SECTION 6

Recommendations

To Governments

Rights-Based AI Governance

An overarching recommendation is for states to adopt a rights-based AI governance framework, ensuring that human rights principles are embedded throughout the development and use of AI systems.[33] At a high level, this includes enacting complementary laws, regulations and institutions that enable the protection, respect, and remedy of rights throughout the AI value chain. This may include mandating risk-based human rights assessments (supported by meaningful external stakeholder engagement) and related preventative and mitigation measures across the AI value chain, mandating state- and non-state based remedy mechanisms, and actively participating in multilateral and multi-stakeholder efforts to shape global AI governance and advocate for the ongoing protection of human rights.[34] The specific recommendations below explore thematic areas and risks highlighted earlier in this Brief.

Restriction of Information

Government mandates related to the inputs and outputs of AI models can both impact human rights.[35] Conditions on inputs that are designed to restrict model outputs are likely to have disproportionate and unintended consequences. While limitations on outputs can be more narrowly tailored, they should focus on content that is illegal. Given the challenges that exist with "re-training" models, governments should be especially careful to design legal and regulatory frameworks so that they avoid creating impacts on rights that will be difficult to remedy retrospectively.

[33] Several GNI members have published thought leadership and have advocated for rights-based AI governance, including Global Partners Digital (GPD) and Article 19, as well as non-members such as Chatham House and Access Now, and multilateral organisations such as B-Tech

[34] Such as the Global Partnership for AI, the Council of Europe Framework Convention on AI, the Global Digital Compact, the World Summit for Information Systems, the G20 AI Dialogues

[35] The potential negative rights impacts of overbroad use of AI in content moderation has been documented extensively, including by ECNL

In addition, governments should be cautious about shifting legal liability for AI-generated content to intermediaries, as this may incentivize over-removal and over-censorship. In line with the legality requirement articulated above, government interventions must clearly define prohibited content and conduct, and allow determinations of responsibility for illegal content to be adjudicated by independent judicial bodies in conformity with due process norms.

Surveillance

Governments using AI technologies to acquire and/or analyze personal data (including biometric data) must ensure that these activities are properly authorized under public and clear legal frameworks, and that appropriate transparency, independent oversight, and remedy/accountability mechanisms exist to guard against misuse. These same safeguards are necessary when governments acquire data from companies that manage AI tools or services, whether through legal requests or via commercial procurement of data. In addition, governments are encouraged to:

- Allow users to interact with AI products or services in ways that protect their identity, including through the use of encryption;
- Avoid requirements that compel or enable tracking, tracing, or proactive monitoring of user activity by companies;
- Minimizing data collection, processing, storage, and retention requirements;[36] and
- Implement rights-protecting data protection laws to ensure users have appropriate awareness and control of their data, as well as access to remedy where their data is misused.
- Refrain from accessing user data, whether directly or indirectly through demands to third parties, without meeting appropriate safeguards.

Export Controls

While international human rights law permits restrictions on freedom of expression and privacy on national security grounds, as noted previously, export controls on critical infrastructure, as well as models themselves, can have unintended and/or disproportionate impacts. It is therefore recommended that any export controls be as targeted as possible and that governments applying such controls:

- Incorporate human rights into export controls policy, including establishing processes to routinely engage with civil society on export controls,

[36] For example, requiring global AI providers to host application or user data locally. Please see page 28 of GNI's Content Moderation Policy Brief

- Strengthen export controls on technologies with an unequivocal dual use to nations with documented human rights violations, including AI-assisted surveillance and censorship technologies,[37]
- Review and implement processes and technologies to more precisely control use cases that meet specific security objectives, instead of blanket export controls on entire nations and their rightsholders,[38] and
- Continue scientific exchange and collaboration on AI technologies to promote cross-jurisdiction understandings and collaborations around risks and empower rights-respecting uses.

Sovereign AI[39]

Governments investing in sovereign AI should ensure that such initiatives are grounded in rights-based governance frameworks, such as those mentioned earlier in this Brief. Specifically, governments should consider evaluating actions that restrict access to information, limit expression, and violate user privacy in line with the three-part test.

Recommendations related to sovereign AI initiatives include:

- Ensure AI models use inclusive datasets that represent minority languages and inputs from marginalized communities;
- Facilitate equitable and rights-respecting access to AI, through open APIs, affordable tools, and AI literacy programs, to narrow digital divides, promote equitable scientific advancement, and empower vulnerable groups;[40] and
- Prioritize opportunities for economic inclusion in AI investments, particularly in underdeveloped or underserved regions.[41]

[37] A position advocated for by Freedom House and Human Rights Watch, among others

[38] As noted by various academics, think tanks, and industry representatives, blanket bans may expedite the development of indigenous capacity in targeted jurisdictions, thereby negating the impacts of export bans, while provoking retaliatory measures that may impact rightsholders in the source nation.

[39] While there are various definitions of "sovereign AI," we refer here to NVidia's definition: "Sovereign AI refers to a nation's capabilities to produce artificial intelligence using its own infrastructure, data, workforce and business networks."

[40] For example, multiple access initiatives such as Canada's AI Compute Access Fund and Singapore's GenAI Sandbox for SMEs.

[41] For example, China's Eastern Data, Western Compute (EDWC) initiative illustrates how infrastructure and AI capabilities can be strategically directed to reduce regional disparities

Public Sector Use Cases

Governments deploying AI in the public sector should ensure that such initiatives are grounded in rights-based governance frameworks, such as those mentioned earlier in this Brief.[42] Specific incremental recommendations include:

- Prohibition of public sector use cases with a strong likelihood of significant and/or irremediable rights impacts;[43]
- Develop mitigations for use cases with lower risks of impact on human rights;[44]
- Consider this guidance in the context of public sector service delivery that involves AI-enabled automated decision making;
- Maintaining a public inventory of AI use cases across government agencies;[45] and
- Implementing remedy mechanisms related to public sector uses of AI.[46]
- Mandatory and meaningful engagement of external stakeholders, especially civil society and affected communities

[42] See this report from the Ada Lovelace Institute on public sector AI procurement, which recommends clearer, consolidated guidance, defined terminologies, stronger governance, built-in ethical and transparency safeguards, public engagement, and support for local government capacity and accountability.

[43] For instance, as in the EU AI Act's Article 5.

[44] For instance, the EU AI Act requires the following types of mitigations: human rights risk assessments (Article 27), third-party evaluations (Article 43), transparency (Article 13), and continuous monitoring (Article 61)

[45] For example: https://github.com/ombegov/2024–Federal-AI-Use-Case-Inventory

[46] This includes executive mechanisms such as the UK's Investigatory Powers Tribunal or the US DOJ complaint mechanism, or judiciary mechanisms

To Civil Society

Civil society has long played a crucial role in safeguarding human rights in the technology sector, and this role is even more vital in the context of AI. Civil society should continue to advocate for rights-based AI governance frameworks that embed international human rights law into both national and international AI regulations and their implementation. This includes active participation in global policy forums to ensure that human rights are central to emerging AI governance structures.

Civil society actors also engage with companies to promote rights-respecting internal AI governance frameworks. This includes: providing input into and feedback on corporate policies and practices to ensure they align with the UN Guiding Principles on Business and Human Rights; engaging with companies on their ongoing human rights due diligence efforts; conducting and publishing research on the impacts of AI-enabled products and services; and participating in accessible remedy mechanisms across the entire AI lifecycle.

In the public sector, civil society should push for public consultation, robust accountability mechanisms, and independent oversight, especially for public sector use cases deployed in high-risk contexts such as recruitment, law enforcement, benefit allocation/social services, border control, and military uses.

Civil society plays an essential role in ongoing engagement with key rightsholders—such as affected communities, journalists, and legal professionals—regarding the human rights implications of AI systems. This close involvement uniquely positions civil society to conduct research, build a credible evidence base, and document, analyze, and elevate the unintended rights impacts arising from AI deployments across both public and private sectors.

Efforts should be made to ensure that civil society and representatives who study, represent, and/or advocate for vulnerable communities or represent marginalized populations are supported (including financial resources) and listened to (including meaningfully incorporating their feedback into product development or use, and policymaking).

To Companies

All companies, including companies in the AI value chain, have a responsibility to respect their users' rights, including the rights to freedom of expression and privacy, and to avoid discriminatory impacts on marginalized groups who are disproportionately impacted by AI systems. They should comply with applicable laws while respecting internationally recognized human rights wherever they operate. In cases where national laws, regulations, or policies fall short of international standards, technology companies are expected to avoid, mitigate, or address the negative impacts of government demands and seek ways to uphold these human rights principles to the greatest extent possible. Furthermore, companies should be able to demonstrate their efforts in fulfilling these responsibilities in line with the UNGPs and the OECD Guidelines for Multinational Enterprises.

To support these efforts, the Global Network Initiative (GNI) Principles on Freedom of Expression and Privacy, along with its more detailed Implementation Guidelines, provide a comprehensive framework offering guidance to the tech sector and other stakeholders in respecting and advancing human rights worldwide. GNI creates space for companies to demonstrate and receive feedback on these efforts, supports cross-industry and multistakeholder learning, supports rights-focused advocacy, and facilitates meaningful stakeholder engagement. The Annex further unpacks how the GNI framework can apply in relation to corporate conduct and decision making related to AI.

Companies should proactively advocate for laws and regulations that align with international human rights norms, refrain from advocating for laws and regulations that are inconsistent with those norms, and engage in proactive joint public policy advocacy with civil society, multilateral organizations, industry bodies, or multistakeholder initiatives in relevant jurisdictions. Companies should conduct ongoing human rights due diligence (HRDD), including meaningful stakeholder engagement, to identify and then take action to avoid or mitigate human rights impacts related to their development and deployment of AI-related technologies, tools, and features. In addition, companies may benefit from conducting detailed human rights impact assessments (HRIA) in certain circumstances, including when developing new products or entering or exiting certain jurisdictions.[47]

As part of this HRDD, companies should understand their potential exposure to diverse forms of government demands, interventions, pressures, and restrictions. When faced with

[47] See, e.g., AI-related HRIAs conducted by Microsoft, Intel and Google

such government action, companies should assess their legality, legitimacy, necessity, and proportionality in line with international human rights law, in order to determine how best to respond. Where government interventions do not meet these criteria, companies should consider how best to push back or otherwise limit compliance, including by engaging in dialogue and advocacy through relevant multilateral or multistakeholder initiatives.

Companies are recommended to maintain transparency towards impacted users and the public in their respective local languages, including by publishing the results of HRIAs, disclosing government interventions where feasible, engaging with rightsholder representatives, and notifying impacted users in affected jurisdictions where permitted by local laws. Additionally, companies should establish grievance mechanisms in line with best practices (UNGP Articles 29 and 31) to allow users to report impacts on them or the rightsholders they represent.

Examples of AI-related, pre- and post-compliance prevention and mitigation measures could include, but are not limited to:

- Conduct impact assessments on AI functionalities (especially high risk use cases such as AI-based facial recognition) in anticipation of and in response to government use and interventions,
- Funding and otherwise supporting independent research and civil society monitoring of the human rights impacts of AI systems in affected regions, especially in contexts where government oversight is weak or absent, and
- Collaborating with governments and/or civil society to provide AI literacy programs or digital security training, especially for vulnerable populations, while supporting the development and access to rights-respecting local AI models in the same regions.