GENEVA GRADUATE INSTITUTE

INSTITUT DE HAUTES ÉTUDES INTERNATIONALES ET DU DÉVELOPPEMENT
GRADUATE INSTITUTE OF INTERNATIONAL AND DEVELOPMENT STUDIES

Geneva Graduate Institute
Masters in International and Development Studies
ARP 35

AI Governance Beyond 2025: UN Pathways and Implications

Team members:

Bhavya Goel (bhavya.goel@graduateinstitute.ch)
Dhruv Nilkanth (dhruv.nilkanth@graduateinstitute.ch)
Shreya Singh (shreya.singh0@graduateinstitute.ch)

Faculty Lead
Prof. Jérôme Duberry

Partner
Prof. Roxana Radu
University of Oxford

Geneva, Switzerland

2025

Contents

1. Executive Summary

1. This report investigates the evolving role of the United Nations and International Geneva in the global governance of artificial intelligence, with a particular focus on institutional pathways, normative frameworks and policy implications beyond 2025. It highlights the urgent need for global coordination to address risks, ensure ethical alignment, and bridge digital divides brought on by the rapid and opaque advancement of AI technologies. Multilateral institutions, particularly the UN, are increasingly challenged to lead or facilitate these governance processes. These dynamics are examined through three central questions throughout the paper:

1. What is the current and potential role of the UN in global AI governance?
2. What are the core challenges and opportunities for the UN to expand its role?
3. What role does International Geneva play, and how should it evolve?

2. The methodology used was a qualitative, interview-based approach, involving three semi-structured interviews with experts from a UN agency, a think tank, and academia. These explored governance challenges, institutional dynamics, and the role of International Geneva in shaping global AI governance. Insights from the interviews were thematically analysed and supported by institutional mapping and desk research.

3. The findings of the study highlight that global AI governance remains fragmented, facilitative and largely normative in nature. Rather than having established binding regulatory frameworks, multilateral institutions, particularly the UN, play a convening and symbolic role, offering ethical guidance but lacking the technical capacity and political coherence required to drive comprehensive governance. While Geneva retains moral legitimacy as a hub for multilateral engagement, its strategic centrality is declining and its capacity to coordinate across fragmented mandates remains limited.

4. Sharp divergences in regional regulatory approaches, such as the market driven model in the United States, the state-centric approach in China, and the rights-based framework of the European Union complicate harmonisation efforts. Similarly, while private sectors are becoming increasingly central in shaping governance agendas, their participation risks regulatory capture without strong co-governance safeguards.

5. The report identifies several strategic challenges currently facing UN led AI governance. These include a widespread lack of in-house AI expertise across UN agencies, growing coordination fatigue due to overlap of mandates and fragmented initiatives, tensions between regulatory sovereignty and global standardisation, persistent asymmetries in Global South participation, and limited resources for local capacity building.

6. Despite these challenges, the report also outlines a range of institutional opportunities for the UN and Geneva-based actors. This includes leveraging soft law instruments such as conducting impact assessments, setting up observatories and undertaking dataset labelling; institutionalising co-governance through structured platforms to include the private sector, academia and civil society; and positioning International Geneva as a normative hub for inclusive rights-based AI governance, especially for underrepresented regions. The report also proposes enhancing foresight capacities through multidisciplinary scenario teams and theme-specific regulatory sandboxes.

7. In response to these findings, the report makes five core recommendations. First, coordinate rather than centralise AI efforts across UN bodies through a flexible and light-touch framework. Second, invest in technical capacity building through recruitment, partnerships and fellowships. Third, institutionalise co-governance platforms for iterative, participatory normal creation. Fourth, promote soft law mechanisms as flexible and scalable governance tools. Fifth, embed anticipatory governance through foresight, experimental pilots, and adaptive regulation.

8. The report concludes that effective AI governance will not emerge from a single treaty or institution. Instead, it must be co-created through iterative, inclusive, and distributed networks that blend ethical commitments with operational flexibility. The UN and Geneva-based actors must invest not only in norms but also in the institutional capacity to govern AI in a rapidly shifting world.

1. Introduction

As artificial intelligence continues to evolve at breakneck speed, it prompts not only technological fascination but also urgent questions of governance, equity, and foresight. The world stands at a juncture where the stakes of AI deployment are no longer speculative. Whether automating critical infrastructure, influencing erstwhile independent deliberation, or transforming global labour markets, AI is beginning to reshape human systems with a velocity that multilateral governance structures, still calibrated for slower challenges, are struggling to match.

This report takes shape in that in-between space: between technological acceleration and institutional inertia, between normative aspiration and policy traction to catch up. It asks how international governance, particularly through the United Nations system and the multi-stakeholder ecosystem of International Geneva, is responding to the fast-evolving, multidimensional challenges of global AI governance. At the heart of this project lies a central

inquiry: can the UN system and its current Geneva-based architecture evolve meaningfully to govern AI beyond 2025? And if so, what are its tools and technologies in place?

The research began with the intent to not only map institutional efforts but to understand their internal logic, historical inclinations, and projected futures. Through an integrated methodology combining literature review, desk-based institutional analysis, and expert interviews, this study uncovers a terrain defined less by clarity than by *complexity*. Global AI governance today resembles a regime complex: a loosely coupled system of overlapping institutions, norms, and initiatives—spanning ethics guidelines, capacity-building programs, soft law instruments, and regional legal regimes.

The literature explored reflects this fragmentation. It orbits around two core poles: normative ideals such as justice and inclusivity, and explanatory theories rooted in international relations (e.g., realism, liberalism, constructivism). While some scholars draw parallels between AI and internet governance[1], others focus on its potential as a global public good[2]. What is clear is that the crucial interrogatives of AI governance are still under construction, driven by actors, ideas, and institutions that are themselves in flux.

Section 4.2 surveys the institutional architecture of the UN and identifies key inflection points, such as the emergence of the Global Digital Compact, the High-Level Advisory Body on AI (HiLAB), and the establishment of the Office of Digital and Emerging Technologies (ODET). These developments mark a growing awareness that AI cannot be governed from the sidelines. Yet, despite normative ambition, the UN's engagement remains largely facilitative: offering ethical guidance, convening multi-stakeholder dialogues, and producing frameworks, but without a robust enforcement capacity or cross-agency coordination.

This challenge of institutional adaptation is echoed in Section 6, which compiles insights from expert interviews with UN officials, researchers, and policy specialists. Interviewees acknowledged the UN's symbolic legitimacy but raised concerns about technical capacity gaps, duplication of mandates, and geopolitical friction, particularly between the US, the EU, and the Chinese models of regulation. Still, they highlighted emerging spaces of promise: soft governance instruments, co-governance platforms, and capacity-building efforts, particularly for the Global South.

Expanding beyond the UN, Section 6.1 and 6.2 examine other international fora, OECD, G20, EU, BRICS, and the WEF, each advancing distinct governance logics. While the EU's AI Act has set a legal precedent, forums like the OECD focus on principles and data-sharing. These regional and plurilateral efforts both complement and complicate the UN's role, revealing tensions between fragmentation and innovation in the global governance landscape.

[1] Raymond & DeNardis, 2015
[2] Kaul et al., 1999; Dignum, 2025

In Section 4.3, Geneva emerges as a symbolic epicenter, home to a dense ecosystem of digital governance actors including ITU, WHO, UNESCO, and the Geneva Science and Diplomacy Anticipator (GSDA). Long respected for neutrality and human rights leadership, Geneva continues to offer a vital space for dialogue. Yet interview data and desk research suggest that its influence is under threat from newer centers of political and technological gravity, such as New York and Riyadh, and hindered by institutional silos and uneven representation.

Rather than presenting a prescriptive roadmap, this report foregrounds tensions, identifies constraints, and gestures toward opportunity spaces. It argues that effective AI governance will require not a central authority but an iterative, distributed ecosystem - one which is anchored in shared values, experimental regulation, and meaningful public–private cooperation. The recommendations (Section 8) veer to propose steps toward institutional coordination, technical capacity-building, inclusive co-governance, and anticipatory regulatory models.

Ultimately, this report is both a map and a mirror: a mapping of existing structures and actors, and a reflection on the assumptions, risks, and possibilities embedded in them. As such, it offers a grounded yet forward-looking contribution to the global conversation on how we might govern AI, not just as a set of technologies, but as a shared human future.

2. Literature Review

3.1 Artificial Intelligence and Governance: The Steering Factors

The rapid advancement of artificial intelligence (AI) has necessitated a move towards processes to develop global mechanisms for its governance. While the institutional and policy space surrounding global governance of AI is moving rapidly, academic literature surrounding the space has been more focused on definitions and applications of AI, as well as ethical standards[3]. While these two strands of research address the question of "what" is being governed and the technical aspects of AI standards, another direction of research aims to understand the "why" and "how" of global governance surrounding AI[4]. Since the processes governing AI are still in a nascent stage, the literature attempting to explain how these processes are coming about, draws heavily from existing understanding of global governance of the "digital" and "internet" spaces[5].

Many theories have been applied to AI global governance in existing literature. For example, the complex interdependence framework shows how the complexity of relations

[3] Tallberg et al., 2023; Veale et al., 2023
[4] Ibid., 6
[5] Raymond & DeNardis, 2015; Cath, 2018.

between states and non-state actors as well as the interconnectedness of various domains with one another requires a multilateral approach to govern AI[6]. Similarly, viewing AI through the global public goods perspective shows how institutional efforts towards AI global governance by actors such as the International Telecommunications Union (ITU) are driven by recognising or framing AI as a public good, which needs to be governed collectively to mitigate risks and harness benefits[7].

The literature on trying to theorise the "how" and "why" of global governance structures and processes surrounding AI can be divided into two broad categories of normative approaches and empirical-based theorising [8]. The first branch or category on AI global governance structures can be summed up as being normative in nature, as it highlights the importance of having concepts of justice, democracy, inclusiveness, representation and legitimacy amongst others as ideals towards designing AI global governance[9].

On the other hand, existing theoretical frameworks of international relations and political science categorise AI global governance through the lenses of realism, liberalism and constructivism, more specifically by looking at power, interests and ideas respectively[10]. Power-centric approaches explain how states prioritising relative gains, combined with the asymmetry in power, contributes to complex governance and regulation structures coming up. An example of this is the simultaneous yet divergent US, China and EU regulation and standardisation efforts over AI[11]. At the same time, the rationalist-functionalist framework suggests that states move towards global governance through international organisations in case of shared interests and to combat barriers in cooperation[12]. Lastly, looking at idea-centric theories explains how historical contexts and norms shape governance[13]. Examples of this are the cross-border convergence on AI ethical standards and the role of civil society advocacy in the case of EU regulations[14].

In contrast to this broad state-centric approach of looking at traditional IR theories, scholars have also tried to explain AI global governance being brought into being by non-state actors, particularly tech companies but also NGOs and CSOs[15]. This kind of multi-stakeholder approach is visible in the EU's method towards governance of many issues, most recently AI, as visible from its EU AI Act where consultations with stakeholders were mandated[16]. The EU AI

[6] Keohane & Nye, 1977
[7] Kaul et al., 1999; ITU, 2024; G20 Global AI Governance, 2023
[8] Tallberg et al., 2023; Veale et al., 2023
[9] Floridi et al., 2018; Buchanan and Keohane, 2006;
[10] Martin and Simmons, 2012; Lake, 2013
[11] Mearsheimer, 1994; Gruber, 2000; Stone, 2011; Dreher et al., 2022,
[12] Keohane, 1984
[13] Tallberg et al., 2023
[14] Jobin et al., 2019; Ulnicane et al., 2021
[15] Raymond and deNardis, 2015; Radu et al., 2015; Veale et al.; 2023
[16] Donders et al., 2018; Potjomkina, 2018

Act also shows how regional legislation can impact global governance as others learn and try to incorporate EU regulations but also are driven to start thinking about governing AI[17]. This indicates that not solely governance, but the mere framing of AI as a policy-priority needing governance beyond domestic boundaries by either states or private companies also contributes to the formation of global governance processes[18]. Therefore, the AI global governance space can be seen as a regime complex with multiple actors, including international organisations as well as various governing agreements (with varying scope and binding power), that interact to govern AI[19].

This regime complex can be seen as a loosely connected network of partially overlapping institutions that address the same issue without a central authority[20]. Rather than a single, unified regime, AI is governed through a patchwork of forums, each engaging with different dimensions of the technology. Various institutions often pursue parallel efforts with overlapping mandates but divergent legal and normative tools, leading to friction or fragmentation. This configuration can be observed in the landscape of AI global governance with the presence of multiple efforts ranging from UNESCO's AI ethical framework[21], the OECD's principles[22] and the EU's legally binding AI Act[23]. Each of these approaches reflects a different institutional history, with varying stakeholder bases and levels of enforcement capability. Two contrasting perspectives with the Complex Regime literature can be applied to the AI governance space. On one hand, scholars such as Alter and Meunier (2009) argue that the expansion of institutional actors in a particular governance space produces overlapping mandates and jurisdictional competition, as different regimes assert authority across varied domains such as ethics, trade, labour rights and security[24]. This diversity reflects the multi-dimensional nature of AI, but also contributes to institutional ambiguity and difficulties in establishing clear governance pathways. A contrasting perspective from Zürn (2018) suggests that this institutional proliferation may be better understood as layered governance, where newer mechanisms are added to existing structures without necessarily replacing them. This results in a system that evolves through accumulation rather than coordination[25].

In this context, the UN's work with its High-Level Advisory Body on AI (HiLAB) has been identified as a case that reflects the broader complexities of coordinating across institutional boundaries and stakeholders communities[26]. While the HiLAB is still active and is expected to contribute further in 2025, early assessments of its process have drawn attention to a range of

[17] Bradford, 2020
[18] Veale et al., 2023; Funtowicz and Ravetz, 1993; Png, 2022; Bareis and Katzenbach, 2022
[19] Tallberg et al., 2023; Gomez-Mera, 2021
[20] Raustiala and Victor, 2004
[21] UNESCO, 2021
[22] OECD, 2019
[23] European Parliament, 2024
[24] Alter & Meunier, 2009
[25] Zürn, 2018
[26] Knight, 2024

challenges. These include questions around how inclusive and representative the advisory process has been, particularly with regard to civil society and actors from the Global South[27]. Commentary has also noted that while the report proposes new coordination mechanisms, it lacks detailed guidance on how these would interact with existing institutions, and how power asymmetries, especially between states and large technology companies, would be addressed[28]. Further concerns relate to the operational ambiguity of proposed structures and the absence of a clear accountability framework for implementation[29]. These reflections suggest that HiLAB not only exemplifies the UN's effort to navigate complex, overlapping institutional mandates, but also illustrates the persistent tensions in building legitimacy and facilitating cross-sector coordination in the global governance of AI.

Further, many institutions tasked with AI governance are not inherently designed for it. AI has been added onto existing mandates, often without a reconfiguration of institutional capacities. This has led to an expansion of scope within entities such as the ITU, UNDP, and even the IMF, but without corresponding shifts in internal structures or processes. As a result, new governance ambitions sit atop old frameworks that are not always agile enough to address fast-evolving technological challenges. This is particularly evident in consensus-driven bodies like the ITU, where procedural legacies limit rapid institutional adaptation despite growing technical relevance[30].

In addition, institutional responses to AI are often shaped by the historical pathways of each organisation. The theory of path dependency helps explain why global institutions tend to respond incrementally rather than transforming their structures outright. Organisational routines, prior policy commitments, and internal power dynamics make it difficult to deviate from entrenched modes of action[31]. For example, the UN is frequently characterised as a system grounded in state-centric processes with diplomatic consensus and intergovernmental negotiations taking precedence. This institutional orientation is shaped by its historical development and mandate, with many of its practices emerging from the post-war context centred on peacebuilding and development. These structural tendencies may present challenges when engaging with the more dynamic, multi-actor environment of contemporary digital and AI governance[32].

Conversely, actors like the European Union have benefited from institutional sequences that favor proactive digital regulation. Its General Data Protection Regulation (GDPR) and the early investment in its digital rights infrastructure laid the groundwork for its leadership in AI governance. Such sequencing exemplifies how prior institutional design shapes present

[27] Global Partners Digital, 2024
[28] Cleeland et al., 2024
[29] The Outpost, 2024
[30] ITU, 2024
[31] Pierson, 2000; Mahoney, 2000
[32] Weiss and Daws, 2007

capabilities, giving some actors a structural advantage in rapidly regulating emergent technologies[33]. Institutions lacking such trajectories often struggle to match the EU's pace or comprehensiveness.

This combination of institutional layering and historical inertia constrains the evolution of AI governance. Even as norms around ethics, safety, and inclusion gain traction globally, they are processed through governance mechanisms shaped by the past. Understanding how regime complexity and path dependency interact offers important insight into why global AI governance remains uneven and fragmented, and why overcoming these structural constraints requires not only new frameworks, but a rethinking of the institutions themselves.

3.2 Development and Implications at the UN level

As multilateralism changes its ways and forms, we realise that its dialogue with AI and governance need to be dynamic too, especially at the epicenter of an organization like the UN. As we set out to examine the developments in AI governance internationally, we began looking at the ideation stage, beginning in May 2023, when The Common Agenda proposed a Global Digital Compact, to advance a "human-centred, open, free and secure digital future" which is grounded in universal human rights[34]. This approach aligned with an active revisit to the Sustainable Development Goals or SDGs.

The Global Digital Compact seeks digital cooperation, on a global scale, with multiple stakeholders in order to overcome digital, data and innovation divides that currently create a deep chasm in the vision of an accessible digital future. Per the policy brief published by the GDC in May 2023, significant gaps persist across regions, gender, income, language, and age. For example, while 89% of people in Europe are online, only 21% of women in low-income countries have internet access. This digital divide creates barriers to equitable participation in the digital world and hinders progress toward achieving the SDGs. Additionally, a growing "data divide" exacerbates inequalities of ownership and control, leaving many developing countries at risk of being mere providers of raw data, while paying for the services their data help to produce. This disparity in data ownership and control further entrenches global inequality. Another major challenge is the governance gap in the digital realm. As digital technologies are primarily developed by private companies, governments often lag in regulating these technologies in the public interest.

Such global challenges to digital equality, and a potential path forward, were acknowledged by the UN's High-level Advisory Board on AI (HiLAB), through their final report, *Governing AI for Humanity*, in September 2024. Initially proposed in 2020, this body was

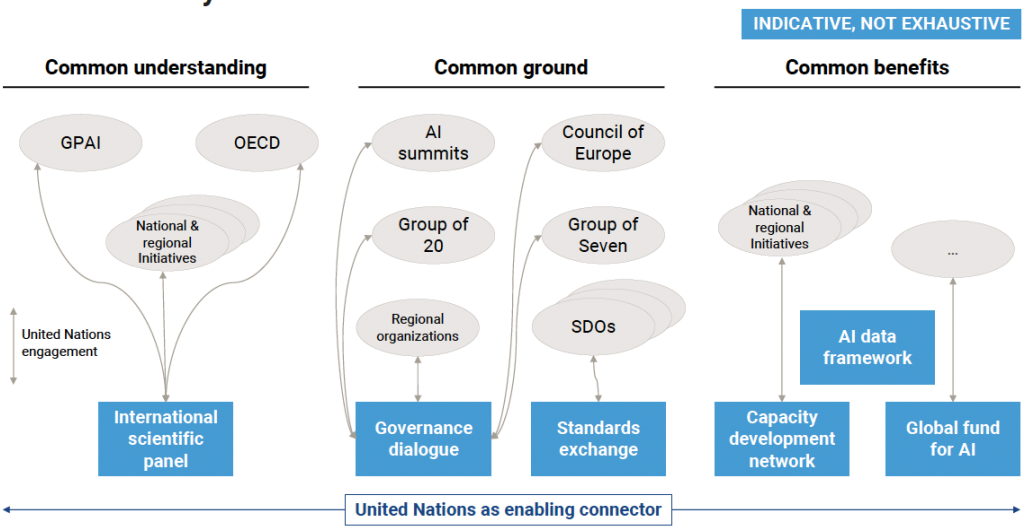
[33] Page, 2006
[34] Global Digital Compact, 2024

formed in October 2023 with the motive to create and provide advance recommendations for the international governance of artificial intelligence.

Recognising the irrefutable argument for the need for global governance, the report also whispers the possibility of unmonitored deployment of AI and its potential misuses, which beckons a holistic, global approach. The report makes seven major recommendations: the 1st recommendation of the report focuses on a "light, agile structure as an expression of coherent effort" in the form of an AI office in the UN Secretariat, working closely with the Secretary-General. The idea for the office is to work closely with both the UN processes and external stakeholders, such as tech companies, civil society, and academia, to ensure a coordinated response to emerging AI issues in as inclusive a manner as possible.
This recommendation for the AI office is sustained by the 2nd recommendation: an independent, international scientific panel on AI, along with an international standards exchange, which draws support from the ITU and UNESCO. The goal is to pool resources and encourage joint investment for global and public interest collaboration(s) and maintain a cohesive inventory of definitions and relevant scientific vocabulary for evaluating AI systems.

The 3rd recommendation focuses on the need of intergovernmental access and policy dialogue for such a multistakeholder tool and project, proposed biannually at the initial stage. The aim would be to exchange best practices on AI governance, development of common understandings to bring both public and private sectors, and to enhance international inoperability.
The 4th recommendation brings to attention a global fund for AI to address the AI divide by providing financial and in-kind support to catalyze local empowerment for the SDGs; it would also be used to support research on data usage and AI models and to create a repository of AI models and curated datasets for SDG applications.

Figure (c): Proposed role of the United Nations in the international AI governance ecosystem



Abbreviations: GPAI, Global Partnership on Artificial Intelligence; OECD, Organisation for Economic Co-operation and Development; SDOs, standards development organizations.

While this diagram[35] aptly summarises the suggestions and propositions of the HiLAB's latest report, it also laid a solid foundation for our research to find and understand how these suggestions materialize, and which ones can be greenlit more feasibly at the earliest.

With the transition of the UN's Envoy on Technology to the newly established United Nations' Office of Digital and Emerging Technologies (ODET), targeted support towards the implementation of the Global Digital Compact is expected, with multi-stakeholder management and facilitation being coordinated by the ODET. As detailed in the Global Digital Compact's main proposal, other UN bodies, in particular the ITU, the United Nations Conference on Trade and Development (UNCTAD), the UNDP and the UNESCO, along with the UNHCR, are mandated to to support the implementation of its guidelines within the UN system. Its final call to the President of the General Assembly at the 82nd Session of the UNGA, to appoint two co-facilitators—one from a developing country and one from a developed country—for this coordination at the GDC, inform us of the scale of their foresight, since the 82nd session is slated to happen in September 2027.

With the establishment of the Office of Digital and Emerging Technologies in January 2025, the specialised focus of the UN on the developments around AI and relevant tech has become quite clear. Hence, the ODET now serves as the central hub for advancing global digital cooperation across five key areas: it provides strategic advice to UN leadership, ensuring informed decision-making on emerging technologies. As the advocacy focal point, ODET also engages

[35] UN High-level Advisory Body on AI, 2024

Member States and stakeholders to shape international digital policies. It fosters multi-stakeholder policy dialogue, creating inclusive spaces for discussions on AI governance, cybersecurity, and other critical issues. Additionally, the office works to strengthen UN system-wide coordination, enhancing collaboration across different agencies on digital matters. Finally, the ODET plays a crucial role in implementing the Global Digital Compact, supporting efforts to translate the commitments made at the Summit of the Future in September 2024 into actionable policies and initiatives. While these mandates do not vary much in essence from the former Tech Envoy to the Secretary-General–which the ODET has come to replace–the constitution of the office itself and its budgetary discussions make the transition and probably the grounding reasons clear. Although there are only estimations available for the budgetary decisions regarding the ODET in the UNGA's Fifth Committee documents, it can be assumed that the decisions would reflect centering significance and personnel both in New York, as compared to the Geneva.

3.3 The Role of International Geneva in AI Governance

The fragmented nature of AI governance stems from differing national priorities, technological capacities, and cultural perspectives[36]. This fragmentation leads to inefficiencies, redundant efforts, and missed opportunities for global collaboration, underscoring the need for more coordinated governance. Despite growing global interest, many organizations remain siloed, struggling to form a unified voice on AI governance[37]. This calls for a more collaborative approach to ensure AI governance reflects the diverse needs of the international community.

Ambassador Valentin Zellweger, Permanent Representative of Switzerland to the Office of the United Nations in Geneva highlights,"The strengths of Geneva are that there are many actors from different stakeholder communities here, with a culture of working together. As the division between political spaces plays into issues of regulation, Geneva with its impartiality and neutrality comes in as a place to get solutions."[38] According to the Geneva Internet Platform as of 2020, 50% of the digital policy issues they have mapped are addressed in Geneva. The city has been a hub for neutral convenings and multi-stakeholder dialogues, particularly around Internet governance. Several Geneva-based institutions and processes are making significant strides in shaping global AI norms and policies, including the United Nations Office at Geneva (UNOG), and the ITU, the World Summit on the Information Society (WSIS) and the Internet Governance Forum. These platforms have facilitated global discussions on AI ethics and policy in recent years.

Other Geneva-based organizations are also contributing. ILO is investigating AI's impact on the future of work, while Diplo Foundation's HumAInism project explores how AI can help draft a social contract for the AI era. The Geneva Academy focuses on ensuring human rights law is central to AI governance, and ICT4Peace has long addressed AI's ethical challenges, including for autonomous weapons. The canton also hosts the Digital Health and AI Research Collaborative (I-DAIR), which promotes responsible AI research for health, and the Geneva Science and Diplomacy Anticipator (GESDA), which is addressing advanced AI's global impact. Swiss Cognitive fosters an AI ecosystem by connecting industries, organizations, and startups to explore AI's potential, and its CognetiveValley initiative promotes the Swiss AI ecosystem worldwide.[39]

The growing number of Geneva-based initiatives places the city into a global hub for AI governance, being a necessary platform for cooperation and exchange. However, escalating geopolitical tensions and the competitive struggle for global leadership in AI innovation, discourages previous efforts to establish global ethical guidelines, due to a lack of international

[36] Kende et al., 2020
[37] Radu, 2024
[38] Kende et al., 2020
[39] Ibid.

consensus. This underlines the growing need for Geneva's convening power in order to establish consensus and a common language of AI governance.

3. Methodology

Our study employs a qualitative research design to explore existing governance structures on AI within the UN and its specialized agencies, and to map how these structures can evolve after 2025. Floridi et al. (2018) argue that AI governance must address ethical, societal, and technical challenges. They emphasize the context-dependent nature of governance, shaped by evolving societal dynamics, regulatory demands, and ethical issues. In line with these observations, our research seeks to understand such complexity through the application of qualitative methodology. As Tuli, F. (2010) notes, qualitative methodologies are optimally utilized in exploring phenomena in their naturalistic environments with focus on the "how" and "why" of emerging patterns. This makes them particularly useful to investigate how and why AI governance arrangements develop and transform.

4.1 Desk Research

We began with an initial desk scan of governance theories, the background of UN initiatives in the field of AI, and Geneva's institutional role towards international governance. This included collecting and reading in close detail documents by UN agencies such as UNESCO, UNICEF, IMF, and the ITU and the UN's HiLab. We also mapped AI-related activities from beyond multilateral fora and partner institutions including the G20, BRICS, the EU, and OECD.

We then conducted a comparative review to examine how various institutions approach AI governance. As guided by George and Bennett (2005), the comparative analysis helped us to uncover patterns and divergences that then led to emergent key themes and shed light on the strengths and shortcomings of AI governance.

At this stage, we also conducted a SWOT analysis to identify the strengths, weaknesses, opportunities, and threats in current UN initiatives on AI governance. The analysis was vital to determine gaps and problems within existing frameworks and evaluate the range of improvement. Outputs from the analysis informed our recommendations for evolving AI governance structure.

Overall, this phase also sought to uncover future challenges and opportunities for international AI governance. Lessons learned through this review provided a solid foundation for our interviews and acted as input for a broader horizon-scanning exercise. These findings also guided the drafting of our forward-looking recommendations.

4.2 Semi-Structured Interviews

We conducted three semi-structured interviews to discuss current UN governance frameworks, what their problems are, and how AI-related policies could shift after 2025. This format allowed us to address pre-formulated questions while remaining flexible enough to uncover new information that came organically during the interviews. (Alsaawi, 2014).

We used a *purposive sampling strategy*, where researchers rely on their judgment in selecting participants that meet the needs of the study. Palinkas et al. (2015) indicate that this helps to ensure participants possess the right expertise to offer perceptive comments. Successive desk research further narrowed down the choice criteria. Our goal was to obtain internal perspectives (inside the UN system and its related specialized agencies) as well as external perspectives (from academics and policy specialists). Possible interview candidates were identified by reviewing academic course rosters, LinkedIn, and through previously studied research reports. From a master list of names, we contacted individuals using LinkedIn and email and ultimately conducted three interviews with AI governance specialists having extensive expertise.

Respondent	Description	Date
A	Researcher in international economic law, with a focus on trade law and data regulation in the context of AI governance.	12.03.2025
B	UN special agency official with over sixteen years of experience; Advisor to the Deputy Director-General in the 2 years; led coordination of AI-related efforts across field offices.	27.03.2025
C	Independent consultant and think tank professional, specializing in internet development and digital governance.	10.04.2025

A written consent form was shared in the initial contact email for signature to ensure ethical research practices.[40]

The interviews covered five thematic sections, including[41]:

- Background and professional experience of the respondent
- AI governance landscape

[40] For the informed consent form, check Appendix 10.1
[41] For the questionnaire, check Appendix 10.2

- Institutional developments and comparative frameworks
- Future of AI beyond 2025 and the role of international Geneva
- Recommendations for future AI governance frameworks

Since the interviewees were geographically dispersed, the interviews were conducted on virtual platforms such as Google Meet, Zoom, and Microsoft Teams. The interviews lasted between 45 minutes to an hour and were conducted under the Chatham House Rule to ensure confidentiality and frankness.

The data from the interviews were analyzed using an inductive, thematic approach (Braun & Clarke, 2006). The texts were coded into sub-themes from which emerging themes were developed. These were synthesized into informing the perspectives, interpretations, and suggestions presented in the paper.

The method acknowledges the evolving nature of AI governance, knowing that the environment is rapidly changing. Our study is conducted until 13 April 2025, since AI governance policies and frameworks continue to unfold globally. The key limitation is the small sample size of the interviews, owing to logistical challenges in accessing high-level experts and obtaining clearance from the primary stakeholders in UN agencies and international organizations. Bureaucratic hurdles and confidentiality concerns contributed to the complexity of scheduling interviews. Additionally, the purposive sampling method, although necessary to select relevant experts, may have introduced bias by excluding some perspectives, particularly because of the inaccessibility of underrepresented organizations or geographies. As AI governance practices evolve, future research may need to revisit and refine these findings.

4. Mapping the AI Institutional Landscape

5.1 Mapping Key Actors: UN Agencies and Organisations

UNESCO, focusing on ethical considerations, produced the first global standard for AI ethics—the Recommendation on the Ethics of Artificial Intelligence—in 2021, adopted by 193 member states. This framework is implemented via the Global AI Ethics and Governance Observatory, which equips tools (e.g., the Readiness Assessment Methodology, RAM) to assess readiness for ethical AI implementation in any given country. Through the UNESCO AI Ethics and Governance Lab, research, toolkits, and case studies are also fostered to encourage responsible innovation. Initiatives such as “Women4Ethical AI” exemplify its focus on inclusivity, both concerning gender underrepresentation in design and implementation of AI.

Other institutions have addressed AI's broader socio-economic implications. During 2023, the International Labour Organization (ILO) inaugurated the Observatory on Artificial Intelligence and Work in the Digital Economy, with the aim of understanding how AI can affect

employment, productivity, and the rights of workers. Analytics around algorithmic management and digital labour platforms demonstrate the ways in which AI can reshape the workplace (International Labour Organization, n.d.). Likewise, the International Monetary Fund (IMF) created the AI Preparedness Index (AIPI) to evaluate the ability of countries to incorporate AI, and highlighting AI's double function to destroy and to improve economic structures (International Monetary Fund, n.d.).

The International Telecommunication Union (ITU) plays a key role in global telecommunications and ICT standard-setting, for example, its various initiatives like the ITU-T Study Groups for international standards, the ITU Focus Group on 'Machine Learning for Future Networks'[42], and the one on Environmental Efficiency for AI and other Emerging Technologies . Since 2024, the ITU has intensified efforts in AI governance, hosting the AI Governance Day at the "AI for Good" Global Summit to facilitate discussions on regulatory trends and policy frameworks. It has worked along with other multilateral AI initiatives, including China's Algorithm Registry, the US Executive Order on AI, and the EU AI Act. Collaborating with WHO, WIPO, and FAO, ITU aims to integrate AI into health, agriculture, and education. Known for its technical and regulatory expertise, the AI governance frameworks[43] that the ITU is attempting to formulate on multiple fronts, with different organisations and initiatives, strengthen its position as a relatively neutral convener for international AI discussions.

In 2024, the UN High Commissioner for Human Rights (UNHCHR) issued a set of practical guidelines on human rights impact assessments (HRIAs) for AI systems, highlighting how to apply them across the lifecycle of AI deployment. The guidelines emphasise safeguards for non-discrimination, privacy, and freedom of expression, and were promoted at the Geneva AI Human Rights Dialogue in December 2024. In parallel, the UN Conference on Trade and Development (UNCTAD), through its Commission on Science and Technology for Development (CSTD), launched the "AI for Development" program in early 2025. This initiative aims to assist developing countries in assessing trade and innovation policy readiness and in aligning national AI strategies with inclusive economic development goals. It also includes regional consultations on the Global Digital Compact and AI capacity-building for least developed countries. The 28th Session of the CSTD held in Geneva in April 2025 also saw multiple discussions on AI governance ranging from AI and human rights, using AI for development as well as building the ideal WSIS to govern AI and digital technology in the near future.[44]

These efforts are primarily orchestrated by a heterogeneous set of institutions which are also responsible for organizing worldwide AI summits and supporting international collaboration. For example, UNESCO is organising the first Global AI Forum in the Asia-Pacific region in

[42] Digital Regulation Platform, The World Bank x ITU
[43] International standards for an AI enabled future, *AI For Good Summit*
[44] UNCTAD, 2025

Thailand in June 2025, and the World Economic Forum Centre for the Fourth Industrial Revolution in Rwanda will organised the Global AI Summit in Africa in April 2025[45]. The work of those institutions is predominantly focused on the ethics of AI, initially focusing on shaping governance structures that are, however, still in their infancy.

5.1.2 The Global Fora: Prominent International Initiatives Beyond the UN

A. OECD

Institutional efforts toward regulating AI are being spearheaded by a number of international organisations, each focused on diverse issues related to AI development, deployment, and ethics. The Organisation for Economic Co-operation and Development (OECD) has been one of the pioneers, starting its work in AI in 2016. By 2019, it introduced the Council Recommendation on AI, which outlined ten principles for trustworthy AI, serving as a cornerstone for international efforts. This was then followed by the creation of the OECD AI Policy Observatory, a dashboard monitoring more than 1,000 AI projects globally and providing real-time information and data. The OECD's Working Party on Governance of AI examines the relative importance of these recommendations and their ability to respond to and adapt to the challenges posed by technological innovation, and its Network of Experts is mobilizing more than 350 international experts to tackle highly pertinent, open questions including AI risks, AI accountability, and AI and climate (OECD, 2024).

B. WEF

The World Economic Forum (WEF) has been a pioneer in its AI Governance Alliance, introducing the Presidio AI Framework, in order to cope with generative AI issues. This framework introduces comprehensive risk mitigation strategies across the entire lifecycle of AI models, from design to retirement. The WEF's AI Governance Alliance has also drawn attention to problems concerning fragmentation of risk management and nebulous interpretations of safety and traceability.[46]

C. G20

The G20 has been an important actor in shaping the global governance of AI through targeted initiatives that emphasize ethical, sustainable, and inclusive AI development. The São Luís Declaration, released by the G20 Working Group on Artificial Intelligence in September 2024, focuses on leveraging AI to advance Sustainable Development Goals (SDGs), reduce poverty, and address inequalities in the digital transition. It calls for the involvement of all

[45] Centre for the Fourth Industrial Revolution Rwanda, 2024
[46] WEF, 2024

stakeholders including states, IOs, NGOs and CSOs in the global governance of AI. The declaration also calls for the mitigation of algorithmic biases related to gender and race as well as tackling disparities in AI accessibility[47].

Similarly, the Maceió Declaration issued during the G20 Ministerial on the Digital Economy in September 2024, raised concerns about the concentration of AI development in the hands of a few multinational companies. It emphasised the need to protect digital sovereignty and personal data security, while also addressing risks posed by AI in areas like social media, where data manipulation can threaten democracy and transparency. It also highlighted the importance of upholding the principles of AI governance put forward by international organisations like the UN, UNESCO and OECD[48].

D. BRICS

BRICS has prioritised AI governance, beginning from the 2015 Memorandum of Understanding on Science, Technology, and Innovation, including AI as a key area of collaboration. In 2017, the group mentioned AI, specifically, in its joint declaration, stressing the importance of new cooperation to promote the development of ICT, and reiterated its understanding of AI as a driving force of, and the key enabler for, economic growth, technological progress, and inclusive development. At the 2023 Summit in Johannesburg, members of the BRICS decided the established AI Study Group within the BRICS Institute of Future Networks marked a significant step toward institutionalising AI cooperation[49].

Similarly, the BRICS-led New Development Bank (NDB) has been putting money into AI applications — for example, into smart city projects in China — indicating growing financial commitment to innovating in AI across member states. China has played a leading role in advocating for the need to develop an AI governance framework and recently launched a China-BRICS Artificial Intelligence Development and Cooperation Center. In his statement at 16th BRICS Summit in Russia, Xi Jinping, President of the People's Republic of China, said that to deepen their cooperation on innovation, they will establish a BRICS Deep-Sea Resources International Research Center, a China Center for Cooperation on Development of Special Economic Zones in BRICS Countries, a China Center for BRICS Industrial Competencies, and a BRICS Digital Ecosystem Cooperation Network. Previously, the Center for Long-term Artificial Intelligence (CLAI), in collaboration with the International Research Center for AI Ethics and Governance hosted at Institute of Automation, Chinese Academy of Sciences, jointly initiated

[47] São Luís Declaration, 2024; Radu, 2024
[48] Ibid.
[49] BRICS Nations to Establish a Study Group to Track AI, n.d.

the development of the ‘AI Governance InternationaL Evaluation Index’ (AGILE Index) to benchmark AI governance globally[50] .

Despite these efforts, China's research tends to be somewhat disconnected from global AI governance networks, highlighting the broader difficulties BRICS faces in engaging with international frameworks[51]. The debate over a unified ethical framework will be crucial, though BRICS' efforts are still lagging the European Union's more established AI Act[52].

E. European Union

The EU became a global pioneer by adopting the EU AI Act in March 2024[53], and becoming the world’s first comprehensive AI law at that scale. It is important to note that the EU’s AI act is more on the defensive of the potential risk of AI and less on the integrative aspect of AI and governance. The Act categorises AI-related risks into Unacceptable, High-Risk and lower levels, while mentioning some transparency requirements to be mandated to regulate its use and steps towards innovation as it moves ahead. The Unacceptable AI groups AI systems that could potentially threaten cognitive or behavioural processes, as well as biometric identification systems. It seeks to outrightly ban them while marking an exception for strict law enforcement purposes. High-Risk AI refers to those which impact safety or fundamental rights, including product safety regulations (toys, medical devices) and AI applications in critical infrastructure which would only be allowed circulation after they have undergone risk assessment to safeguard the fundamental rights of EU citizens. High-impact, general-purpose AI models, most popularly GPT-4, are supposed to be thoroughly evaluated and their usage be monitored and declared to avoid any systemic risks that might permeate in the long run. Finally, the Act also aims to foster innovation by providing opportunities for start-ups and small businesses to develop and test AI models before making them publicly available.

At length, this Act is a cautious, vigilant step towards accepting the large unknown of the AI while ensuring maximum regulation for the realms which are in the know already. This Act certainly foresees one aspect of the future with AI which can be emulated to an extent at the UN level—the utter respect for user safety and privacy[54]—but it does seem to be too cautious and less reliant on research itself to counter any potential threats.

In addition to its regulatory strategy, the EU has stated a more general strategic goal to establish itself as a world leader in reliable, human-centered AI. The European Commission described plans to increase investments in AI research, innovation, and infrastructure in its 2024 AI

[50] Centre of Long-Term Artificial Intelligence, n.d.
[51] Xin & Xin, 2024.
[52] BRICS Wants to Shape Global AI Governance, Too, 2024
[53] European Parliament, 2024
[54] Mügge, 2024

Continent communication, with a focus on public-sector use and industrial deployment [55]. The EU is making significant investments in AI ecosystems, skill development, and cross-border cooperation through programs like the Digital Europe Programme and Horizon Europe[56]. By supporting technological sovereignty and bolstering the EU's long-term vision of digital leadership rooted in democratic values, this strategy enhances the AI Act.

5.2 The Shift to ODET: UN's Future-Proof System of Digital Cooperation

While the Global Digital Compact was adopted in late 2024, the emergence of the ODET and its primary agenda focussed on the implementation of the GDC, leave ground to explore the political significance of the Compact. While the erstwhile UN Office of the Secretary-General's Envoy on Technology (or, Tech Envoy) had primarily played a convening and agenda-setting role since it was convened in 2021, particularly in the lead-up to the GDC, the establishment of ODET reflects a broader strategic imperative: to move from dialogue and norm-building towards direct coordination, implementation, and policy integration within the UN system. As part of the Executive Office of the Secretary-General (EOSG), ODET is now positioned at the heart of multilateral digital governance efforts, with a clear mandate to operationalise global digital cooperation in line with the shifting geopolitical and technological tectonics. This transition signals the cross-cutting priorities of AI and its imminent implications for peace and security, development, climate, and human rights.

One of the most significant developments shaping this transition is the expanding role of private sector actors, particularly major technology companies, in the UN's digital governance ecosystem. Over the past few years, big tech companies have moved from being peripheral stakeholders to central participants in multilateral forums. Through initiatives like the Tech Envoy's Roadmap for Digital Cooperation and now through ODET's multi-stakeholder tracks, these companies are increasingly being brought into structured dialogues around setting global norms—particularly on issues like AI ethics, data stewardship, platform accountability, and algorithmic transparency. This involvement is both strategic and pragmatic. On one hand, the UN recognises that many of the digital infrastructures and AI systems that shape global societies are developed, owned, and deployed by private companies, making their participation in governance design essential. On the other hand, the active inclusion of big tech in norm-setting spaces has raised concerns about regulatory capture, conflicts of interest, and power imbalances—especially from civil society and Global South governments. ODET's challenge presently appears to be to navigate these tensions, ensuring that the GDC and its associated frameworks reflect public interest values, protect rights, and prioritise equity, while still benefiting from the technical expertise and resources of the private sector.

[55] European Commission, 2024a
[56] European Commission, 2024b

In this context, ODET serves as both a technical coordination hub and a political negotiation space[57], bringing together governments, industry leaders, academic experts, and civil society under a common framework. Its success will depend on maintaining legitimacy across diverse stakeholder groups while enabling meaningful implementation mechanisms for digital governance. By anchoring the 2025 Paris AI Action Summit's outcomes within the GDC and elevating private sector accountability alongside government commitments, the UN is making a visible attempt at creating a more balanced, enforceable, and future-proof system of digital cooperation: one that acknowledges the realities of power and infrastructure, without compromising on the foundational principles of equity, inclusion, and multilateralism.

6. Findings and discussion

The following sections analyse the interview and data findings in relation to the research questions, to interpret the role of the UN in AI governance, the structural and strategic challenges it faces, and the evolving function of International Geneva in this landscape.

6.1 What is the current and potential role of the UN in global AI governance?

Through the interviews we conducted, there was a shared understanding of the fragmented, evolving, and largely facilitative nature of attempts at AI governance throughout the UN system. While all the experts brought a unique institutional view, a general agreement was on the observation that the UN functions more as a convener and a facilitator of dialogues than as a regulator. The current institutional set-up is defined by fragmented initiatives led by various UN agencies, i.e. ITU, UNESCO, UNDP, UNCTAD, ILO, and the ODET. These initiatives range from ethics codes to observatories and forums but are predominantly normative in character and non-enforceable. Speaker A described these efforts as necessarily multi-stakeholder in design but diffuse in impact.

In terms of temporal orientation, all three experts put current UN activities in a sequence of incremental development. In the short-term future, the focus appears to be laid on internal capacity development, stakeholder mapping, and building soft tools such as ethics guidelines or observatories, though as one speaker noted, "we already have a lot of those." Speaker C emphasised the requirement for Global South capacity development and public outreach through town halls and expert panels. In the medium term, there is some cautious optimism about the potential consolidation of ethical norms and voluntary standards. These may not be binding but

[57] A new UN Office for Digital and Emerging Technologies, UN Press Release

could nevertheless acquire global traction through widespread adoption, especially when reinforced by regional legislation like the EU AI Act. In describing the long term, each of the three experts questioned the likelihood of binding global frameworks emerging under the UN. Speaker A noted that even though the UN may mediate tensions between different AI governance models, it cannot be expected to serve as a treaty-broker entity – limited by rising geopolitical divides and the principle of regulatory sovereignty being upheld by powerful states.

The challenges to bolder UN leadership were addressed extensively by all experts. The most frequent structural barrier mentioned was the absence of in-house technical expertise in AI at the UN agencies. Most employees in relevant UN agencies were not originally recruited with AI or data science skills, and reskilling so far has been spotty and insufficient. Speaker B labeled this lack of skill as a considerable limitation to the UN's ability to meaningfully interact with AI innovation. Subscribed to these are political challenges. Speaker A emphasized the position of great powers, referencing China's growing influence in the ITU, which has led to decreased trust among certain Western actors. Meanwhile, Speaker C referred to the challenge of reconciling fundamentally different global data governance models i.e., the firm-based paradigm of the United States, the citizen-centric paradigm of the European Union, and the state-centric paradigm being advocated by China. Such paradigmatic differences largely undermine prospects for regulatory harmonisation.

Along similar lines, the idea of “coordination fatigue” was mentioned. Speaker B illustrated that though the UN often hails inter-agency collaboration, little is ever measured for effect. As they put it, “Collaboration is only as good as what it delivers. We should be focused on results and looking at the means to do that rather than the means as a result”. In all the interviews, there emerged a unanimous opinion that coordination remains more rhetorical than real, and it results in duplicative mandates, redundant efforts, and fragile institutional cohesion.

The UN’s present role in global AI governance, as supported by all interviewees, aligns most closely with a norm-framing, convening, and facilitative model. This is in line with thinking that regime complexity really highlights (Raustiala & Victor, 2004), where instead of a strong central regulator, governance is a fragmented ecosystem of overlapping players and norms. We observed that the UN, through bodies like ITU and UNESCO and now ODET, has focused on developing ethical guidelines, establishing observatories, and getting different stakeholders together to dialogue. However, the interviews show that these efforts still amount to soft law- non-binding, discursive, and often duplicative.

Literature suggests a key way of improving government cooperation is to frame artificial intelligence as a public good, and several UN documents (e.g., the GDC and HiLAB reports) have attempted to do just that. However, the interviewees were skeptical about the UN’s capacity to move from convening to coordination, let alone enforcement, due to entrenched geopolitical

tensions, regulatory sovereignty, and institutional inertia. These worries align with some ideas in International Relation theories that emphasise that since states vary so much in how powerful they are and in their interests, this disparity means it's hard to form effective collective governance. The divergence between the U.S., China, and EU approaches to AI regulation, mentioned by both literature and interviewees, reflects these divides.

Yet, this limitation does not negate the UN's utility. Interviewees pointed to the UN's symbolic legitimacy, especially on ethical and social justice questions, as a unique comparative advantage. As Floridi et al. (2018) emphasise, AI governance must be as much about justice and inclusiveness as it is about rules, suggesting that the UN's role may not be regulatory but normative, helping define the moral vocabulary of global AI governance.

6.2 What are the core challenges and opportunities for the UN to expand its role?

In spite of these problems, the interviews highlighted several strategic opportunities. Speaker A emphasised the UN's convening power and moral authority, notably on contentious ethical issues such as algorithmic bias and social credit scores. Speaker C emphasised the growing salience of soft law mechanisms such as observatories, impact assessments, and arrangements such as 'dataset nutrition labels'[58] that disclose the composition, potential biases, and representativeness of AI training data. Although non-binding, these tools were considered part of a broader movement towards incorporating transparency and accountability into the design and deployment of AI systems. Another area of opportunity was in helping under-resourced countries. All the experts spoke of the UN's special role to offer capacity-building and set up inclusive governance processes. This role was considered the basis for the long-term building of equitable AI governance. Public-private engagement was identified as an under-exploited field too. Speaker B stated that private sector stakeholders have expressed their interest in participating in regulatory debate.

The most visible challenges, consistently raised by the interviewees, were institutional. The UN system was described as under-equipped technically, and overburdened structurally, with AI governance layered onto pre-existing mandates without institutional adaptation. It reflects the argument often made in literature on path dependency (see Pierson 2000), which tells how the legacies of history shape evolutionary changes in institutions. Bodies like the ITU and ILO were

[58] The Dataset Nutrition Label, similar to a food's Nutrition Facts Label, offers a standardized overview of a dataset's ingredients, ensuring better data quality before AI model development. This helps to spot biases, explainability, and encourages better data collection practices to enhance AI model outcomes.

not designed with emerging technologies in mind, and the ad hoc reskilling of staff cannot match the technical pace of AI.

Politically, the problem of coordination fatigue echoes criticism that's in the literature's critique of multi-stakeholderism, that lacks enforcement and coherence (Donders et al., 2018). Furthermore, there seems to be an absence of structured platforms for public-private collaboration, in a way that people who actually build the technology are involved in policy discussions too. The regulatory vacuum, coupled with limited technical legitimacy, weakens the UN's ability to meaningfully influence industry norms.

Yet, there are possibilities. The UN can strengthen its mandate for capacity building, especially for the developing world. This would reinforce the Global Digital Compact's emphasis on closing the digital and data divide, and is within the constructivist perspective of global governance as shaped by ideas, not just power. Moreover, interviewees referred to growing momentum for soft governance mechanisms; such as dataset nutrition labels, transparency regimes, and observatories, that require no binding agreements but can shape practice. Such mechanisms echo the shift toward reflexive regulation: iterative, inclusive, and norm-creating but not prescriptive.

In addition, the scholarship on co-governance models (Harvard Law Review, 2024) presents a more participatory counterpart to top-down rule-making. Interviewees implicitly echoed this logic in calling for multi-stakeholder processes, town halls, and Global South representation. While the UN may not constitutionally be well-suited to mandate a one-size-fits-all approach, it can test or facilitate networked governance models where state and non-state actors iteratively co-develop adaptive standards.

6.3 What role does International Geneva play, and how should it evolve?

Lastly, the potential of International Geneva as a hub for AI regulation was discussed with some reservation. While all of the speakers praised Geneva's time-honored norm-setting legacy, they also signaled towards indicators of decreasing prominence. One spoke of Geneva as a "workshop" in which ideas are crafted but decisions increasingly get acted out elsewhere in the world, such as New York and Riyadh. Speaker B testified to institutional silos' resilience even in Geneva, while Speaker C noted that participation in AI-related activities has a strong bias towards affluent states at the expense of smaller states. However, all three speakers highlighted that Geneva's potential, comparative to New York, serves as a better bridge between Global North and South to enable more representative debate about global digital futures. As one expert put it, "Governments are already moving faster on AI than they did with privacy or social media", a development which underscores both the need and the possibility for Geneva to return to being a trusted, values-based node in global AI governance.

We observed that Geneva is still a critical site within global digital governance, but increasingly one that is only symbolic in its power, rather than strategic. The research showcases institutional layering (Mahoney, 2000) and the fragmentation of digital regimes (Zürn, 2018) in the context of cities like Geneva where overlapping of different forums is observed, but does not have a central authority to lead them. Geneva has unique strengths, historical neutrality, record of human rights, and a dense ecosystem of agencies, that still render it a natural context for norm-setting processes, especially contexts that are ethical and rights-based. However, two structural limitations were raised. First, institutional silos have emerged, which is a real concern alongside asymmetrical participation across forums, and declining political engagement against New York. This is solidified due to a shift in funding and a stronger concentration of tech companies, as compared to Europe.

With that being said, Geneva's future may be less about reclaiming relevance, and more about recasting its value proposition. Interviewees and other commentators note the possibility of Geneva serving as a bridge, including: (1) between the Global North and South; and (2) between policy and ethics. It was noted that Geneva's ability to bring together those voices that have been underrepresented, and to frame emerging norms for AI in terms of human rights considerations, may provide a normative ethical basis for a future framework for governance processes, even if the technical decisions are made elsewhere if at all.

While inspecting the role of International Geneva in 2025 and beyond, with respect to the regulations around AI governance, the role of ITU beckons a mention in detail. The ITU is deeply rooted in Geneva's multilateral ecosystem and has always worked closely with the UN agencies, global development initiatives and other regulatory bodies – it remains the UN's primary standard-setting body for technological innovations and stemming issues of global frameworks for the same. Parallelly, the more agenda-oriented ODET shapes high-level policy discussions in New York, including the implementation of the Global Digital Compact and the WSIS+20 Review. To better understand the same, the following SWOT analysis of the ITU in Geneva and the ODET in New York would better demonstrate the current interaction, friction and complementary tangents of the two bodies.

Strengths: The ITU excels in technical standard-setting, regulatory frameworks, and AI governance implementation, benefiting from Geneva's multilateral ecosystem and longstanding expertise in digital infrastructure. It ensures neutrality in international discussions on AI governance. ODET, on the other hand, has strong political influence due to its position within the UN Headquarters, enabling it to shape global AI policies through diplomatic engagement.

Weaknesses: The ITU's bureaucratic structure and consensus-based decision-making can slow down the implementation of AI governance policies. It also struggles with balancing commercial interests with public-sector needs. ODET, meanwhile, lacks technical depth compared to the ITU

and depends on external agencies for expertise, which can limit its ability to enforce governance standards effectively.

Opportunities: The ITU has the potential to expand its role as the global AI standard-setting body, fostering collaboration with emerging tech governance organisations and increasing its influence in developing economies. ODET can strengthen a multi-stakeholder engagement on AI governance, leveraging New York's policy-driven environment to advance global agreements on AI ethics and security.

Threats: The ITU faces competition from private-sector entities like Big Tech companies, which often push their own governance frameworks. Additionally, geopolitical tensions could impact funding and regulatory consensus. ODET risks becoming politically fragmented as AI governance discussions are influenced by national interests and evolving global policies, making diplomatic negotiations more complex.

Concluding it, both agencies are increasingly interdependent: the ITU provides the technical backbone, while ODET sets the political vision and global priorities. However, New York increasingly holds more sway in normative leadership and global legitimacy, particularly through the GDC and post-Paris AI governance pathways.

7. Recommendations

AI governance evolves rapidly, moves across borders with ease, and impacts all of us. Based on our empirical data and analytical insights, we make the following recommendations to the United Nations and Geneva-based stakeholders which are informed by a grasp of institutional path dependence, regime complexity, and geopolitical dynamics:

7.1 Strategically coordinate, rather than centralise, AI initiatives across the UN system

Because of the regime complexity of global AI governance (Raustiala & Victor, 2004), unification into a single institution does not seem to be an actual prospect. The UN should, rather, encourage greater strategic coordination between its existing institutions.

While fragmentation will encourage inefficiencies, it also makes room for norm innovation and adaptability. A light-touch coordinating framework, perhaps integrated into the Global Digital Compact's post-2024 implementation or placed beneath the ODET, might be capable of doing so by:

- Mapping agency requirements and AI initiatives
- Facilitating information sharing and thematic coordination
- Acting as a soft hub (not top-down directive)

Such coordination would prevent duplication, ease norm diffusion, and preserve the UN's credibility in the AI governance realm.

7.2 Address institutional path dependencies through technical capacity-building

As discussed, institutions tend to stick to inherited structures (Pierson, 2000), and many UN bodies were not designed to grapple with AI. To break from these limitations, they must invest in sustained, cross-disciplinary expertise. We recommend:

- Hiring professionals with a background in engineering and AI expertise
- Partnering with universities, think tanks, and innovation hubs for fellowships and secondments

This will boost operational relevance and epistemic legitimacy, a crucial element for soft governance instruments to gain global traction.

7.3 Institutionalise inclusive co-governance frameworks

As experimental governance scholarship has claimed (Harvard Law Review, 2024), public–private hybrid models offer more flexible, adaptive governance frameworks. The UN should create formal, institutionalised co-governance platforms over informal advisory processes. These could include:

- Joint working groups on algorithmic accountability
- UN–industry partnerships rooted in the public interest
- Technical roundtables for experimenting and co-designing voluntary standards

This is a step towards reflexive governance, wherein legitimacy is built on participation, transparency, and collective problem-solving.

7.4 Promote soft governance instruments of normative sway

Soft law tools, e.g., non-binding frameworks, impact assessments, and reporting directives, are especially suitable for fast-changing fields like AI (Abbott & Snidal, 2000). The UN can foster the development and uptake of:

- Dataset nutrition labeling (provenance, bias risk, representativeness)
- Algorithmic impact assessments (most notably in the public procurement)

- Transparency reporting for AI within sensitive or high-stakes domains

Although non-binding, these instruments shape action by establishing normative expectations, particularly when underpinned by multilateral legitimacy or linked to funding and procurement obligations.

7.5 Leverage Geneva's ecosystem to embed foresight processes and pilot adaptive governance models

As we approach frontier and potentially superintelligent AI[59], anticipatory, proactive governance becomes more pressing by the minute. Early work suggests that the landscape for this is politically contested: while there exist actors prioritising long-term existential risks, such as those posed by superintelligent AI, there exist other actors who claim that devoting attention, resources, and time to such a focus takes it away from more urgent, near-term harms, such as algorithmic discrimination, surveillance, and automation of work (Hoes & Gilardi, 2025). These competing agendas affect institutional agendas, research directions, and funding agendas.

But the report shows that stories of existential risk do not overshadow more immediate dangers to society. Exposure to long-term risk scenarios increases awareness of speculative harms but does not lower concern for harms in the present. Concern among the public is actually higher for harms to occur in the near future for both perceived likelihood and impact. This highlights the need for an equilibrated approach to regulation of AI, one that balances long-term vision and near-term sensitivity.

Geneva-based organisations, uniquely located at the intersection of diplomacy, ethics, and technical expertise, ought to develop inclusive foresight processes that span both perspectives. Doing so would strengthen Geneva's role as a parallel force to New York in AI governance. For this to be realised, it is crucial that these pilots direct greater institutional attention and resources towards this part of the world. Such pilots promote cross-sectoral communication, enable iterative norm-building, and allow for regulatory experimentation within a rapidly evolving field.

We recommend establishing dedicated foresight teams with:

- Technical AI experts
- Social scientists and ethicists

[59] Artificial superintelligence (ASI) is a hypothetical software-based artificial intelligence (AI) system with an intellectual scope beyond human intelligence. At the most fundamental level, this superintelligent AI has cutting-edge cognitive functions and highly developed thinking skills more advanced than any human. (IBM, 2023)

- Policy analysts and global legal specialists

These clusters can avail themselves of scenario planning, risk analysis, and visioning into the distant future to maintain governance frameworks contextually relevant and dynamic. Following the path of experimentalist governance theory (Sabel & Zeitlin, 2012), Geneva has to become a regulatory incubator. Viable mechanisms include:

- AI policy sandboxes for piloting sensitive applications
- Local ethics advisory councils engaging the UN
- Flexible multilateral task groups on emerging matters (e.g., emotion identification, autonomous choice-making)

8. Conclusion

This report, grounded in both institutional mapping and critical reflection, has attempted to situate current UN efforts in global AI governance within a wider theoretical and political landscape. Rather than prescribing a fixed blueprint, it has approached the governance conversation as iterative, contested and contingent. The study acknowledges the growing complexity of international AI governance as a product of fragmented mandates, overlapping jurisdictions and diverging national strategies. Yet it also identifies emerging signals, such as the institutionalisation of co-governance mechanisms, the mainstreaming of soft law tools and the rise of inclusive foresight exercises, that together sketch the contours of a more anticipatory and democratic governance architecture.

Empirically, this report has shown that AI governance within the UN remains largely facilitative and norm-based, with notable gaps in enforcement, coherence and equity. And yet, these very gaps become spaces for innovation. Drawing from ideas on co-governance[60], the study finds that meaningful and effective governance will likely not emerge from top-down hierarchy alone, but through distributed, adaptive networks of accountability. The future of governance may therefore lie less in global rules, and more in global capacities to deliberate, iterate and co-create across borders and sectors.

Several thematic priorities have emerged from this analysis. First, the legitimacy of governance frameworks will depend on whether they centre the voices of historically underrepresented actors, particularly from the Global South. Second, governance must evolve not just in response to known risks, but in anticipation of emergent harms. The pivot towards "governance as innovation" should not be dismissed as idealism. It is, in fact, a strategic necessity in an age where AI capabilities can change faster than the law[61]. Third, governance must move beyond performative stakeholderism. Co-governance only holds promise if

[60] Harvard Law Review, 2025
[61] Dignum, 2025

participation mechanisms are binding, well-resourced and accountable to those they claim to include.

The study recognises its own limitations, both methodological and conceptual. With a relatively small interview pool and a primary focus on multilateral institutions and processes centred around the UN system, it risks reproducing a Global North institutional bias. Future work should explore whether the principles discussed here, such as soft governance, co-regulation and institutional experimentation, are applicable across political and economic contexts, especially in underrepresented regions.

Still, the findings here reinforce that effective global governance of AI cannot be reduced to technical standard-setting or diplomatic declarations alone. It requires long-term investment in inclusive institutional processes, political will to redistribute influence, and sustained effort to translate ethical commitments into operational practices. Building an AI governance framework that is transparent, fair and responsive is not only a normative goal, it is a practical foundation for ensuring that AI systems serve broad societal interests rather than narrow ones.

9. References

Alsaawi, A. (2014). A Critical Review of Qualitative Interviews. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.2819536.

Alter KJ, Meunier S. (2009). The Politics of International Regime Complexity. Perspectives on Politics. 2009;7(1):13-24. doi:10.1017/S1537592709090033.

Abbott, K. W., & Snidal, D. (2000). Hard and soft law in international governance. *International Organization*, *54*(3), 421–456. https://doi.org/10.1162/002081800551280

Bareis, J., & Katzenbach, C. (2022). Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics. *Science, Technology, & Human Values*, *47*(5), 855–81. https://doi.org/10.1177/01622439211030007.

Bradford, A. (2012). The Brussels Effect. Northwestern University Law Review, 107(1), 1–68.

Buchanan, Allen, & Keohane, R.O. (2006). The Legitimacy of Global Governance Institutions. *Ethics & International Affairs*, *20*, 405–37.

Cath, C. 2018 Governing artificial intelligence: ethical, legal and technical opportunities and challenges. Phil.Trans. R. Soc. A 376: 20180080.http://dx.doi.org/10.1098/rsta.2018.0080

Centre for the Fourth Industrial Revolution Rwanda. (2024). Global AI Summit on Africa Rescheduled to April 2025. *Centre for the Fourth Industrial Revolution Rwanda.* https://c4ir.rw/global-ai-summit-on-africa-rescheduled-to-april-2025.

Cleeland, B., Stauffer, M. & K. Seifert. (2024). Response to the Interim Report of the UN Secretary-General's High-Level Advisory Body on Artificial Intelligence. *Simon Institute for Longterm Governance*. https://www.simoninstitute.ch/blog/post/response-to-the-interim-report-of-the-un-secretary-general%E2%80%99s-high-level-advisory-body-on-artificial-intelligence/

Daigneault, P. (2016). Process Tracing: From Metaphor to Analytic Tool. In A. Bennett and J.T. Checkel (Eds.), *Canadian Journal of Political Science*, *49*(4), 795–97. https://doi.org/10.1017/s0008423916000895.

Dignum, V. (2025). Beyond the AI Race: Why Global Governance Is The Greatest Innovation. *AI Policy Lab*. Umeå University. http://aipolicylab.se/2025/04/08/beyond-the-ai-race-why-global-governance-is-the-greatest-innovation/.

Donders, K., H. Van den Bulck, & T. Raats. (2018). The Politics of Pleasing: A Critical Analysis of Multistakeholderism in Public Service Media Policies in Flanders. *Media, Culture & Society*, *41*(3), 347–66. https://doi.org/10.1177/0163443718782004.

Dreher, A., Valentin Lang, B., Rosendorff, P. et al. (2022). Bilateral or Multilateral? International Financial Flows and the Dirty Work-Hypothesis. *Journal of Politics*, Online First. https://www.journals.uchicago.edu/doi/epdf/10.1086/718356.

European Commission. 2024a. *AI Made in Europe: The AI Continent – European Approach to AI Excellence and Trust*. Brussels: European Commission. Available at: https://digital-strategy.ec.europa.eu/en/library/ai-made-europe-ai-continent-european-approach-ai-excellence-and-trust

European Commission. 2024b. *Digital Europe Programme: Artificial Intelligence*. Brussels: European Commission. Available at: https://digital-strategy.ec.europa.eu/en/activities/digital-programme-ai

European Parliament. (2023, June 1). *Artificial Intelligence Act: EU outlines risk-based approach* [PDF]. https://www.europarl.europa.eu/pdfs/news/expert/2023/6/story/20230601STO93804/20230601STO93804_en.pdf

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., *et al.* (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Fletcher Russia and Eurasia Program. (2024, April 12). BRICS wants to shape global AI governance, too. *Fletcher Russia and Eurasia Program*. https://sites.tufts.edu/fletcherrussia/brics-wants-to-shape-global-ai-governance-too/

Funtowicz, S. O., & Ravetz, J. R. (1993). The emergence of post-normal science. In R. von Schomberg (Ed.), *Science, politics and morality: Scientific Uncertainty and Decision Making* (Vol. 17, pp. 85–123). Springer. https://doi.org/10.1007/978-94-015-8143-1_6

Global Partners Digital. (2024). *The final report of HiLAB-AI: Our analysis and thoughts*. https://www.gp-digital.org/news/the-final-report-of-hlab-ai-our-analysis-and-thoughts/

Gómez-Mera, L. (2021). International Regime Complexity. *Oxford Research Encyclopedia of International Studies*. https://oxfordre.com/internationalstudies/view/10.1093/acrefore/9780190846626.001.0001/acref

ore-9780190846626-e-648

Gruber, L. (2000). *Ruling the World*. Princeton University Press.

Harvard Law Review. (2025). Co-governance and the future of AI regulation. *Harvard Law Review*. https://harvardlawreview.org/print/vol-138/co-governance-and-the-future-of-ai-regulation/

High-Level Advisory Board on AI. (2024). *Governing AI for Humanity*. https://www.un.org/en/ai-advisory-body

Hoes, E., & Gilardi, F. (2025). Existential risk narratives about AI do not distract from its immediate harms. *Proceedings of the National Academy of Sciences*, *122*(16). https://doi.org/10.1073/pnas.2419055122

IBM. (2023). What is artificial superintelligence? https://www.ibm.com/think/topics/artificial-superintelligence

International Labour Organization. (n.d.). Artificial Intelligence and Work in the Digital Economy. https://www.ilo.org/artificial-intelligence-and-work-digital-economy

International Monetary Fund. (n.d.). AI Preparedness Index (AIPI). https://www.imf.org/external/datamapper/datasets/AIPI

International Telecommunication Union. (2025, March 7). *International standards for an AI-enabled future*. AI for Good. https://aiforgood.itu.int/international-standards-for-an-ai-enabled-future/

International Telecommunication Union. (2024). *AI and the Environment – International Standards for AI and the Environment*. ITU Publications. https://www.itu.int/dms_pub/itu-t/opb/env/T-ENV-ENV-2024-1-PDF-E.pdf

International Telecommunication Union. (2024). *AI Governance Day – From Principles to Implementation*. ITU Publications. https://s41721.pcdn.co/wp-content/uploads/2021/06/2401225_AI_Governance_Day_2024_Report-E.pdf

Jobin, A., Ienca, M., & Vayena, E. (2019). The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kaul, I., Grunberg, I., & Stern, M. A. (1999). *Global Public Goods: International*

Cooperation in the 21st Century. Oxford University Press. https://doi.org/10.1093/0195130529.001.0001

Keohane, R. O. (1984). *After Hegemony*. Princeton University Press.

Keohane, R. O., & Nye, J. S. (1977). *Power and Interdependence: World Politics in Transition*. Little Brown.

Kende, M., Fondation pour Genève, & Graduate Institute of International and Development Studies. (2020). *Internet Governance in International Geneva* [Report]. https://www.graduateinstitute.ch/sites/internet/files/2020-09/FPG_Bulletin%20Internet%20Governance-DIGITAL.pdf

Knight, W. (2024). The United Nations wants to treat AI with the same urgency as climate change. *Wired*. https://www.wired.com/story/united-nations-artificial-intelligence-report/

Lake, D. A. (2013). Theory is dead, long live theory: The end of the Great Debates and the rise of eclecticism in International Relations. *European Journal of International Relations*, *19*(3), 567–587.

Mahoney, J. (2000) Path dependence in historical sociology. *Theory and Society* 29, 507–548 (2000). https://doi.org/10.1023/A:1007113830879.

Martin, L. L., & Simmons, B. A. (2012). International Organizations and Institutions. In W. Carlsnaes, T. Risse, & B. A. Simmons (Eds.), *Handbook of International Relations* (pp. 360–378). SAGE.

Mearsheimer, J. J. (1994). The False Promise of International Institutions. *International Security*, *19*(3), 5–49. https://doi.org/10.2307/2539078

Müge, D. (2024). EU AI Sovereignty: For whom, to what end, and to whose benefit? *Journal of European Public Policy*, *31*(8), 2200–2225. https://doi.org/10.1080/13501763.2024.2318475

O'Shaughnessy, J. (2007). Book Reviews: Case Studies and Theory Development in the Social Sciences: Alexander L. George and Andrew Bennett. *Journal of Macromarketing*, *27*(3), 320–323. https://doi.org/10.1177/0276146707305480

OECD. (2024). AI governance and the role of the OECD. *OECD Events*. https://www.oecd-events.org/gpn2024/en/session/45ece119-96d0-ee11-85fa-0022488a6322/ai-governance-and-the-role-of-the-oecd

Page, Scott E. (2006), "Path Dependence", Quarterly Journal of Political Science: Vol. 1: No. 1, pp 87-115. http://dx.doi.org/10.1561/100.00000006.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2013). Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Pierson, P. (2000) 'Increasing Returns, Path Dependence, and the Study of Politics', *American Political Science Review*, 94(2), pp. 251–267. doi:10.2307/2586011.
Potjomkina, D. (2018). Multistakeholderism in the EU's trade governance. *GREMLIN: Global and Regional Multistakeholder Institutions*, 2018/01.

Radu, R. (2024). The G20 and global AI governance. *IE CGC*. https://static.ie.edu/CGC/G20_Global_AI_Governance.pdf

Radu, R., Zingales, N., & Calandro, E. (2015). Crowdsourcing ideas as an emerging form of multistakeholder participation in internet governance. *Policy & Internet*, *7*(3), 362–382. https://doi.org/10.1002/poi3.99

Raustiala, K. and Victor, D.G. (2004) 'The Regime Complex for Plant Genetic Resources', International Organization, 58(2), pp. 277–309. doi:10.1017/S0020818304582036.

Raymond, M., & DeNardis, L. (2015). Multistakeholderism: Anatomy of an inchoate global institution. *International Theory*, *7*(3), 572–616.

Sabel, C. F., & Zeitlin, J. (2012). Experimentalism in the EU: Common ground and persistent differences. *Regulation & Governance*, *6*(3), 410–426. https://doi.org/10.1111/j.1748-5991.2012.01157.x

Stone, R. W. (2011). *Controlling Institutions: International Organizations and the Global Economy*. Cambridge University Press.

Tallberg, J., Erman, E., Furendal, M., Geith, J., Klamberg, M., & Lundgren, M. (2023). The Global Governance of Artificial Intelligence: Next Steps for Empirical and Normative Research. *International Studies Review*, *25*(3), https://arxiv.org/pdf/2305.11528

The Outpost. (2024). UN Advisory Body proposes Global AI Governance Framework Amid Tech Giants' Challenges. *The Outpost.* https://theoutpost.ai/news-story/un-advisory-body-proposes-global-ai-governance-framework-am

id-tech-giants-challenges-5318/

T20 Brasil. (2024). *G20 Sao Luis Working Group Declaration on AI*. https://www.t20brasil.org/media/filemanager/20240910-Sao-Luis-Declaration-Artificial-Intelligence2.pdf

Tuli, F. (2010). The basis of distinction between qualitative and quantitative research in social science: Reflection on ontological, epistemological, and methodological perspectives. *Ethiopian Journal of Education and Sciences*, *6*, 1–12. https://www.scirp.org/reference/referencespapers?referenceid=3194148

Ulnicane, I., Eke, D.O., Knight, W., Ogoh, G. and Stahl, B.C. (2021), "Good governance as a response to discontents? Deja vu, or lessons for AI from other emerging technologies", Interdisciplinary Science Reviews, Vol. 46 Nos 1-2, pp. 71-93, doi: 10.1080/03080188.2020.1840220.

UNCTAD. (n.d.). Commission on Science and Technology for Development, 28th session. https://unctad.org/meeting/commission-science-and-technology-development-28th-session

UNESCO. (n.d.). Ethics of artificial intelligence. https://www.unesco.org/ethics-ai/en

UNESCO. (2024). 2025 Global Forum on AI and Digital Transformation for the Public Sector. https://www.unesco.org/en/articles/2025-global-forum-ai-and-digital-transformation-public-sector

United Nations. (2024). *Global Digital Compact, Main Text*. https://www.un.org/en/global-digital-compact

Veale, M. M. (2023). AI and Global Governance: Modalities, Rationales, Tensions. *Annual Review of Law and Social Science*, *19*. https://doi.org/10.1146/annurev-lawsocsci-020223-040749

Weiss, T. G., & Daws, S. (2007). *The Oxford handbook on the United Nations*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199279517.001.0001

World Economic Forum. (n.d.). AI Governance Alliance. https://initiatives.weforum.org/ai-governance-alliance/home

Xin, L., & Xin, L. (2024, September 26). China is churning out AI research but 'decoupled' from global networks, report finds. *South China Morning Post*. https://www.scmp.com/news/china/science/article/3279907/china-churning-out

Zürn, Michael. (2008). A Theory of Global Governance: Authority, Legitimacy, and Contestation (Oxford, 2018; online edn, *Oxford Academic*, 19 Apr. 2018), https://doi.org/10.1093/oso/9780198819974.001.0001.

10. Appendix

10.1 Informed Consent Form

I, [Participant's Name], have been invited to participate in a research interview by the Geneva Graduate Institute research teams and Blavatnik School of Government, University of Oxford, under an applied research project on AI Governance: UN Pathways and Implications Beyond 2025. Before agreeing to participate, I confirm that I have read, understood, and had the opportunity to ask questions regarding the following information related to my participation in this research.

Voluntary Participation

My participation in this survey is entirely voluntary, and I understand that I have the right to withdraw at any time without providing a reason. My decision to participate or withdraw will not result in any negative consequences.

Confidentiality

I understand that my responses will be treated with strict confidentiality. All information collected will be anonymized and aggregated for analysis. No personally identifiable information will be disclosed without my explicit consent.

Use of Data

I acknowledge that the data collected from this survey will be used for academic research purposes only. The findings may be reported in academic publications or presentations only, ensuring that my identity remains confidential.

Contact Information

If I have any questions or concerns regarding the survey, I can contact the research team at arp2024.aigovernance@graduateinstitute.ch.

Consent

I have read and understood the information provided in this consent form. By proceeding with the survey, I voluntarily consent to participate in the study AI Governance: UN Pathways and Implications Beyond 2025. Involvement under the conditions outlined in this form.

10.2 Questionnaire

Section 1: Background and professional experience of the respondent

1.1 Could you tell us about your work as a [Designation] and your interactions with various UN organisations?

1.2 Are there any ongoing or upcoming research initiatives you're particularly excited about?

Section 2: AI Governance Landscape

2.1 In your view, what are some of the most pressing challenges in AI governance at the global level?

2.2 How do you see the role of the UN evolving in AI governance?

> 2.2.1 There is a shift in diplomatic terms in recent times: in this environment of diplomatic mistrust, do we think the UN can emerge together to find solutions and pathways for AI in global governance?

2.3 What governance models (state-led, multistakeholder, regional) are most effective? Where do we see the challenges effectively addressed?

Section 3: Institutional Developments

3.1 How do international frameworks (e.g., UN AI Office, EU AI Act, G20 initiatives) influence AI governance?

3.2 What gaps exist in AI governance structures, and how can they be addressed?

3.3 How do private actors (big tech firms, civil society) interact with governance mechanisms?

3.4 What is being done at the organisational level to incorporate and/or understand AI?

3.5 What is unique about understanding the scale at which AI may impact and become useful, when at the scale of an international organisation?

Section 4: Future of AI Governance Beyond 2025

4.1 What trends do you foresee in AI governance in the next 5–10 years, with a special focus on the UN?

4.2 Do you believe International Geneva has a unique role to play in the global governance of AI?

Section 5: Closing & Additional Insights

5.1 What is one recommendation you would give for improving AI governance?