CELE
Centro de Estudios en Libertad de Expresión

Documento de posición

30

Public Consultation on Freedom of Expression and Artificial Intelligence

Vladimir Garay,
Matías González Mama,
Nicolás Zara y
Agustina Del Campo

28 de julio de 2025

Vladimir Garay, Matías González Mama, Nicolás Zara y Agustina Del Campo,"ublic Consultation on Freedom of Expression and Artificial Intelligence", Documento de posición No. 30 (ENG), Centro de Estudios en Libertad de Expresión (CELE), Buenos Aires (2025)

Facultad de Derecho

Centro de Estudios en Libertad de Expresión

UP
Universidad de Palermo

Public Consultation on Freedom of Expression and Artificial Intelligence

Vladimir Garay

Matías González Mama

Nicolás Zara

Agustina Del Campo

CELE/UP

cele@palermo.edu

July 28, 2025

The following document contains CELE's contribution to the call for input by the International Freedom of Expression Mandate Holders — the United Nations (UN) Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, the Organization for Security and Co-operation in Europe (OSCE) Representative on Freedom of the Media, the Organization of American States (OAS) Special Rapporteur on Freedom of Expression, and the African Commission on Human and Peoples' Rights (ACHPR) Special Rapporteur on Freedom of Expression and Access to Information in Africa — for a joint declaration on Artificial Intelligence and freedom of expression.

The Center for Studies on Freedom of Expression, CELE, is an academic legal research center. The Center develops research on cutting edge issues affecting the development of freedom of expression and access to information. Our research encompasses local, regional, global ,and comparative studies on state and private conduct affecting these rights.

On AI, the Center has been tracking discussions around legal developments regionally and internationally. In tracking these discussions, an alarming factor has been the disproportionate concern that AI use has created vis-à-vis the creation of content, particularly with GenAI. Because the concern has been mostly focused on the content produced -its potential for disinformation for example has been highlighted as a main concern- there is a special need for experts in the field of freedom of expression to clarify the extent to which traditional rules of international freedom of

expression standards should apply to these discussions; how those standards should be interpreted; and what novel issues need further analysis.

1. What do you consider to be the biggest risks and challenges for freedom of expression, including media freedom, independence and pluralism, when it comes to the design, development and deployment of AI and in particular GenAI? Please consider the entire lifecycle of the AI (and GenAI) and specify the layer where the risk or challenge is to be identified. If possible, please also provide examples on what you, in your field of work, expertise or region, would consider to be the most important issue(s).
(word count: 500)

Conversations and discussions around AI are, for the most part, polarized. There is a great deal of "technosolutionism" on one side of the scale and a great deal of panic on the other. There is barely anything in between, and public-policy conversations have been, in our experience, heated and speculative rather than technical. Few people in public policy discussions understand the technology or have the technical abilities to audit, oversee, or even assess the technology involved in AI or GenAI. Although these technologies are "experimental", as described by their own developers and creators, there is little understanding of what that should mean for its adoption in different spheres or by different actors. Therefore, most discussions on Artificial Intelligence in international, regional, and national instruments do not provide an in-depth evaluation of the relationship between AI and human rights.

There are different ways in which the adoption and evolution of AI may impact freedom of expression and access to information. There is more information available online than ever before, and this impacts for good the exercise of the rights to education, culture, access to information (on many aspects, including health, public accountability, environment, etc), and freedom of expression. AI, we are told, provides new tools to make searches, content creation, learn, creativity, and innovation much more efficient, quicker, and more accurate and draw the benefits of aggregated knowledge for an individual to take advantage of. It also has potential risks for freedom of expression that are deeply linked to its uses:

Content moderation and curation:

One of the biggest risks to freedom of expression in the context of AI use is the overreliance of online platforms on automated systems for content curation and moderation. These technologies lack the ability to understand the specific context in which content appears, which often leads to errors—typically the over-removal of legal content. The most well-documented cases involve automated moderation related to copyright infringement, where the mere appearance of protected material triggers algorithms to remove or suppress content without recognizing legitimate exceptions that may render the content not only legal but also publicly relevant. For example, there are documented cases in the United States where police officers play loud, copyrighted music during patrols so that videos of them are more likely to be taken down if posted online, weaponizing both copyright law and automated systems for content moderation to suppress criticism and civil accountability.

Automated decision-making systems in public surveillance, such as facial recognition technology:

Facial recognition technology, powered by AI subjects the entire population to a disproportionate surveillance regime and can be used to identify individuals who participate in public demonstrations and protests, creating a chilling effect on freedom of expression, especially when deployed without an adequate legal framework that respects human rights and without proper measures for control, accountability, and independent oversight. An example of this happened in Buenos Aires, Argentina, where a facial recognition system originally implemented to locate fugitives was used to surveil political opponents.

AI replacing search without sources:

AI services create summaries of different information found in different sources available to their systems. They generate automated answers to users' queries, which reduces traffic to primary sources and, consequently, limits revenue for media and journalistic outlets. This shrinks the diversity of media voices and weakens the quality and pluralism of public debate. AI users lose the ability to see and check by themselves the sources used in the construction of answers that AI systems offer, breaking the link between the original sources and the readers. This may contribute to ongoing issues like disinformation.

State use of experimental tech like AI:

There are different instances where State actors have been adopting AI to develop their everyday work. Judges and lawyers have been using AI to generate arguments, summarize information, and even draft decisions. The Police has been using AI to scroll social media for a while now, developing non-regulated practices of OSINT that lack accountability and oversight. Different public offices have been adopting AI as aids in their everyday work, without oversight or guidelines or any transparency as to what systems they may be using, what the use entails for privacy and data protection, or copyright for that matter.

2. What key recommendations should the Freedom of Expression Mandate Holders issue to respond to the risks and challenges you identified? For which stakeholders should the recommendations be developed?

Conduct assessments of AI's potential positive and negative impacts on the exercise of human rights, particularly freedom of expression, to better understand why this technology matters and what aspects are worth protecting.

For online platforms: Avoid overreliance on fully automated content moderation systems, especially when fundamental rights like freedom of expression and access to information are at stake. Establish robust human-in-the-loop mechanisms to review decisions made by automated tools, particularly for borderline or context-sensitive content. Implement clear, transparent, and expedited procedures to appeal decisions made by automated content moderation systems. Be transparent as to when AI is used and what content it impacts.

For States: Introduce clear legal frameworks, rooted in Human Rights, that explicitly ban or strictly limit the use of AI-based surveillance in contexts that may unduly restrict privacy, freedom of expression, and assembly, such as protests. Ensure transparency procedures about the acquisition, deployment, scope, and uses of surveillance systems. Companies that develop, commercialize, or deploy AI systems should be mandated to conduct periodic Human Rights impact assessments of their service, to evaluate their potential impact on Freedom of Expression and other rights. These assessments should be made public. In the event that a potential adverse impact was identified, companies should address it in a way that is adequate for the nature of the system and the specific use case. The measures taken to address these potential negative impacts must always be the least rights-restrictive alternative. For instance, in the case of AI systems used for algorithmic content moderation, these measures could include appeals mechanisms for content

removal decisions (false positives) and flagging systems to identify illegal content that was missed by the algorithm (false negatives).

Regarding the sustainability of media and journalistic outlets, given the loss of revenue opportunities resulting from the deployment of generative AI tools in search engines, states and private companies should explore and implement alternative forms of financing for the media ecosystem, thereby ensuring a more diverse and robust landscape of information and public debate. Search engine platforms should implement design choices that systematically promote and credit original sources, including clear citations and direct links to primary content. State agents must foster respect for media and journalists, even when disagreements may arise, and promote a healthy environment for journalistic expression.

3. What key recommendations should the Freedom of Expression Mandate Holders issue to tackle the role of transparency with regards to the design, development and use of AI and GenAI? Please consider providing inputs on what you believe to constitute the minimum required information that should be public, on how to guarantee its accessibility and intelligibility and/or on any other standard on transparency that you consider relevant.
(word count: 300)

Different actors have different transparency obligations. States should be transparent about their acquisition, development, use, and oversight mechanisms developed for AI and GenAI. Specific guidelines should be adopted where necessary to make sure that employees and organisms at different levels of the State understand their obligations to follow, disclose, and be transparent about what technology they use and depend upon to develop their everyday work. States should also set clear guidelines for instances where AI and GenAI technologies are used to communicate with citizens, like chatbots or other tools that they may be using or deploying.

Companies should be transparent about the use of AI and GenAI in the development of their businesses. For example, if AI is used to curate or moderate content, companies should be transparent about the kinds of technology they are using, the contents that are being impacted and the systems in place to provide oversight.

Where the company provides AI tools for third party users, it should be transparent about their experimental nature, the accuracy that the systems have, and rates of error. Under UNGP,s companies should abstain from selling AI services to states for uses that may compromise human rights of the people under their jurisdictions, particularly when the technology is experimental and the risks outweigh the benefits. Should they decide to sell these services to State agents or organisms, they should conduct oversight as to the uses and impact of these services on human rights and be transparent about them.

4. What do you consider to be the major opportunities for freedom of expression, including media freedom, independence and pluralism, when it comes to the design, development and deployment of AI and in particular GenAI? Please specify and, if possible, provide examples on what you, in your field of work, expertise or region, would consider to be the most important issue(s). Where relevant, please specify also what measures and guarantees must be put in place to preserve these advantages.
(word count: 500)

The process of massification of AI is still unfolding, with countless opportunities for both success and failure still to be discovered. In Latin America, a noteworthy instance of AI being used in favor of freedom of expression is that of "Operación Retuit," a Venezuelan YouTube channel where two AI-generated avatars ("La Chama" and "El Pana") delivered news flashes with information gathered from different local and international sources. The utilization of avatars was vital to protect the anonymity, privacy, and physical integrity of the journalists behind the project.

5. What guarantees should be put in place to ensure that the opportunities that AI offers for advancing freedom of expression benefit everyone, including women, minorities, vulnerable, underrepresented, and other marginalised groups and communities across the globe? (word count: 300)

To ensure that AI promotes freedom of expression for all, it is essential to counteract the monopolistic tendencies of the current industry landscape. Concentrated control over large AI systems risks embedding narrow cultural, political, or economic biases. Encouraging the development and coexistence of multiple AI systems, built on diverse datasets and trained with different epistemological, linguistic, and cultural perspectives, can offer users a plurality of viewpoints and a wider range of responses to the same queries.

Open AI systems can play a key role in this process. They allow communities, including marginalized groups, to audit, adapt, or develop models that better reflect their values, languages, and lived experiences.

States could promote diversity and deconcentration through specific rather than general regulation. AI offers different opportunities and challenges depending on its uses. Risks and opportunities for the scientific community are not the same as those for state policing or surveillance. Public facing AI, like Chatgpt create challenges for privacy and freedom of expression that are different from those AI developed for private use, like writing aids. The AI regulatory conversation can benefit from nuances and a purpose approach that are currently missing in the discussion.

6. What are the specific challenges and needs for freedom of expression and freedom of the media that arise when AI systems are designed, developed and deployed by the State, either alone or within the framework of public-private partnerships?
(word count: 300)

A major concern can be opacity.

When State use of AI lacks transparency, it makes it difficult to assess how AI systems function, what data they rely on, and whether they are used to suppress dissent, silence minorities, or target critical media and journalists. Public-private partnerships can further complicate accountability, as companies could invoke trade secrets to avoid disclosure, therefore affecting transparency.
AI-enabled surveillance (i.e., facial recognition among others), whilst affecting privacy, can create a chilling effect on freedom of expression, especially in protests or politically sensitive contexts. This can have a chilling effect on journalists, activists, and citizens who may prefer to self-censor due to fears of being monitored or profiled.

States must ensure transparency when designing, developing, and deploying AI systems, conduct mandatory human rights impact assessments, and establish robust accountability mechanisms for the public and private sectors. Civil society, academia, and the technical community must be involved in the oversight and design processes of AI systems.

When States set standards for and design, develop, and deploy AI system,s they should be wary of promoting ideological conformity under the guise of neutrality. As seen in certain proposals that

claim to defend freedom of expression, by removing references to topics deemed ideological, such as misinformation or climate change, from public AI frameworks, and by requiring systems to reflect a notion of "truth" free from "social engineering", States may inadvertently chill speech and distort the information ecosystem. While support for open-source AI is encouraging, conditioning public contracts on vague political criteria threatens both developers' autonomy and public access to diverse viewpoints. Genuine leadership in AI requires safeguarding pluralism, not enforcing a single agenda through governmental pressure.

7. A variety of international, regional and national instruments (among others: the UN Global Digital Compact, the Council of Europe Framework Convention on AI and Human Rights, the Africa Declaration on Artificial Intelligence, the OECD AI principles, the UNESCO Recommendation on the Ethics of Artificial Intelligence, the EU AI Act, the Ministers and high-level authorities of science and technology of the Americas' Declaration and Plan of Action: Towards the Safe, Secure and Trustworthy Development and Deployment of Artificial Intelligence in the Americas) have already been put in place to deal with the impact of AI on freedom of expression and, to a certain extent, on the media. What are the main strengths and weaknesses of these instruments and what are the gaps still to be filled?
(word count: 300)

International, regional, and national initiatives on AI governance show both strengths and notable gaps regarding the impact of AI on freedom of expression and the media.

Among their strengths, the UNESCO Recommendation on the Ethics of AI explicitly promotes freedom of expression, academic freedom, and access to information, emphasizing transparency and content moderation practices to counter disinformation. The OECD AI Principles also highlight the importance of safeguarding freedom of expression in the face of potential risks and biases. UN General Assembly Resolution 78/265 introduces important commitments to transparency, bias mitigation, protection of linguistic and cultural diversity in language models, as well as increasing media and information literacy. The Council of Europe's Framework Convention further reinforces the role of AI governance in supporting democracy and the free formation of opinions, with a focus on transparency in AI-generated content.

However, there are significant weaknesses. The GDC promotes transparency and the mitigation of disinformation in AI-generated content. Its weaknesses include the lack of robust commitments to human rights and independent oversight, as well as ambiguity regarding sensitive uses, which could negatively impact freedom of expression and informational diversity. The EU AI Act, while ambitious, relies heavily on a risk-based approach that often defers human rights protection to self-assessments by developers and providers. There are also serious risks of disproportionate censorship, as vague regulatory language around "harmful content" or disinformation can lead to overreach and the restriction of legitimate speech.

Developing countries have had limited participation in shaping these frameworks, raising concerns about their relevance and applicability across diverse contexts.

Unresolved issues remain around the attribution of legal responsibility and copyright for AI outputs and the extent to which these outputs are protected under freedom of expression. Some instruments, such as the OAS Declaration, lack specificity and concrete commitments, limiting their effectiveness. To address these gaps, future governance efforts must ensure binding safeguards, inclusive participation, especially from the Global Majority, and a stronger alignment with international freedom of expression standards.

8. What is the key impact of the current concentration of power across the AI, and in particular the GenAI, stack on freedom of expression and media freedom, independence and pluralism? What market and business dynamics (including market concentration, vertical integration, structural or economic dependencies) need to be addressed to solve the issue(s) and what are the specific instruments to be used? Please specify and, if possible, provide examples on what you, in your field of work, expertise or region, would consider to be the most important issue(s).
(word count: 300)

Companies in the already hyper-concentrated markets of social media platforms and search engines are increasingly relying more and more on Artificial Intelligence to curate and moderate content, i.e., to apply the rules that govern what information is accessed and how expression is treated in the online realm. The vertical integration of these services makes entry barriers higher for potential competitors who lack the knowledge or resources needed to adopt and deploy these tools. The effect of this is the difficulty for new actors, particularly those with alternative approaches to

content moderation or substantive commitments, to enter the market and compete for the space these companies occupy. This accentuates the market concentration already referred to, and narrows the number of fora in which discussions can be had. As already flagged, this makes their operations more and more opaque and hard to scrutinize for governments, which can't match their technical capacities, and challenging for users affected by the way these systems operate and the individual decisions they make about user-generated content.

9. What are, in your opinion, the key safeguards and measures to be put in place to guarantee that the design, development and deployment of AI, and in particular GenAI, guarantee and enhance human agency, rather than limiting or suppressing it? Please specify and, if possible, provide examples on what you, in your field of work, expertise or region, would consider to be the most important issue(s).
(word count: 300)

It is essential that potential users can understand what AI tools can and cannot do. In the case of GenAI systems, users should have access to clear and concise explanations of what each GenAI system is and what it is not, including how it works and its technical capacities and limitations. This information can empower users to make informed decisions about whether and how to use these systems effectively, and to be aware of the potential shortcomings in the outputs they produce.

10. What are the main purposes that AI, and in particular GenAI, should serve with regards to the media sector? Are there any tasks, processes and functions in the media sector that should never be delegated to or performed by AI systems? Please specify and, if possible, provide examples on what you, in your field of work, expertise or region, would consider to be the most important goals and/or red lines.
(word count: 300)

This is a difficult question because the key issue lies in business models and economic incentives. In theory, generative AI could be used by media outlets to handle routine and repetitive tasks, such as transcriptions, translations, formatting, or initial data processing. This could free up journalists and editorial staff to focus on more substantive reporting, investigative work, or simply have more time to rest, improving both quality of life and journalistic depth. However, an entirely different

scenario is also plausible: a media outlet might choose to reduce costs by replacing most of its staff with AI tools, delegating not only support tasks but the most substantial ones to automated systems. This could lead to a sharp decline in the quality of information, but if it remains profitable, then probably that would be a non issue.

Regarding the functions that should never be delegated to AI systems, AI systems should not be granted final decision-making authority, especially in contexts where accountability, ethical judgment, or legal liability are involved.

Not answered:

11. Can you identify best practices – whether from the private sector, governments or civil society organisations – related to the use of AI (and GenAI) in the media ecosystem? If so, please indicate them. Links to these materials are appreciated.