

Contribución del Centro de Estudios en Libertad de Expresión a la Consulta - Impacto de la Inteligencia Artificial en los Derechos Humanos en las Américas de la RELE/CIDH

CELE

Febrero de 2026

CELE, "Contribución del Centro de Estudios en Libertad de Expresión a la Consulta - Impacto de la Inteligencia Artificial en los Derechos Humanos en las Américas de la RELE/CIDH", Documento de posición No. 40 (ESP), Centro de Estudios en Libertad de Expresión (CELE), Buenos Aires (2026)

Facultad de Derecho

Centro de Estudios en Libertad de Expresión

UP
**Universidad
de Palermo**



Contribución del Centro de Estudios en Libertad de Expresión a la Consulta - Impacto de la Inteligencia Artificial en los Derechos Humanos en las Américas de la RELE/CIDH

Buenos Aires, Argentina

2 de marzo de 2026

CELE/UP

cele@palermo.edu

Introducción

El presente documento contiene la contribución del Centro de Estudios en Libertad de Expresión (CELE) a la convocatoria de aportes realizada por la Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos (CIDH), en el marco de la elaboración de estándares y análisis sobre inteligencia artificial y libertad de expresión.

El CELE es un centro académico de investigación jurídica dedicado al estudio de cuestiones contemporáneas que inciden en el desarrollo y la protección de la libertad de expresión y el acceso a la información. El Centro lleva adelante investigaciones de alcance local, regional, global y comparado sobre conductas estatales y privadas que afectan estos derechos.

En materia de inteligencia artificial, el CELE sigue de cerca los debates regulatorios y los desarrollos normativos tanto a nivel regional como internacional. En este contexto, resulta particularmente relevante que los especialistas en libertad de expresión contribuyan a precisar en qué medida los estándares tradicionales del derecho internacional en la materia resultan aplicables a estos debates; cómo deben interpretarse frente a los nuevos desarrollos tecnológicos; y qué cuestiones novedosas requieren un análisis adicional y más específico.

5. Impacto del uso de inteligencia artificial, incluyendo la inteligencia artificial generativa, en derechos específicos

- a. *¿Cuál ha sido el impacto del desarrollo y/o uso de sistemas de inteligencia artificial, y de inteligencia artificial generativa sobre el derecho a la libertad de expresión, el pluralismo y la libertad de prensa en la región? ¿Cuáles son los principales retos o desafíos?*

Los sistemas de inteligencia artificial pueden asistir y facilitar la producción y circulación de distintas formas de discurso, en distintos formatos (texto, imagen, audio, video). La posibilidad de generar “medios sintéticos” —que simulan o recrean eventos, personas o sonidos mediante algoritmos, como los *deepfakes*— puede ser una herramienta útil para ciertas formas de expresión artística, la innovación cultural, la sátira y la crítica política. Además, la IA ofrece herramientas útiles para tareas tales como la traducción, la transcripción, el análisis de datos y la producción de informes, entre otros.

Por ello, esta tecnología presenta un potencial que puede ser usado en favor del ejercicio de la libertad de expresión, en general, y para el ejercicio del periodismo, en particular. En América Latina hay casos documentados¹ de medios de comunicación independientes que están explorando los usos y beneficios de las herramientas de inteligencia artificial.

De forma general, se entiende también que la inteligencia artificial podría tener impactos problemáticos o riesgosos, vinculados con la utilización de la tecnología para cometer ilícitos civiles o penales, para facilitar la violencia de género mediante la difusión no consentida de imágenes íntimas sintéticas, o a través de estafas y suplantación de identidad. Otra arista problemática tiene que ver con la manera en que los sistemas de IA pueden reforzar las inequidades existentes; por ejemplo, el entrenamiento con un set de datos sesgados puede reproducir dichos sesgos y estereotipos, mientras opera bajo la ilusión de neutralidad.

Sin embargo, es importante reconocer que la inteligencia artificial se encuentra todavía en estado de desarrollo, no solamente en el aspecto técnico, sino también en términos regulatorios y respecto de su modelo de negocio. Un claro ejemplo de esto han sido las

¹ “IA y periodismo en América Latina: ¿quiénes innovan?” Deutsche Welle, August 11, 2025, <https://akademie.dw.com/es/ia-y-periodismo-en-américa-latina-conoce-a-quienes-impulsan-la-innovación-en-el-continente/a-73559562>.

propuestas de *backtracking* regulatorio propuestas por la Unión Europea,² luego del informe Draghi,³ y el cambio de política de la administración estadounidense luego de la asunción de Trump.⁴ Todos estos ámbitos inciden sobre el contexto sociotécnico que, a su vez, determina las formas posibles en que su potencial se desarrollará.

Por ejemplo, la IA puede ser un apoyo para el ejercicio del periodismo, no solamente asistiendo en tareas como el análisis de grandes volúmenes de datos, sino que además podría —por ejemplo— usarse para la realización de tareas tediosas y rutinarias, como la confección de reportes, liberando tiempo valioso que puede ser invertido en cuestiones más significativas, como la investigación periodística. Como resultado, se obtiene la circulación de información de mayor calidad. Esto supone, entre otras cosas, la imposición de un criterio periodístico. Sin embargo, la misma tecnología puede tener un impacto completamente diferente si el criterio a partir del cual se despliega es distinto. Por ejemplo, un sistema de IA puede implementarse con la finalidad de abaratar costos y prescindir de un número importante de periodistas, delegando al sistema la confección de la información. Muy probablemente el resultado sea la puesta en circulación de un producto informativo de inferior calidad.

Por ello es importante evitar análisis basados en un errado determinismo tecnológico, sobre todo en el estadio de desarrollo actual de la inteligencia artificial. No solamente es muy pronto para evaluar sus impactos, sino que estos no están determinados por un criterio únicamente técnico. Pensemos, por ejemplo, en los aspectos de índole económica: desde una perspectiva de libertad de expresión existe un interés en favor de que una diversidad de actores, trabajando en distintas escalas, disponible acceso a distintos sistemas de inteligencia artificial que hayan sido entrenados de distinta forma y que puedan ofrecer *outputs* diversos. Para ello, es necesario que una diversidad de actores pueda acceder a los datos necesarios para poder entrenar dichos sistemas. Sin embargo, actualmente existe un cuestionamiento —principalmente, por parte de empresas titulares de derechos de autor, pero también de distintos tipos de creadores— respecto a la posibilidad de que sus obras sean incluidas en los sets de entrenamiento de la IA, sin compensación monetaria. Usualmente, lo que se propone es la creación de algún sistema de licenciamiento que habilite el uso. Sin embargo, este tipo de medidas crea un umbral de acceso más alto que impide que aquellos actores pequeños

² Matias Gonzalez Mama et al., “Global AI Governance: Who Regulates, with What Approach, and for Whom?,” *SSRN Electronic Journal* CELE Research Paper No. 67 (June 2025), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5311517.

³ https://commission.europa.eu/topics/competitiveness/draghi-report_en

⁴ <https://www.mofo.com/resources/insights/250728-new-federal-ai-action-plan-prioritizes-deregulation>

—empresas de desarrollo local, universidades y centros de investigación, que no cuentan con los recursos suficientes para pagar las licencias puedan participar del desarrollo de sistemas de IA, lo que colabora a que el campo sea cooptado por un grupo pequeño de grandes empresas, lo que a su vez tiene un impacto negativo sobre el ejercicio de la libertad de expresión.

La regulación, el desarrollo de mercados, la percepción pública son algunos de los factores que van a determinar las posibilidades concretas de uso de los *affordances* técnicos. Desde este punto de vista, resulta valioso que en América Latina se estén desarrollando agendas de investigación y desarrollo vernáculos, y estos esfuerzos tienen el potencial de bajar las barreras de entrada a una tecnología disruptiva pero prometedora.

b. ¿Cuál ha sido el impacto del desarrollo y/o uso de sistemas de inteligencia artificial y de inteligencia artificial generativa sobre el derecho a la libertad de expresión artística y el derecho a la libertad académica? ¿Cuáles son los principales retos o desafíos?

Los sistemas de inteligencia artificial y de inteligencia artificial generativa pueden contribuir a ampliar el campo de experimentación estética y artística.⁵ Pueden asistir en los procesos de creación artística mediante el acceso y la aplicación de diversos medios, técnicas y métodos, y pueden funcionar como herramientas de investigación, comparación y fusión de teorías y tradiciones de la producción artística y de trabajos de artistas en particular. También pueden ampliar los medios de experimentación artística al facilitar la utilización de materiales sintéticos que antes no resultaban fácilmente asequibles para los procesos de creación. En algunos casos pueden servir como inspiración para sortear el problema de la hoja en blanco y proporcionar material en bruto sobre el cual trabajar.⁶

⁵ Eric Zhou and Dokyun Lee, “Generative Artificial Intelligence, Human Creativity, and Art,” *PNAS Nexus* 3, no. 3 (2024), <https://doi.org/https://doi.org/10.1093/pnasnexus/pgae052>.

⁶ *Embracing Creativity: How AI Can Enhance the Creative Process*, <https://www.sps.nyu.edu/about/news-and-ideas/articles/etc/2024/embracing-creativity-how-ai-can-enhance-the-creative-process.html#:~:text=Inspiration%20and%20Idea%20Generation:%20AI,springboards%20for%20crafting%20engaging%20narratives>.

Pueden contribuir a reducir ciertos costos de la producción artística y asistir en procesos legales o administrativos vinculados a ella, así como en la definición de objetivos y en planificación y organización de procesos y planes de trabajo. Además, puede servir como herramienta para que personas con inquietudes artísticas, pero sin entrenamiento formal, se expresen a través de distintos medios y soportes. Esto no implica que la inteligencia artificial reemplace o sustituya el proceso de formación y de aprendizaje necesario para devenir artista, ni que las obras resultantes —cuando son realizadas por un aficionado— tengan una calidad similar a una producida por un profesional. Más bien, los sistemas de inteligencia artificial pueden complementar y asistir en la labor y producción de obras.

Algunos estudios y reportes que señalan que los *outputs* de la IA reproducen los sesgos de los *datasets* empleados en su entrenamiento,⁷ lo que podría tener efectos condicionantes en la expresión artística. En esta línea, Alexandra Xanthaki, Relatora Especial sobre los derechos culturales, sostuvo que los “productos culturales generados por IA reflejan las normas dominantes y sus sistemas de recomendación producen resultados distorsionados, sea porque están configurados por modelos con fines de lucro o porque reflejan lagunas en los datos. Como las herramientas de IA reproducen los datos con los que han sido entrenadas de manera acrítica e incontestada, no sorprende que su contenido sea parcial, estereotípico y discriminatorio”.⁸ También destacó que la inteligencia artificial tiende a privilegiar aquellas expresiones creativas que resultan comercialmente exitosas, lo que podría limitar la diversidad cultural y favorecer procesos de homogenización por reforzar los estilos y patrones más frecuentes.⁹

Si bien se trata de preocupaciones atendibles, consideramos que estas conclusiones pueden resultar apresuradas. En primer lugar, porque aún no contamos con evidencia suficiente que respalde tales afirmaciones respecto de la tecnología en sí misma, con independencia de cada sistema en particular y de los contextos regulatorios, económicos y culturales en los que se implementa. En segundo lugar, porque preocupaciones análogas podrían plantearse respecto de cualquier plan de un estudio de una carrera artística. Éstos están estructurados sobre tradiciones y criterios igualmente selectivos y sus efectos en la formación y en la producción artística podrían ser comparables en términos de condicionamientos.

⁷ *Datasets used in algorithmic decision-making systems*, European Parliamentary Research Service, 2022, [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/729541/EPRS_STU\(2022\)729541_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/729541/EPRS_STU(2022)729541_EN.pdf)

⁸ Alexandra Xanthaki, *A/80/278: Informe de La Relatora Especial Sobre Los Derechos Culturales*, Alexandra Xanthaki - Inteligencia Artificial y Creatividad (Naciones Unidas, 2025).

⁹ *Ibíd.*

Una cuestión distinta es el impacto que pueden tener los sistemas de inteligencia artificial empleados para la moderación automatizada de contenidos. En estos casos, resulta fundamental que quienes los implementan realicen una evaluación específica de su potencial incidencia sobre la libertad de expresión artística, tomando en consideración, en particular, que no operen como mecanismos de filtrado previo o upload filters.

En el ámbito de la libertad académica, estos sistemas pueden facilitar el acceso a herramientas de investigación y ampliar las capacidades de análisis disponibles. Según diversos autores, pueden contribuir a la generación de ideas y al diseño de investigaciones, ayudar con el contenido y su estructuración, apoyar la revisión y síntesis de la literatura especializada, optimizar la gestión y el análisis de datos, colaborar en tareas de edición, revisión y publicación, y asistir en los procesos de comunicación y difusión de resultados.¹⁰

Mediante el *text and data mining*, puede contribuir al análisis de grandes conjuntos de datos identificando patrones, tendencias y correlaciones.¹¹ También pueden ayudar en la creación de otras herramientas de IA que, en función de necesidades y saberes específicos, resulten de utilidad para la investigación científica y académica. Además, pueden facilitar el acceso a investigaciones y obras en otros idiomas que no dispongan de traducciones promoviendo una mayor circulación del conocimiento.

Sin embargo, existen desafíos relevantes: si el entrenamiento de los sistemas de IA queda sujeto a modelos cerrados de licenciamiento, los costos de entrada aumentan, afectando particularmente a universidades y desarrolladores independientes, especialmente en países en desarrollo. Las crecientes demandas de protección de derechos de autor podrían restringir prácticas de investigación académica facilitada por la inteligencia artificial.

¹⁰ Mohamed Khalifa and Mona Albadawy, *Using Artificial Intelligence in Academic Writing and Research: An Essential Productivity Tool*, las (2024), https://www.sciencedirect.com/science/article/pii/S2666990024000120?__cf_chl_tk=qRnqUZ3a1TNCCbCnroROCyGh.eWRkT4gwo0zJYVCorg-1772225677-1.0.1.1-yZ0X1Az5d_N.njnnFLW.z3pffmq1k4IdD5BQPtO58GQ.

¹¹ Text and Data Mining (TDM) for Researchers: Policies, Tools, Use Cases, <https://library.hkust.edu.hk/news-events/news/text-and-data-mining-tdm-researchers-policies-tools-use-cases#:~:text=In%20today's%20data%2Ddriven%20research,Offering%20TDM%20Services%20for%20>

Por cuestiones de brevedad, y en tanto resultan igualmente aplicables a esta sección, remitimos a las advertencias formuladas en la pregunta anterior respecto del determinismo tecnológico, la evaluación de los impactos de la tecnología y los efectos restrictivos sobre el derecho a la libertad de expresión de un potencial licenciamiento de las obras utilizadas.

- c. En el caso de plataformas digitales, ¿cuáles son las consecuencias del uso de inteligencia artificial en la moderación de contenidos u otras decisiones que afectan el derecho a la libertad de expresión de los usuarios? ¿Existen salvaguardias para evitar discriminación o arbitrariedad?

Desde hace varios años, las principales plataformas digitales utilizan sistemas automatizados para moderar los contenidos publicados por sus usuarios. Esta es una respuesta a un problema de escala: las grandes plataformas reciben diariamente millones de publicaciones nuevas, al punto de que es simplemente imposible que estas sean evaluadas solamente por revisores humanos.

Sin embargo, los sistemas automatizados son incapaces de comprender el contexto —cultural, lingüísticos, políticos, sociales, históricos y económicos— en el que se inscribe la comunicación,¹² así como la proxémica, el tono y el registro de la enunciación. Aspectos como la ironía, el cinismo, el humor y la crítica son difíciles de discernir para los sistemas de IA.¹³ Esto resulta en errores, que van desde la remoción de contenidos legítimos (que afectan diferencialmente a poblaciones históricamente marginalizadas), hasta la difusión no controlada de discursos prohibidos por el sistema interamericano.

Esta dificultad, sumada al volumen ingente de publicaciones que requieren revisión, ha llevado a algunos especialistas a sostener que se trata de una tarea imposible de realizar a la perfección,¹⁴ lo que no significa mantener una actitud complaciente respecto al desempeño de las empresas que ofrecen servicios en internet. Existen algunas medidas que las compañías

¹² Ibidem.

¹³ Compare Council of Europe study ‘Algorithms and Human Rights: Study on the human rights dimensions of automated data processing techniques and possible regulatory implications’, DGI(2017)12, 21.

¹⁴ Mike Masnick ha escrito al respecto. Ver por ejemplo “Masnick’s Impossibility Theorem: Content Moderation At Scale Is Impossible To Do Well,” Techdirt, November 20, 2019, <https://www.techdirt.com/2019/11/20/masnicks-impossibility-theorem-content-moderation-scale-is-impossible-to-do-well/>.

pueden implementar en favor de mejorar la moderación de contenidos: mejores medidas de transparencia respecto de las decisiones tomadas y la manera en que se adoptaron, establecer mecanismos efectivos de apelación, garantizar la revisión humana posterior cuando sea necesaria.

Un elemento que es importante mantener a la vista es que las decisiones en materia de moderación de contenido —incluyendo el despliegue de soluciones técnicas con esta finalidad— son producto de distintos incentivos, desde los propios intereses comerciales de las plataformas (generar un espacio suficientemente atractivo para atraer tanto a usuarios como a empresas dispuestas a pagar publicidad) hasta requerimientos normativos. Por ejemplo, los sistemas de remoción de contenido por violación a los derechos de autor de las grandes plataformas de internet están hechos siguiendo los lineamientos de la DMCA estadounidense, que habilita el sistema de notificación privado para la remoción de contenido infractor y que ha demostrado ser fácilmente explotable, incluso con fines de censura.¹⁵ En ese sentido, uno de los principales desafíos vinculados a los sistemas de IA proviene de respuestas regulatorias desproporcionadas, que no toman en cuenta la potencial afectación a la libertad de expresión que podrían ocasionar.

- d. En relación con la integridad de la información, ¿qué medidas han adoptado los Estados y actores privados para contrarrestar la generación y difusión de contenido falso o engañoso generado mediante inteligencia artificial (por ejemplo, *deepfakes*)?

En relación con esta pregunta, es necesario recordar que “integridad de la información” no constituye una categoría jurídica reconocida en el derecho internacional ni genera obligaciones para los Estados distintas de las ya existentes en materia de libertad de expresión. Se trata de un concepto reciente, de carácter programático y político, no de una categoría normativa consolidada.

Desde esa perspectiva, corresponde advertir que muchos de los daños asociados a la generación y difusión de contenido engañoso mediante inteligencia artificial ya se encuentran cubiertos por marcos regulatorios vigentes. La manipulación electoral debe evaluarse conforme a la legislación electoral; los *deepfakes* de carácter sexual pueden encuadrarse en

¹⁵ Courtney Radsch, «Weaponizing Privacy and Copyright Law for Censorship», *SSRN Electronic Journal*, advance online publication, 2023, 12, <https://doi.org/10.2139/ssrn.4464300>.

normas sobre violencia de género, protección del honor o delitos contra la intimidad; conductas como el fraude, la suplantación de identidad o la falsificación están tipificadas en la mayoría de los ordenamientos jurídicos; y la incitación a la violencia o al odio se encuentra regulada bajo estándares internacionales consolidados.

En consecuencia, la discusión no debería centrarse en la creación de nuevas categorías amplias e indeterminadas como “integridad de la información”, sino en la correcta aplicación de las normas existentes conforme al principio de legalidad y al test estricto de necesidad y proporcionalidad en materia de libertad de expresión.

Adoptar marcos regulatorios expansivos bajo una noción imprecisa de “integridad” implica riesgos estructurales: introduce estándares abiertos sobre la “calidad” o “veracidad” del discurso que no cuentan con delimitación jurídica clara en el derecho internacional; amplía el margen de discrecionalidad estatal en materia de control del contenido; y facilita su utilización para fines de censura o silenciamiento político, especialmente en contextos institucionales frágiles.

Asimismo, regular tecnologías específicas por el solo hecho de su novedad —en este caso, la IA generativa— tensiona el principio de neutralidad tecnológica, que exige que el derecho regule conductas y efectos jurídicamente relevantes, no herramientas en abstracto.

- e. ¿En caso de uso de inteligencia artificial para mecanismos de vigilancia por parte del Estado, ¿ha sido evaluado el impacto en los derechos a la libertad de expresión, asociación y reunión? ¿Qué medidas se han implementado para evitar violaciones de estos derechos?

La adquisición e implementación de tecnologías de vigilancia con fines de seguridad por parte de los Estados es una práctica que se ha incrementado en los últimos años. Este fenómeno ha sido advertido por organismos internacionales de derechos humanos. En 2019 el Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión de Naciones Unidas dijo que estas prácticas “ya no está[n] reservad[as] a los países que disponen de los recursos necesarios para llevar a cabo una vigilancia en masa o selectiva

mediante instrumentos elaborados a nivel interno”,¹⁶ sino que se ha expandido a Estados con menores recursos por el significativo aumento de oferta. Asimismo, “al desarrollar y desplegar nuevas tecnologías y herramientas de comunicaciones de formas determinadas, los agentes empresariales también han adoptado voluntariamente medidas que facilitan la vigilancia de las comunicaciones por el Estado”.¹⁷

La utilización de tecnologías en contextos de seguridad ha sido evidente, por ejemplo, en las políticas contraterroistas de los estados. La capacidad de vigilancia masiva como herramienta rutinaria en estrategias antiterroistas ha crecido enormemente por tendencias como la caída en costos de tecnología y almacenamiento de datos, la omnipresencia de dispositivos digitales conectados y el aumento exponencial en potencia computacional.¹⁸

La vigilancia digital tiene un impacto directo sobre el derecho a la privacidad, pues implica un alto nivel de intromisión en la esfera íntima de la persona que es objeto de esta. David Kaye explicó como “la privacidad y la libertad de expresión están entrelazadas en la era digital, y la privacidad en línea es el punto clave para garantizar el ejercicio de la libertad de opinión y de expresión”.¹⁹ La interdependencia de los derechos humanos en los contextos de uso de tecnologías digitales ha conllevado a que el impacto que tienen sus usos, sean relevantes en el debate público, especialmente cuando se trata de usos estatales y de herramientas cada vez con mayores capacidades, permitidas ahora por la incorporación de sistemas basados en inteligencia artificial. Es por ello que la adquisición y uso de tecnologías por parte de los Estados requieren de evaluaciones de derechos humanos antes, durante y después de su implementación.

Asimismo, las normativas que regulan la adquisición y uso de estas tecnologías en algunos casos son deficientes o ambiguas. Un ejemplo es el de la aplicación de técnicas de investigación

¹⁶ Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión. La vigilancia y los derechos humanos Informe del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión. A/HRC/41/35. 41er período de sesiones. Consejo de Derechos Humanos, 2019. <https://docs.un.org/es/A/HRC/41/35>.

¹⁷ Frank La Rue. *Informe del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y expresión*, Frank La Rue. A/HRC/23/40. Consejo de Derechos Humanos, 2013.

¹⁸ United Nations Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism. Civil Society Under Attack A global study in defense of civic space. 2024. <https://defendcivicspace.com/>

¹⁹ David Kaye. *Informe del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión*, David Kaye. A/HRC/29/32. Consejo de Derechos Humanos, 2015. <https://docs.un.org/es/A/HRC/29/32>.

de fuentes abiertas (OSINT) a través de herramientas tecnológicas. En varios países (Colombia, Argentina, Bolivia, Uruguay, México y Brasil), la implementación de OSINT en contextos de seguridad ha sido llamada ciberpatrullaje, un concepto “cuyos límites y alcances no terminan de ser claramente delimitados. Las autoridades suelen sugerir que se trata simplemente del típico patrullaje de las calles pero trasladado a internet, cuando la verdad es mucho más compleja.”²⁰

La utilización de sistemas de inteligencia artificial para la investigación en fuentes abiertas supone un incremento considerable en las capacidades de vigilancia, lo que podría impactar negativamente en los derechos fundamentales de quienes son objetos de esta vigilancia. Particularmente por su alto nivel de intromisión y por sus características masivas antes que selectivas o específicas, incluso cuando el objetivo sea el seguimiento de una persona o grupo específico. En algunos países, como Argentina, se ha optado por normas expresas sobre el tema, que generan preocupaciones respecto a las facultades amplias que otorgan a los mecanismos de seguridad sin adecuar evaluaciones de derechos humanos.

En el año 2024, el Ministerio de Seguridad de Argentina autorizó vía Resolución N° 428/2024 a las fuerzas federales de policía y seguridad para realizar “tareas preventivas” en el ciberespacio. Un mes después, mediante Resolución N° 710/2024, se constituyó la Unidad de Inteligencia Artificial Aplicada a la Seguridad (UIAAS), orientada a prevenir, detectar, investigar y perseguir delitos —y sus vínculos— mediante IA. Dicha Unidad, liderada por la Dirección de Ciberdelito y Asuntos Cibernéticos, integrada por fuerzas federales, tiene funciones como patrullar redes sociales públicas, apps e internet; analizar en tiempo real imágenes de cámaras de videovigilancia para identificar sospechas o personas buscadas con reconocimiento facial; emplear algoritmos de machine learning sobre datos históricos delictivos para predecir crímenes futuros; y monitorear actividades en redes sociales para anticipar amenazas, rastrear grupos criminales o prevenir disturbios.

En Colombia, si bien la regulación es más difusa, algunas investigaciones han advertido sobre la adquisición e implementación de estos sistemas, sin que resulte claro el alcance de su uso por parte de las autoridades. En 2021, la Policía Nacional adquirió un “sistema de ciberinteligencia basado en inteligencia artificial”, desarrollado por la misma empresa que le

²⁰ Lucia Camacho. *Ciberpatrullaje en Argentina: Análisis de una Resolución problemática*. Centro de Estudios en Libertad de Expresión. s. f.
<https://observatoriolegislativocele.com/ciberpatrullaje-en-argentina-analisis-de-una-resolucion-problematica-por-lucia-camacho-g/>.

vendió años atrás al Ejército un sistema que fue utilizado para realizar vigilancia ilegal a periodistas, defensores de derechos humanos y políticos. En ese mismo año, la Policía reportó a la Comisión Interamericana de Derechos Humanos más de 21 mil horas de ciberpatrullaje facilitado por inteligencia artificial realizado en el marco de protestas sociales.

Los potenciales riesgos de esta actividad de vigilancia por parte del Estado no terminan en las vulneraciones de la privacidad o el efecto inhibitorio que pueden producir, sino que se extienden a lo que hace el Estado con los productos de inteligencia. Así, se generan dudas de para qué realizan estas actividades y qué hacen con la información que recopilan, creando alertas de, por ejemplo, la posibilidad de que estos datos se utilicen en perfilamientos ilegales. También en Colombia, se ha advertido la realización de actividades OSINT, sin un estricto cumplimiento del principio de legalidad del test tripartito del sistema interamericano. Esto ha tenido como consecuencia la criminalización de expresiones legítimas y la persecución de personas que ejercen su derecho a la protesta.²¹

Estos riesgos se agudizan en contextos sociales conflictivos, así lo evidencia una investigación de *Amnistía Internacional*²² que ha examinado documentación de los archivos públicos del Departamento de Seguridad Nacional de Estados Unidos, y documentos previamente develados sobre adquisiciones y evaluaciones de la privacidad, que demuestran la adquisición y uso de las herramientas de inteligencia artificial Babel X, proporcionada por Babel Street, e Immigration OS de Palantir. Estas aplicaciones cuentan con capacidades automatizadas que permiten el monitoreo, la vigilancia y la evaluación constantes de personas, a menudo de quienes no tienen la ciudadanía estadounidense y que han sido utilizadas en el contexto de las protestas a favor de Palestina.

Lo mismo sucede con el uso de sistemas de reconocimiento facial en contextos de seguridad. No hemos evidenciado prácticas por parte de los Estados direccionadas a medir el impacto en derechos humanos de estas tecnologías que se basan en inteligencia artificial, a pesar de que plantean riesgos sustanciales para el ejercicio de los derechos humanos, siendo el más

²¹ Lina Palacios. Se están metiendo al perfil de al lado: un análisis de la política criminal y el uso de tecnologías para judicializar personas durante el paro nacional de 2019 y 2021 en Colombia. Fundación Karisma, 2024. <https://biblio.karisma.org.co/book/209>

²² Amnistía Internacional. Estados Unidos/Global: La tecnología de Palantir y Babel Street amenaza con someter a vigilancia a manifestantes y migrantes en favor de Palestina. 2025 <https://www.amnesty.org/es/latest/news/2025/08/usa-global-tech-made-by-palantir-and-babel-street-pose-surveillance-threats-to-pro-palestine-student-protestors-migrants/>

advertido el derivado de los sesgos algorítmicos, que vulneran el principio de igualdad y perpetúan prácticas discriminatorias fundadas en categorías sospechosas.²³ Otro riesgo relevante atañe al derecho de reunión pacífica, dado el creciente empleo de esta tecnología en tales contextos, lo cual puede inducir un *chilling effect* entre los manifestantes.²⁴

Aun cuando la precisión tecnológica ha mejorado de manera notable en años recientes, persisten vulnerabilidades a errores. Por ejemplo, una imagen podría ser identificada erróneamente como coincidencia (*falso positivo*), con repercusiones graves para los derechos individuales, tales como la detención o procesamiento penal de personas inocentes. Incluso con tasas de error reducidas, su aplicación a grandes volúmenes poblacionales puede resultar en la identificación inexacta de cientos de individuos.²⁵

Por ejemplo, la app Mobile Fortify utilizada desde el año 2025 por el Department of Homeland Security de Estados Unidos, funciona con un “sistema que convierte imágenes faciales captadas en la calle en plantillas biométricas que se comparan con bases de datos gubernamentales, devolviendo una serie de “candidatos” que superan un umbral de similitud. Ese umbral puede ajustarse para priorizar velocidad y volumen, a costa de aumentar el riesgo de falsos positivos, especialmente cuando las imágenes se toman fuera de condiciones controladas”.²⁶

Esta tecnología habilita un régimen de vigilancia masiva y desproporcionada sobre la población, al facilitar la identificación de participantes en protestas y manifestaciones públicas. Ello genera un efecto inhibitorio sobre la libertad de expresión, especialmente en

²³ Clare Garvie Frankle Jonathan, «Facial-Recognition Software Might Have a Racial Bias Problem», The Atlantic, 7 de abril de 2016, <https://www.theatlantic.com/technology/archive/2016/04/the-underlying-bias-of-facial-recognition-systems/476991/>; David Leslie, «Understanding bias in facial recognition technologies» (Zenodo, 26 de septiembre de 2020), <https://doi.org/10.5281/ZENODO.4050457>; Sidney Perkowitz, «The Bias in the Machine: Facial Recognition Technology and Racial Disparities», MIT Case Studies in Social and Ethical Responsibilities of Computing, n.o Winter 2021 (5 de febrero de 2021), <https://doi.org/10.21428/2c646de5.62272586>.

²⁴ Alto Comisionado de las Naciones Unidas para los Derechos Humanos. Impacto de las nuevas tecnologías en la promoción y protección de los derechos humanos en el contexto de las reuniones, incluidas las protestas pacíficas. Consejo de Derechos Humanos, 2020. <https://docs.un.org/es/A/HRC/44/24>

²⁵ Íbidem.

²⁶ Nicholas Dale Leal. La ‘app’ de reconocimiento facial del ICE, en la mira por fallos y por amenazar un auge de la vigilancia estatal. El País, 2026. <https://elpais.com/us/migracion/2026-02-12/la-app-de-reconocimiento-facial-del-ice-en-la-mira-por-fallos-y-por-amenazar-un-auge-de-la-vigilancia-estatal.html>

ausencia de marcos normativos que protejan los derechos humanos, junto con mecanismos de control, rendición de cuentas y escrutinio independiente.

El uso excesivo de la vigilancia biométrica produce un efecto disuasorio inexorable: el temor a un monitoreo invasivo desincentiva la participación en actividades cívicas, vaciando los espacios públicos esenciales para la reunión, el diálogo, las protestas y el debate democrático.²⁷ En este sentido, la Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos ha afirmado que “el hecho de grabar, analizar y conservar las imágenes faciales de alguien sin su consentimiento constituye una injerencia en el ejercicio del derecho a la vida privada. Al desplegar la tecnología de reconocimiento facial en las reuniones y manifestaciones, la injerencia adquiere una escala enorme e indiscriminada, ya que se requiere la recopilación y el procesamiento de imágenes faciales de todas las personas captadas por una cámara equipada con un sistema de tecnología de reconocimiento facial o conectada a un sistema de este tipo”.²⁸

Las grandes capacidades de los sistemas biométricos conllevan riesgos como el chilling effect anteriormente mencionado y se extienden a otros, como las prácticas abusivas de perfilamientos, que tienden a basarse en sesgos discriminatorios, y que preocupan por darse en escenarios de poca certeza respecto a para qué son usadas. Regresando al ejemplo de la app Mobile Fortify usada por el DHS, las preocupaciones en materia de derechos humanos de este tipo de tecnologías aumentan en contextos de persecución a poblaciones históricamente marginalizadas.

A lo anterior se suma un problema grave de transparencia:²⁹ la adopción de tecnologías de inteligencia artificial por parte de los Estados con fines de seguridad pública suele enmarcarse en regímenes de reserva legal, vinculados a materias de orden público, lo que dificulta el

²⁷ United Nations Special Rapporteur on the promotion and protection of human rights and fundamental freedoms while countering terrorism. Civil Society Under Attack A global study in defense of civic space. 2024. <https://defendcivicspace.com/>

²⁸ Alto Comisionado de las Naciones Unidas y para los Derechos Humanos. Impacto de las nuevas tecnologías en la promoción y protección de los derechos humanos en el contexto de las reuniones, incluidas las protestas pacíficas. Consejo de Derechos Humanos, 2020. <https://docs.un.org/es/A/HRC/44/24>

²⁹ Grupo de trabajo: Transparencia y derecho humanos en las, políticas en torno a las tecnologías de vigilancia, y políticas en torno a las tecnologías de vigilancia. «Transparencia y derecho humanos en las políticas en torno a las tecnologías de vigilancia». s. f. <https://www.derechosdigitales.org/wp-content/uploads/Recomendaciones-para-la-transparencia-y-anticorrupcion-en-la-compra-y-uso-de-tecnologias-de-vigilancia-por-parte-de-los-Estados-americanos.pdf>

seguimiento de las capacidades reales de vigilancia estatal por la escasa disponibilidad de información.

6. Lagunas, buenas prácticas y recomendaciones

- a. ¿Qué cambios jurídicos y/o políticos han introducido los Estados de la región para hacer frente a un contexto donde la inteligencia artificial y la inteligencia artificial generativa juega un papel cada vez más prominente en la vida cotidiana de las personas?

Actualmente, en la región algunos estados (Brasil, Argentina, Chile, Colombia, México y Uruguay) han avanzado en la confección de lineamientos, estándares y estrategias nacionales en materia de inteligencia artificial. A la fecha, Perú es el único país que cuenta con una ley aprobada específicamente dedicada a la materia. Otras propuestas regulatorias actualmente en discusión abordan aspectos específicos derivados del uso e impacto de la IA como, por ejemplo, los *deepfakes*. En este sentido, consideramos importante señalar que las respuestas regulatorias orientadas a imponer prohibiciones estrictas a estos sistemas pueden resultar problemáticas y potencialmente restrictivas de la libertad de expresión por, al menos, dos razones.

En primer lugar, porque pueden arrasar de manera indiscriminada con usos legítimos y socialmente valiosos de estas tecnologías, al no distinguir adecuadamente entre prácticas dañinas y expresiones creativas, críticas o satíricas. En segundo lugar, porque corren el riesgo de convertirse en herramientas funcionales para regímenes autoritarios, que podrían instrumentalizar la lucha contra las “falsedades” para perseguir la disidencia, restringir el debate público y consolidar mecanismos de censura bajo la apariencia de protección de la verdad. Un principio metodológico recomendable debería consistir en examinar, antes de promover nuevas regulaciones específicas, qué herramientas jurídicas ya existen en el ordenamiento vigente y en qué medida pueden dar respuestas adecuadas —o ser razonablemente adaptadas— a los desafíos y oportunidades que introduce la tecnología.³⁰

³⁰ Vladimir Garay and Matias Gonzalez Mama, “¿Prohibir Lo Falso?: Deepfakes y Libertad de Expresión,” La Silla Vacía, June 14, 2025, <https://www.lasillavacia.com/red-de-expertos/red-de-democracia-y-tecnologia/prohibir-lo-falso-deepfakes-y-libertad-de-expresion/>.

En el contexto global existen diversos lineamientos desarrollados en el marco de instancias multinacionales, tanto de carácter vinculante como no vinculante, siendo estos últimos los más numerosos. Ambos pueden tener efectos modelizadores sobre las regulaciones locales en desarrollo. Dentro del grupo de los vinculantes, los más relevantes son la AI Act de la Unión Europea y el Convenio Marco del Consejo de Europa sobre IA, Derechos Humanos, Democracia y Estado de Derecho. En el caso de los no vinculantes, están las recomendaciones surgidas en el marco de la OCDE, incluyendo el Proceso de Hiroshima del G7 y la Global Partnership on AI, el Global Digital Compact, el proceso de la UNESCO sobre ética de la IA y cabe mencionar también la primera resolución de la Asamblea General sobre IA.³¹

En términos generales, las iniciativas y regulaciones existentes revelan una tendencia clara hacia la adopción de enfoques basados en riesgos para la regulación de la inteligencia artificial; cuando son mencionados, los derechos humanos son relegados a un segundo plano y las referencias a la protección de la libertad de expresión son escasas, cuestión que fue ya advertida en su momento por el Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión de Naciones Unidas en 2018.³²

El riesgo que encierran las regulaciones basadas en riesgos es que los Estados, a través de la regulación, deleguen en actores privados la responsabilidad de identificar, valorar y mitigar dichos riesgos sin establecer las salvaguardas adecuadas. En ese sentido es importante incorporar obligaciones específicas destinadas a garantizar el respeto y protección de los derechos humanos a lo largo del ciclo de vida de los sistemas de IA, alineadas con las normas y estándares del derecho internacional de los derechos humanos.

7. ¿Desea añadir alguna información o comentario adicional sobre el tema de este cuestionario?

Desde el Observatorio Legislativo del Centro de Estudios en Libertad de Expresión hemos realizado el rastreo de propuestas regulatorias sobre inteligencia artificial promovidas en los

³¹ Matias Gonzalez Mama et al., “Global AI Governance: Who Regulates, with What Approach, and for Whom?,” *SSRN Electronic Journal* CELE Research

³² Asamblea General de las Naciones Unidas. A/73/348. Informe del Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, párr. 43. Disponible en: <https://documents.un.org/doc/undoc/gen/n18/270/45/pdf/n1827045.pdf>

congresos y asambleas legislativas de 12 países de la región: Argentina, Bolivia, Brasil, Chile, Colombia, Costa Rica, Ecuador, Guatemala, México, Paraguay, Perú y Uruguay, desde el año 2023. La matriz puede consultarse [aquí](#).

Este monitoreo nos ha permitido recolectar información de 194 proyectos de ley y 1 ley sancionada en 10 de los 12 países observados. Además de recopilar las propuestas en una base de datos descargable, las hemos clasificado de acuerdo con algunos elementos valiosos para analizar el posible impacto de las mismas en materia de libertad de expresión.

Cada proyecto está clasificado de acuerdo con las temáticas que aborda. En la matriz descargable están disponibles tres columnas de tema, siendo la primera el tema central y las otras los temas secundarios; las gráficas de temática expresan los resultados del tema central. Esta clasificación es relevante porque permite interpretar cuáles son las tendencias legislativas respecto a los sistemas basados en inteligencia artificial, da lugar a una mayor comprensión de los enfoques adoptados y también permite tener bases sólidas para analizar los cambios que estas regulaciones pueden generar en las dinámicas de gobernanza de internet, tales como el régimen de responsabilidad de intermediarios. Sugerimos el análisis de datos de manera integral y teniendo en cuenta la pluralidad temática de las propuestas.

Además, hemos destacado si los proyectos modifican alguna ley, si incluyen sanciones (en caso afirmativo, qué tipo de sanción) y si las disposiciones normativas cumplen con el test tripartito. Por último, tuvimos en cuenta si el texto presentado podría influir en el panorama de moderación de contenidos, por ejemplo, incluyendo obligaciones a los intermediarios.

Asimismo, aprovechamos esta oportunidad para compartir con la RELE los siguientes documentos en los que el CELE ha desarrollado su posición respecto de los sistemas de inteligencia artificial:

Referencias

Global AI Governance: Who Regulates, with what Approach, and for Whom?, disponible [aquí](#).

CELE submission to the Committee of Ministers' public consultation on the Draft Guidance Note on the Implications of Generative Artificial Intelligence for Freedom of Expression, disponible [aquí](#).

Public Consultation on Freedom of Expression and Artificial Intelligence, disponible [aquí](#).

Beyond Ethics: Why a Rights-Based Framework is Essential for AI Governance, disponible [aquí](#).