

Civil Society Read Out Report

1st United Nations Global Dialogue on AI Governance

Global Digital Rights Coalition

Global Digital Justice Forum

Global South Alliance

MAP-AI

Civil society members of the Gender in Digital Coalition

Introduction

The inaugural UN Global Dialogue on AI Governance (AI Dialogue) marked a critical convergence point, reinforcing that Artificial Intelligence (AI) has the potential to reshape economies, societies, and daily life. However, amidst the hype of AI as a game-changer for development, the Dialogue also brought much-needed attention to the risks, inequalities, and governance challenges associated with this emerging paradigm.

The **power asymmetries** in AI capabilities and the uneven distribution of AI benefits and harms between the Global North and Global South are pronounced, necessitating open conversations about equity in AI governance and international cooperation. This makes a coordinated international and democratic governance approach to AI a critical consideration for the future.

Established by the UN General Assembly, the AI Dialogue created a platform for governments and stakeholders to discuss international cooperation, share best practices and lessons learned, and facilitate open, transparent, and inclusive discussions on AI governance. This first Dialogue was an important start, but some gaps need to be overcome before the second Dialogue.

Several critical voices, including those of women and Indigenous Peoples, were insufficiently represented throughout the Dialogue. Discussions on human rights were also uneven across thematic areas, and opportunities for direct engagement between governments and other stakeholders remained limited. As preparations begin for the second Dialogue, addressing these gaps will be essential to advancing the UN's vision of AI governance as "inclusive, human-centric, and development-oriented." This requires ensuring the **meaningful participation of historically underrepresented groups**, strengthening multistakeholder dialogue, and grounding all discussions on AI governance in international human rights standards.

We commend the Co-Chairs for their careful stewardship of this first iteration and their commitment to bringing diverse stakeholders into the Dialogue. Drawing on these collective experiences, **this report offers both substantive and procedural dimensions of AI governance, which deserve recognition in the Co-chairs' summary** based on the

Civil Society Report was developed collectively by members of civil society organizations from five coalitions - the Global Digital Rights Coalition, the Global Digital Justice Forum, the Global South Alliance, MAP-AI and the Civil society members of the Gender in Digital Coalition - which have engaged consistently throughout the process, with many of their members participating directly in the inaugural Dialogue.¹

collective read-out of the discussions heard during the first Dialogue and the relevant gaps also identified. Acknowledging the Dialogue as a welcome beginning to a broader process of international coordination, this report offers recommendations to strengthen its second iteration and support a more inclusive, rights-based, and effective global AI governance framework.

Cluster 1: AI opportunities and implications: social, economic, cultural, ethical, linguistic, and technical dimensions

Member States should recognize that AI systems are developed within existing social, economic, and historical contexts and may reinforce global and local inequalities, including those affecting developing countries. AI governance should therefore promote inclusive and equitable AI development by **ensuring that diverse knowledge systems**, languages, histories, and lived experiences, including those of women, Indigenous Peoples, and other historically marginalized communities, **are meaningfully integrated throughout the design, development, and deployment of AI systems.**

Robust **data governance** is foundational to harnessing AI opportunities. The governance of the non-personal data commons requires a societal approach that includes **safeguards for collective privacy and the rights of communities to steward the use and re-use of their data** resources in innovation ecosystems, together with strong personal data protection rights. This recognition should be aligned with the work currently underway by the Working Group on Data Governance at All Levels, at the UN Commission on Science, Technology and Development.

Member States should promote regulatory and oversight measures to address **manipulative and harmful AI-enabled design practices.** Such measures should include appropriate transparency and reporting obligations, including access to relevant disaggregated data and mechanisms for reporting identified risks. Particular attention should be given to **AI systems used in education and to their potential impacts on children**, taking into account their specific rights and vulnerabilities. Efforts to strengthen AI literacy should be complemented by measures to prevent harmful and manipulative design practices and to promote the safe and rights-respecting development and use of AI systems.

Impacts related to extraction of critical minerals, large-scale energy consumption, and exploitative labour conditions are borne disproportionately by developing countries, while the economic and strategic benefits of AI are concentrated elsewhere. **The environmental, labour and social impacts of AI should be treated as core governance issues**, and the **precautionary principle** adopted, with mandatory environmental and human rights risk and impact assessments across the value-chain of AI systems, with internationally comparable standards of measurement, public disclosure requirements and the participation of affected communities. Recognizing and explicitly naming the environmental footprint of AI should be a priority for the Second UN Global Dialogue. AI's environmental footprints are structural, not incidental, and they often impact workers, communities and territories far from the main beneficiaries of AI deployment. Addressing the environmental impacts of AI implies serious commitments from multiple stakeholders and a need to discuss the conditions under which

Civil Society Report was developed collectively by members of civil society organizations from five coalitions - the Global Digital Rights Coalition, the Global Digital Justice Forum, the Global South Alliance, MAP-AI and the Civil society members of the Gender in Digital Coalition - which have engaged consistently throughout the process, with many of their members participating directly in the inaugural Dialogue.²

AI is developed and deployed. We need to **strengthen coordination with other international processes and institutions** (such as the IPCC) to ensure the subject is not addressed through siloed AI-specific mechanisms.

Cluster 2: bridging AI divides: capacity-building, access and digital foundations

Member States should commit to capacity development in a broad sense, encompassing **technical, institutional, and human rights capacities**, while ensuring the meaningful participation of all stakeholders, including civil society. The thematic co-chairs' summary should **reaffirm the mandate of the Office of the High Commissioner for Human Rights (OHCHR)** to continue its critical work of anchoring the Global Digital Compact (GDC), the AI Dialogue, and the WSIS+20 outcomes in international human rights standards, and to strengthen its capacity to provide advisory services on human rights in the digital space, as recognized in paragraph 24 of the GDC.

Member States should recognize that AI divides extend beyond gaps in connectivity to include persistent documentation and data divides. The 2.2 billion people who remain offline are not only excluded from the benefits of AI but are also largely absent from the data on which AI systems are developed, with women disproportionately affected by this exclusion. The evidence does not indicate that this gap will be solved solely with connectivity. As AI systems are increasingly trained on the outputs of previous models, existing gaps in representation risk becoming further entrenched over time, limiting the ability of future AI systems to reflect the experiences and knowledge of under-documented populations.

Therefore, Member States should recognize provenanced and consent-based data relating to under-documented populations as a core digital foundation, alongside connectivity infrastructure, and treat its development as a time-critical priority rather than an incremental process. Capacity-building efforts should include support for countries and communities not only to access and use AI systems, but also **to generate, steward, and govern the data resources necessary for the development of AI systems** that reflect their own realities, knowledge, and priorities. It should also prioritize institutional resilience by strengthening the capabilities of public administrations, regulators, and oversight bodies to have responsible AI systems throughout all their development and deployment cycle.

Member States should recognize that the inclusiveness and equitable benefits of AI systems depend on the quality, diversity, and **representativeness of the data** on which they are developed. Gaps or imbalances in training data may result in the systematic underrepresentation of certain populations, languages, and domains of knowledge, thereby limiting the ability of AI systems to serve all communities fairly and effectively. To address these challenges, Member States should promote the development and curation of diverse, high-quality, and representative datasets, ensuring that women, underrepresented linguistic communities, and other historically marginalized groups are adequately reflected throughout the AI development lifecycle.

They should also support **contextually appropriate AI models and applications that respond to local needs and priorities**, recognizing that trust in AI is strengthened when

Civil Society Report was developed collectively by members of civil society organizations from five coalitions - the Global Digital Rights Coalition, the Global Digital Justice Forum, the Global South Alliance, MAP-AI and the Civil society members of the Gender in Digital Coalition - which have engaged consistently throughout the process, with many of their members participating directly in the inaugural Dialogue.³

systems are developed and adapted with meaningful local participation and control. The promotion of **open-source software, open data, and open AI models** can contribute to innovation, transparency, and capacity-building, where embedded in pro-commons governance.

Cluster 3: Safe, secure and trustworthy AI: interoperability and compatibility of approaches

On Transparency and Measurable Benchmarks, we trust that the Dialogue can evolve towards discussing overall transparency obligations that extend beyond model outputs to training corpora. Trustworthiness must be measured comparably across jurisdictions using standardized benchmarks grounded in international human rights law rather than vendor-led metrics. High-stakes deployments in clinical, judicial, and social protection settings require mandatory demographic performance thresholds.

Member States should promote secure-by-design and secure-by-default approaches across the AI lifecycle, and commit towards making Global governance converge on shared human rights frameworks—such as the Council of Europe Framework Convention on AI, international agreements like the Hiroshima process and national/regional legislation—for **transparency, risk assessment, interoperability and incident reporting**. Content provenance and Authenticity infrastructures (e.g. C2PA) must prioritize privacy over surveillance, with independent third-party verification. We also recommend that the Dialogue builds up on the warnings from the Independent International Scientific Panel on AI about how Artificial Intelligence can erode the shared reality, and mandate and resource a multistakeholder working group on content authenticity and information integrity, on the road to New York, with specific deliverables and civil society at its centre.

With regards to **Socio-Technical Safety and Accountability**, we urge Member States to move governance beyond technical safety to encompass social considerations, including social audits, gender impact assessments, and localized evaluations. We call for the Dialogue to set up the basis for binding **public accountability standards** across the entire AI value chain, ensuring legal requirements and technical controls are clearly linked.

Current "gold standard" datasets **largely exclude the Global South due to centralized infrastructures**. In order to **foster Inclusivity and global equity**, the best practices and norms approached by the Dialogue must be developed with active participation from developing countries to account for diverse contexts. Research ecosystems must be funded to generate local evidence, drawing on expertise from informal economies, multilingual societies, and infrastructure-constrained settings.

We also **urge Member States to develop Gender-Responsive Governance and approaches**. Gender must be treated as a structural dimension of safety and no longer be addressed as an add-on risk category. Technology-facilitated gender-based violence (TFGBV), including non-consensual intimate deepfakes and gendered disinformation, are safety failures in their own right, and result in heightened safety and security risks for women and gender-diverse people - therefore deserving zero tolerance. Data extraction without consent and biased datasets that repeat old colonial patterns in AI development and Civil Society Report was developed collectively by members of civil society organizations from five coalitions - the Global Digital Rights Coalition, the Global Digital Justice Forum, the Global South Alliance, MAP-AI and the Civil society members of the Gender in Digital Coalition - which have engaged consistently throughout the process, with many of their members participating directly in the inaugural Dialogue.⁴

deployment disproportionately impact women and gender-diverse people, who bear the most harm. This requires mandatory gender impact assessments prior to deployment, clear "red lines" for rights-violating uses, and direct funding for affected groups to participate in oversight.

Lastly, we call on Member States to acknowledge that AI harms have sectoral impacts. AI is reshaping welfare, agriculture, and healthcare delivery across the Global South, faster than institutions can monitor. Further, the impacts in these sectors are also greatly gendered. We need to build norms that speak to the specific trust, safety and security considerations in each of these sectors to enable the translation of principles into practice, and their operationalisation on the ground.

For all of the points listed above, we urge the Chairs' Summary to enshrine a '**precautionary approach**' as the baseline for AI development and deployment.

Cluster 4: Respecting, protecting and promoting human rights: transparency, accountability and human oversight

With regards to the **development of stronger accountability for human rights governance**, international human rights law still serves as the essential framework for consensus. Therefore, we request that Member States ground their AI national policies and regulations in these existing legal obligations, treating governance as a fundamental human rights issue rather than solely a developmental one, and must demonstrate accountability through concrete strategies, such as the Universal Periodic Review. Added to that, and looking forward to the next Dialogue, we welcome stronger commitments from Member States on centering AI Governance in human rights, as well as a broader cognisance of governments' and private sector's respective obligations and responsibilities under international human rights law. Further, we would **welcome that the Chair's Summary explicitly acknowledges that international law, including international human rights law, already applies to the way in which AI systems are designed, commercialised, deployed, used, sold and governed.**

As stated in the concept note for this cluster, **accountability after harm occurs is insufficient**. This means that ex ante safeguards are indispensable. The Chair's Summary must also point towards the direction of a stronger adoption of human rights due diligence and human rights impact assessments by States and companies to identify and address harms before they occur, applying these practical tools across the AI value chain and embedding them in institutional decision-making. Additionally, as observed by the Independent International Scientific Panel on AI, the adequacy of existing liability frameworks to hold transnational corporations accountable for harms stemming from their AI products and services is questionable. In this sense, **we suggest that the Chairs' Summary must call for global cooperation to urgently address the challenges AI systems pose to prevailing liability frameworks, such as (a) tracing harms to specific design choices (b) transparency obligations in relation to making risk management processes public and (c) appropriate distribution of accountabilities in multi-actor AI value chains.**

With regards to AI redlines, we suggest that the co-chair's summary must call for clear prohibitions or moratoria on AI systems and uses that pose unacceptable risks to human rights. This is explicitly stated in the 2024 UN General Assembly Resolution, *Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development*. The summary should restate that international human rights law is the basis for determining which uses or design choices cannot be justified as legally grounded, necessary or proportionate, and which are therefore fundamentally unacceptable. Added to that, the summary should state specific cases where there is clear or emerging consensus that moratorium or prohibition is an appropriate legal response. This list should be illustrative, not exhaustive, and should include the use of biometric mass surveillance (A/HRC/RES/58/23; A/HRC/RES/54/2), social scoring systems that lead to detrimental or adverse treatment, manipulative or deceptive design practices, AI systems that enable lethal force, AI systems that target decisions without meaningful human control, and AI systems that meaningfully enhance chemical, biological, radiological, and nuclear (CBRN) threats. The purpose of this proposal is for the Dialogue to acknowledge recommendations prescribed by international law, and part of the obligations of existing justice systems and accountability mechanisms. **Implementing red lines is a critical component of ensuring the rule of law and good governance: clarifying and enforcing what is permitted, and what must never be allowed.**

In order to ensure that the trajectory of AI development and deployment actively advances economic, social, and cultural rights—particularly within the Global South—it is imperative to **establish defined baseline norms across the entire AI value chain**. These standards must function as non-negotiable requirements that address critical areas including labor practices, ecological sustainability, public accountability, and equitable value distribution, ensuring that justice and equity remain central to the lifecycle of AI production.

Frequently, decisions are made via AI systems, yet those adversely affected lack a clear avenue to challenge them. This issue can be intensified as agentic AI becomes ubiquitous. Additionally, laws and regulations must mandate that an accountable actor is integrated into decision-making workflows. Such measures should also establish clear and reasonable timelines to ensure that affected individuals can effectively access remedies. In this sense, we reinforce that **enabling human oversight remains paramount to ensure accountability and mitigate risk** and we also **urge the Dialogue to evolve on discussions on enabling the right to remedy in light of AI harms**. Humans must maintain meaningful oversight and input regarding AI systems, with an accountable actor integrated into decision-making workflows to guarantee both ad hoc and post hoc accountability, including the ability to access remedies. This necessity becomes increasingly urgent as AI systems grow more agentic—particularly with the rise of AI assistants capable of operating across diverse applications, devices, and operating systems—as the inherent risk of harm significantly escalates.

With respect to the discussion on differentiated and disproportionate AI impacts across the value chain and communities, we urge the co-chair's summary to also address the differentiated and disproportionate impacts of AI based on gender, age, disability, climate change and other protected characteristics, including those in possession of multiple or intersecting characteristics. AI systems can perpetuate and exacerbate discrimination by

Civil Society Report was developed collectively by members of civil society organizations from five coalitions - the Global Digital Rights Coalition, the Global Digital Justice Forum, the Global South Alliance, MAP-AI and the Civil society members of the Gender in Digital Coalition - which have engaged consistently throughout the process, with many of their members participating directly in the inaugural Dialogue.⁶

design, as evidenced by tools that yield lower diagnostic accuracy or skewed judicial outcomes. AI governance should strengthen their protection through dedicated risk mitigation, capacity-building, and access to effective remedies. Beyond systemic bias, the proliferation of large-scale synthetic media threatens fundamental human rights, including agency and dignity, while automation risks disrupting critical sectors like journalism—a cornerstone of democracy. To address these disproportionate impacts, governance must extend beyond identifying harms to actively fostering public digital infrastructure standards and innovation policies that prioritize equity and public interest over existing technical biases.

Procedural recommendations

Building on the substantive discussions of the inaugural session, the following procedural and structural recommendations are essential to enhance the effectiveness of the upcoming AI Dialogue. These proposals focus on strengthening the UN ecosystem's engagement with civil society, improving the inclusivity of dialogue processes, and ensuring that future meetings are grounded in transparency and equitable participation.

UN Ecosystem & other relevant events

1. The global IGF is at the midpoint of the two Dialogues. We strongly encourage the Co-Chairs of the second Global Dialogue on AI Governance and their teams to attend the IGF, making use of this valuable multistakeholder mechanism and space to consult with different communities and refine the priorities of the first Dialogue and to develop the agenda for the second Dialogue.
2. The Co-Chairs and the Secretariat should emphasise the need to adequately resource the OHCHR and UN human rights mechanisms, including universal periodic reviews, special procedures, treaty bodies and the Human Rights Council, to continue their work to monitor, document and respond to AI-related harms. Specifically, the OHCHR requires adequate resourcing to continue its critical work of anchoring the GDC, the Dialogue and the WSIS+20 outcomes in human rights, and to enable it to fulfil its function of providing an advisory service on human rights in the digital space (Para 24, GDC).
3. The Secretariat should ensure logistical support for critical voices to obtain the relevant visas to participate in this process, especially for Global South actors.
4. The Co-Chairs and the Secretariat should continue the established practice of holding consultations online and at regional levels to allow for as many voices as possible to be heard in the run-up to the second Dialogue.
5. Lastly, we would welcome it if the Dialogue attempted to establish stronger links with other Artificial Intelligence-related events such as the AI Summit series. This could ensure that the discussions and developments of the Dialogue are taken forward and centered in other spaces.

Format of the Dialogue

6. In the allocation of plenary speaking roles to representatives from developing countries, it is important to ensure that civil society organizations are adequately represented. The Dialogue format must follow the UNESCO Recommendation on

Civil Society Report was developed collectively by members of civil society organizations from five coalitions - the Global Digital Rights Coalition, the Global Digital Justice Forum, the Global South Alliance, MAP-AI and the Civil society members of the Gender in Digital Coalition - which have engaged consistently throughout the process, with many of their members participating directly in the inaugural Dialogue.⁷

Open Science in terms of ensuring the meaningful engagement of civil society actors. Special attention must be given to ensuring the effective participation of organizations representing marginalized people and groups, such as ethnic groups, Indigenous Peoples and other small-scale food providers, workers, women and girls, LGBTIQ+, and Youth.

7. We ask the Secretariat to ensure all speakers are explicitly informed and confirmed once they are placed on the speaker list, ensuring that any changes to panel modalities are made at least 2 weeks before the event date.
8. Foster a genuine and facilitative Co-Chair role that ensures inclusive, balanced, and active engagement across all stakeholder groups during the sessions. Communicate more clearly about the role of the co-chairs and cluster's co-facilitators in order to set the correct expectations towards the role and improve transparency.
9. We also recommend establishing ways for those unable to attend the Dialogue or those attending but unable to take the floor, to offer their input in a meaningful and inclusive way into the Dialogue.

We extend our sincere thanks to the co-facilitators for their openness and the opportunity to engage in these vital discussions. Our coalitions look forward to discussing these suggestions further and are eager to see the next iteration of the Dialogue unfold.