis on legally binding treaties, hitherto the primary mechanism for global action. This is partly a reflection of the nature and pace of AI development which challenge regulatory instruments built on quantitative, verifiable commitments. It mirrors domestic struggles among governments, industry and civil society actors to balance innovation with regulation, private with public interest. Many bureaucracies lack technical AI knowledge to design effective governance frameworks. Industry transparency and voluntary disclosure initiatives do not fill the gap. The range of components that make up AI systems presents dilemmas for competition policy, while liability faces novel questions of intent. The evidence and the tools for state management of AI are still being built even as the technology is being rolled out in every aspect of human life.[20]

Geopolitical rivalries and US–China AI competition also factor. Some western states are concerned that the UN's global membership gives authoritarian states outsize influence in shaping AI governance and limits scope for multistakeholder approaches. They query the UN's technical capacity to shepherd AI policy development. China's support for AI governance at the UN and the Trump administration's opposition reinforce western caution.[21] Many developing states, for their part, push back against Guterres' concern with AI risks as a potential obstacle to accessing and adopting AI. Some are sceptical of the organisation's ability to deliver capacity development at scale. Divisive UN bureaucratic competition and budget constraints do little to encourage confidence. The planned General Assembly review of new AI bodies and processes, as part of a broader assessment of GDC implementation in 2028, is indicative of the fragility of consensus on global AI governance and the UN's role in it.

---

[17] A/RES/78/265, 21 March 2024 and A/RES/78/311, 1 July 2024.

[18] *Pact for the Future: Global Digital Compact*, (Annex I), A/RES/79/1, 22 September 2024. Terms of reference for the Independent International Scientific Panel on AI and the Global Dialogue on AI governance, A/RES/79/325, 26 August 2025.

[19] A/79/966, 2 July 2025, and A/80/817, 24 August 2026.

[20] Binding regulations on AI include auditing and reporting requirements for advanced AI models, as in the EU AI Act (2024) and California Senate Bill 53, *Transparency in Frontier AI Act* (2025). New York's RAISE Act takes effect on 1 January 2027.

[21] State Council Information Office of the People's Republic of China, "More Just and Equitable Global Governance: China's Principles, Proposals and Actions", June 2026; Keynote speech by President Xi Jinping, 2026 World AI Conference, July 2026; Remarks by Vice-President Vance at the AI Action Summit, Paris, France, 11 February 2025.

Recommendations

A new Secretary-General will inherit the principle but little practice of global AI governance. She will face a turbulent AI environment and a short timeline to deliver action. Neither moral exhortation nor scientific evidence will suffice. The new UN leadership must build a political constituency, engage new partners, and harness public concern to drive AI governance action at industry, national, and international levels. The Secretary-General must make the UN the facilitator, not the centrepiece, of global action. This paper recommends four actions in the first 100 days of the new Secretary-General's tenure to chart the course for the rest of her term.

1. Mobilise political support for AI governance action

Few champions or coalitions have yet emerged for AI governance action at the UN.[22] A new Secretary-General could use her initial political capital to encourage diverse member states to advance collective action potentially through a Group of Friends on AI Governance. An informal cross-regional grouping could give the Secretary-General a platform to shape priorities and test ideas while building trust between the Secretariat and member states. A grouping of AI-engaged and AI-capable states could give the Secretary-General the backing she needs to navigate geopolitical divides and to connect with emerging governance initiatives, such as the recent international call for control of frontier AI models led by Norway and Finland.[23]

Such a Group could leverage the concern of many states for shared AI standards and risk management. It could harness the interest of middle-income states, such as Brazil, India or Kenya, to access AI infrastructure and capacity without being forced to choose between US or Chinese-only systems as their respective international initiatives—Pax Silica and the World AI Cooperation Organisation (WAICO)—appear to reflect.[24] By focusing on areas of shared interest, such a group could champion practical initiatives for evaluation and safety standards. Advocacy for transparent reporting by AI developers on training, testing, and auditing of open-weight models, or public procurement initiatives for AI services could provide tangible targets to mobilise action by diverse states. Collaboration on child safety or multilingual data could send a visible demand signal to industry and provide practical help to states in building locally tailored AI tools. Mobilisation around specific issues could offer an entry point to addressing safety and diffusion challenges while avoiding divisive debate on new international institutions. A new Secretary-General could make clear to China and the US her intention to advance safe AI diffusion through practical cooperation and not to propose the creation of a new UN regulatory body, unless explicitly requested by member states. She could publicly advocate—and privately offer her good offices—for US–China bilateral dialogue to mitigate the misuse, misapplication, or loss of control of the most advanced AI systems and highlight how multilateral action could complement a bilateral track.

[22] Two New York groupings were established in 2024: the Group of Friends of AI for Sustainable Development, convened by Morocco and the US, and the Group of Friends for International Cooperation on AI Capacity-Building, chaired by China and Zambia. The proposed cross-regional grouping would seek political backing for practical initiatives across AI governance beyond their respective development and capacity-building focuses.

[23] A Call for Control of Frontier AI Models, 21 September 2026. Leaders representing 29 states, together with the President of the European Commission, had endorsed the call as of 30 September 2026.

[24] The US launched Pax Silica in December 2025 to secure supply chains for advanced technologies and infrastructure development. The US Department of State reported 25 declaration signatories when Italy joined on 31 July 2026. China's World Artificial Intelligence Cooperation Organisation, established in July 2026 and focused on AI cooperation and capacity building, had 37 member states according to the China Centre for International Science and Technology Exchange on 5 August 2026. The US has argued that participation in both initiatives is incompatible; September reporting identified Kazakhstan as the only country participating in both. See "US Department of State announcement; China Centre for International Science and Technology Exchange, 5 August 2026," *Reuters*, 14 August 2026 and *The Diplomat*, 8 September 2026.

2. Establish a standing mechanism for private sector engagement

If collective action on AI access and benefits is to be pursued, industry must be involved. Most AI knowledge, components, and resources are in private hands: governments cannot negotiate their distribution (as the experience of the WHO Pandemic Agreement has illustrated). How to engage private actors while avoiding undue influence is not a new dilemma. But it is one that the UN secretariat has tended to avoid. A new Secretary-General must tackle it.

One option she could consider is to build on the 2000 Global Compact established by Kofi Annan to encourage companies to sign up to principles on human rights, labour, and the environment.[25] The Compact could be reformed as a framework for engaging the private sector in solving global problems or, more realistically, the new Secretary-General could establish a dedicated roundtable on AI governance within it. The functions of such a mechanism could include sharing insights on AI developments; providing input to AI access initiatives, and mobilising resources for capacity development, including in-kind AI compute and infrastructure support.

Participation could be open to private companies that have formally endorsed the GDC. Conflict of interest guidance could be drawn from other UN frameworks, for example, the WHO Framework of Engagement with Non-State Actors (2016). The mechanism would be independent of the activities of UN technical agencies, such as ITU, which regularly engage private companies in standard-setting.

3. Secure financing for the Independent International Scientific Panel on AI

The new 40-member panel presented its first, preliminary, report in 2026: a full report will be submitted to the Global Dialogue on AI Governance in May 2027.[26] The panel's composition and initial output have been broadly welcomed as balanced and independent, although states and stakeholders continue to closely monitor it for autonomy and neutrality.

If the panel is to provide an authoritative evidence basis for dialogue and action, it will be essential that it has secure and predictable financing to maintain an independent secretariat, the resources to engage with scientists and research institutions around the world, and the capability to respond at pace to major AI developments. Speed and response demands distinguish the AI panel from the IPCC and must include the potential to produce interim or issue-specific briefs, including on major AI incidents, in addition to its annual report to the Global Dialogue.[27]

Currently, the panel is funded through one-off contributions by a handful of countries and tech philanthropies to an existing UN fund that supports activities of the Office of Digital and Emerging Technologies (ODET). The Secretary-General could establish a dedicated voluntary trust fund open to government, intergovernmental and, potentially, closely screened philanthropic funding. That would enhance the panel's transparency and independence while also ensuring it has the funding required to work effectively. Contributions should be pooled, preferably multi-year and unearmarked for substantive work, with published donor information and safeguards.

---

[25] The only commitment is annual communication on progress by participants against ten principles on human rights, labour, the environment and anti-corruption, see Homepage | UN Global Compact.

[26] *Preliminary report of the Independent International Scientific Panel on AI*, 1 July 2026.

[27] The Panel's first thematic brief, *AI Agents, Misalignment and the Risk of Losing Human Control: Evidence from the OpenAI-Hugging Face Incident*, was released on 21 September 2026.

4. Advocate for the Global Dialogue on AI Governance to serve as a platform for actionable governance initiatives

The first Global Dialogue on AI Governance, described by its Co-Chairs as a "distinctive convening platform", took place in July 2026.[28] Some 3,000 people and 163 states participated in a range of plenary and thematic discussions over two days in Geneva. The 2027 Dialogue will take place in May in New York. A new Secretary-General could make it a priority to prevent the convening becoming a UN-centric talk shop. She could engage with the President of the General Assembly and the Co-Chairs of the Dialogue whom he appoints to share her vision of a dynamic platform for exchange that helps governments and stakeholders advance concrete action in specific areas of AI governance. She could stress, in public statements, that this is not about bringing AI governance action inside the UN but about facilitating national, regional and multistakeholder cooperation beyond it.

To make the Dialogue a platform for action, she could propose that its Co-Chairs establish an informal multistakeholder advisory group of representatives from regional organisations, civil society, private sector and technical communities to identify a subset of thematic priorities. She could encourage them to consider inviting the convenors of minilateral and regional AI initiatives to contribute so as to facilitate cross-exchange and promote successful cases of collective action. This could also provide a way of addressing sensitive issues. For example, the Organisation for Economic Co-operation and Development (OECD) could be invited to brief on its (open-source) AI incident monitor. The UK-led International Network for Advanced AI Measurement Evaluation and Science could update on safeguards developments among national AI safety and security institutes.

She could direct the UN offices involved in supporting Dialogue planning—ITU, ODET and UNESCO—to focus less on showcasing UN activities and more on facilitating exchange among existing AI governance initiatives, including mapping international governance commitments.[29] She could task them to explore scope for intersessional work to prepare concrete commitments and actions, such as the launch of the AI Child Safety Pledge at this year's Dialogue. Bringing global actors together is insufficient. The Dialogue must demonstrate how the effective curation of cross-cutting interests and actors can drive AI governance action and, in so doing, show the UN's value in tackling problems that no state can solve.

AI has captured the global imagination: the Secretary-General must voice the hopes and the fears that it generates. Uncertainty about its trajectory is not sufficient reason to defer its governance. If the UN can help put in place frameworks and tools for safe and equitable development, AI governance may provide a foundation for the Organisation's future.

---

[28] *Co-Chairs' Summary: UN Global Dialogue on Artificial Intelligence Governance*, Geneva, 6-7 July 2026.

[29] The AI Dialogue Partnerships Hub, launched in May 2026 to showcase practical AI initiatives, currently registers 359 projects against the 17 Sustainable Development Goals, mostly focused on AI capacity development.

DAG HAMMARSKJÖLD FOUNDATION

The Dag Hammarskjöld Foundation is an independent, non-governmental organisation established in 1962 in memory of the second Secretary-General of the United Nations. Building on Dag Hammarskjöld's legacy, we advance dialogue, principled policy, and effective practice to support peace and sustainable development.

THE TODA PEACE INSTITUTE

The Toda Peace Institute is an independent, nonpartisan institute committed to advancing a more just and peaceful world through policy-oriented peace research and practice. The Institute commissions evidence-based research, convenes multi-track and multi-disciplinary problem-solving workshops and seminars, and promotes dialogue across ethnic, cultural, religious and political divides. It catalyses practical, policy-oriented conversations between theoretical experts, practitioners, policymakers and civil society leaders in order to discern innovative and creative solutions to the major problems confronting the world in the twenty-first century (see www.toda.org for more information).

CONTACT US

Toda Peace Institute
Samon Eleven Bldg. 5th Floor
3-1 Samon-cho, Shinjuku-ku, Tokyo 160-0017, Japan

Email
contact@toda.org

Sign up for the Toda Peace Institute mailing list
https://toda.org/policy-briefs-and-resources/email-newsletter.html

Connect with us on the following media.
YouTube @todapeaceinstitute3917
X https://x.com/TodaInstitute
Bluesky @todapeaceinstitute.bsky.social
Facebook https://www.facebook.com/TodaInstitute