

ARTICLE 19 Submission to the High-Level Advisory Body on Artificial Intelligence (AI) on key issues of Global AI Governance

In this submission, ARTICLE 19 responds to the call for papers on key issues on global AI governance, in advance of the first meeting of the Multistakeholder Advisory Body on AI (Advisory Body). We encourage the Advisory Body to consider a human rights-centered framework—particularly as it relates to the right of freedom of expression—as one of its thematic pillars. We further offer a suggested ten-point focus plan for this pillar, based on ARTICLE 19’s analysis of existing guidance, best practices, and risk areas on AI and freedom of expression from a variety of stakeholders.

In the context of escalating and overwhelming application of AI across a variety of fields from government to the private sector, ARTICLE 19, as part of its ongoing advocacy about the impact of AI on freedom of expression,¹ welcomes the efforts to include a diversity of important themes to be addressed through global governance.

As a number of human rights bodies have recognized already, AI carries both benefits and risks for the enjoyment of fundamental rights and therefore implicates States’ obligations under human rights law.² The development and application of AI by both the private and public sectors also heavily implicates the rights to freedom of opinion and expression, both directly and indirectly. These issues are particularly critical given recent calls³ by a coalition of UN special procedures and other experts for urgent action on the “alarming” use of AI to undermine journalists and human rights defenders, as well as its use in the mass production of synthetic content to spread disinformation or promote incitement to hatred, discrimination or violence.

ARTICLE 19 therefore proposes that one of the thematic pillars for the work of the Advisory Body should be a human rights-centered framework for AI. This—as we provide in the ten-point framework—would serve as a reference for States, private actors, and civil society as they engage with a variety of timely topics that impact freedom of expression. In analyzing these issues, ARTICLE 19 echoes the calls of the High Commissioner for Human Rights to examine “AI’s entire lifecycle”, evaluating how technical standards may contribute to or undermine human rights.⁴ The ten-point framework should address in particular the following issues:

- 1. AI and content-based interferences with freedom of expression.** ARTICLE 19 observes that applications of machine learning algorithms and limitations on their use have the potential to limit expressive activity online. Where this occurs, it may constitute interference with freedom of expression that must be analyzed pursuant to Article 19(3) of the International Covenant on Civil and Political Rights. Such content-based interferences include, but are not limited to, content moderation on social media platforms as they increasingly utilize AI for automated multimedia content analysis, moderation, or blocking.⁵ AI is often poor at detecting nuance—especially for content deemed to be hate speech or ‘disinformation’, and thus may be excessive in its removal of legitimate expressive activity.⁶ Already vague notions of hate speech⁷ may be exacerbated by AI systems that struggle with contextual nuance. At the same time, AI may contribute to the problem by amplifying inherently unfair, discriminatory, or biased trends in training datasets.⁸

Therefore, the Advisory Body could reiterate that any measures to address these problems must comply with international standards,⁹ including the guidance of the UN Special Procedures and Human Rights Council.¹⁰

2. **AI and surveillance.** ARTICLE 19 observes how AI and biometrics are used for facial recognition in public spaces and municipal infrastructure, to develop profiles on individuals, monitor movements and relationships, and even predict criminality.¹¹ ARTICLE 19 recommends that the Advisory Body provide a framework on the collection, use, and sharing of biometrics consistent with international standards; i.e. that surveillance only be conducted on a targeted basis on grounds of reasonable suspicion, and personal data protections must be in place.¹²
3. **AI and safety of human rights defenders, journalists, and activists.** AI impacts multiple groups—who are often subject to intimidation, harassment, and threats of violence in a transforming media environment¹³—via means such as bot network harassment,¹⁴ doxing, the use of generative AI to create materials for blackmail,¹⁵ and AI-based surveillance (see above). AI can also be utilized to ‘de-anonymize’ individuals, undermining journalist-source relationships. ARTICLE 19 suggests the Advisory Body provide best-practices for oversight and mechanisms for remedies to protect these groups.
4. **AI and media freedom.** AI impacts the work of newsrooms in novel ways. These include automated news creation, promoting broader dissemination (such as quickly translating stories for new audiences), or curating access to stories based on reader patterns.¹⁶ These should not be used as a pretext for media regulation, and as such ARTICLE 19 suggests the Advisory Body monitor any attempts of governments to regulate the media. ARTICLE 19 recommends media self-regulation on how it deploys AI in order to promote a pluralistic media environment.
5. **AI industry best practices.** ARTICLE 19 suggests the Advisory Body collect and share existing ethical codes and various industry standards on artificial intelligence. These have important implications for the protection and promotion of freedom of expression by directly impacting the manner in which the private sector develops and deploys AI. Such efforts are already underway through the private sector and civil society.¹⁷
6. **Respect for human rights safeguards.** As States increasingly adopt long-term strategic plans relating to their implementation of AI,¹⁸ ARTICLE 19 urges that these plans reference existing rights obligations and safeguards. The High Commissioner on Human Rights has stressed the urgent need to pause the use and sale of AI negatively impacting human rights until adequate safeguards are in place.¹⁹ We suggest the Advisory Body provide best-practice safeguards for States to include in their strategic plans.
7. **AI and transparency.** The opacity of machine learning algorithms presents particular challenges for individuals, regulators, civil society, and even designers of systems, as it is often unclear when and how systems are utilized, and therefore difficult to audit their human rights implications.²⁰ ARTICLE 19 suggests that the Advisory Body recommend standards for developers of “high risk” AI systems,²¹ which are particularly prone to impact human rights, to provide meaningful public

and civil society access to those activities, including but not limited to requirements for public registration either nationally or internationally.

8. **Impact assessments.** ARTICLE 19 suggests that the Advisory Body provide that users of high-risk AI systems have an obligation to conduct and publish human rights impact assessments *prior* to their deployment. These proposals also further the aim of transparency, and have been echoed at the national level.²²
9. **Accountability.** ARTICLE 19 suggests that the Advisory Body develop and recommend mechanisms to empower individuals whose rights are violated, including a right to lodge complaints, a right of representation, and rights to effective remedies.
10. **Prohibition of dangerous AI.** At the broader level, there must be a full ban on certain AI systems that go beyond “high risk” but pose a fundamental, unacceptable risk for rights, consistent with rights standards. These include all types of remote biometric identification, emotion recognition, and biometric categorization using sensitive attributes. We invite the Advisory Body to define standards for unacceptable AI systems.

In sum, a rights-centered pillar on freedom of expression and privacy would accomplish several key objectives and aid the Advisory Panel in the following ways:

- It would offer consistency, clarity, and guidance for States as to their human rights obligations in this complex field;
- It would provide a participatory mechanism for stakeholders, including the private sector, Special Procedures, and civil society, to engage with creating human rights-centered best practices;
- It would reinforce the critical importance of protecting and promoting human rights, including rights to freedom of expression, through the continued development and application of AI;
- It would provide a process for transparency and accountability in the application and any abuses of AI.

ARTICLE 19 is prepared to offer any additional assistance and expertise that would be helpful to the Advisory Body as it considers these topics.

¹ ARTICLE 19 has monitored key developments in AI, including participating in a coalition for proposals to the European Parliament and Council for the Artificial Intelligence Act. We have made expert submissions to, e.g., the Institute of Electrical and Electronics Engineering (IEEE), the UK House of Lords Select Committee on AI, and the Article 29 Working Party of the European Data Protection Board. Together with Privacy International, we published [Privacy and Freedom of Expression in the Age of Artificial Intelligence](#), in April 2018.

² See e.g. the UN Special Rapporteur on freedom of expression, [Report on Artificial Intelligence technologies and implications for freedom of expression and the information environment](#), U.N. Doc. A/73/348, 29 August 2018; or Council of Europe, [Artificial Intelligence: Ensuring respect for democracy, human rights and the rule of law](#).

-
- ³ UN Office of the High Commissioner on Human Rights, [New and emerging technologies need urgent oversight and robust transparency: UN experts](#), 2 June 2023.
- ⁴ UN Office of the High Commissioner for Human Rights, [Artificial intelligence must be grounded in human rights, says High Commissioner](#), 12 July 2023.
- ⁵ Carey Shenkman, Dhanaraj Thakur, Emma Llansó, [Do You See What I See? Capabilities and Limited of Automated Multimedia Content Analysis](#), Center for Democracy and Technology, May 2021.
- ⁶ Office of the OSCE Representative on Freedom of the Media, [Freedom of the Media and Artificial Intelligence](#), 16 November 2020, p.1.
- ⁷ ARTICLE 19, [Self-regulation and ‘hate speech’ on social media platforms](#), 2018, p. 4.
- ⁸ Michelle Hampson, [Combating Hate Speech Online With AI](#), IEEE Spectrum, 21 February 2023.
- ⁹ Esha Bhandari, [Regulation of generative AI must protect freedom of expression](#), Open Global Rights, 2 June 2023.
- ¹⁰ [Disinformation and freedom of opinion and expression](#), Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, U.N. Doc. A/HRC/47/25, 13 April 2021; Human Rights Council, [Role of States in countering the negative impact of disinformation on the enjoyment and realization of human rights](#), U.N. Doc. A/HRC/49/L.31/Rev.1, 28 March 2022.
- ¹¹ ARTICLE 19, [Emotional Entanglement: China’s emotion recognition market and its implications for human rights](#), January 2021; ARTICLE 19, [EU: AI Act must protect prioritise fundamental rights](#), 19 April 2023.
- ¹² UN Human Rights Council, [Report of the Special Rapporteur on the Rights to Freedom of Peaceful Assembly and of Association](#), U.N. Doc. A/HRC/41/41, 17 May 2019, para. 57.
- ¹³ Council of Europe, Conference of Ministers responsible for Media and Information Society, [Artificial Intelligence—Intelligent Politics, Resolution on the safety of journalists](#), 10-11 June 2021.
- ¹⁴ Reporters Without Borders, [Online Harassment of Journalists: Attack of the trolls](#), 2018, p. 13.
- ¹⁵ Paul M. Barrett and Justin Hendrix, [Safeguarding AI: Addressing the Risks of Generative Artificial Intelligence](#), NYU Stern Center for Business and Human Rights, June 2023.
- ¹⁶ Council of Europe, [Implications of AI-Driven Tools in the Media for Freedom of Expression](#), 28-29 May 2020, p. 8, 16.
- ¹⁷ OECD, [Artificial Intelligence & Responsible Business Conduct](#), 2019; Amnesty International and Access Now, [The Toronto Declaration: Protecting the right to equality and non-discrimination in machine learning systems](#), 16 May 2018; Partnership on AI, [About Us](#), 2023; IEEE, [Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems](#), 12 December 2017, pp. 83-112.
- ¹⁸ Foundation for Law & International Affairs, [China’s New Generation of Artificial Intelligence Development Plan](#), 30 July 2017.
- ¹⁹ Human Rights Council, [Report of the United Nations High Commissioner for Human Rights](#), U.N. Doc. A/HRC/48/31, 13 September 2021, paras. 39, 42, 59(c); Human Rights Council, [New and emerging digital technologies and human rights](#), U.N. Doc. A/HRC/53/L.27/Rev.1, 12 July 2023, p. 4
- ²⁰ [Side-event of the Internet for Trust Global Conference](#), 7 February 2023.
- ²¹ The European Commission, for instance, defines “high-risk” AI to include systems of law enforcement that may interfere with fundamental rights, or migration, asylum, and border control management. The Commission recommends that these systems, if deployed, be subject to strict obligations before they can be taken to market. European Commission, [Proposal for a Regulation laying down harmonised rules on artificial intelligence](#), 21 April 2021.
- ²² Australian Human Rights Commission, [The Need for Human Rights-centred Artificial Intelligence](#), 26 July 2023, p. 49-50.