

Cuestionario de consulta sobre el impacto de la Inteligencia Artificial en los Derechos Humanos en las Américas, Relatoría Especial para la Libertad de Expresión (RELE) de la Organización de Estados Americanos OEA.

Marzo 02, 2026

Derechos Digitales¹ es una organización regional latinoamericana sin fines de lucro fundada en 2005, que se dedica a la defensa y promoción de derechos humanos en el entorno digital para contribuir con sociedades más justas, inclusivas e igualitarias. Sus acciones combinan investigación, incidencia en políticas públicas y privadas, análisis de tecnologías, campañas y formación en los derechos digitales y la seguridad digital, entre otros.

Aportes a la consulta

1. Uso de inteligencia artificial por parte de Estados de la región. (Respuesta conjunta a punto A y punto B)

Punto A. ¿Qué instituciones públicas utilizan actualmente sistemas de inteligencia artificial, incluyendo inteligencia artificial generativa, para apoyar u adoptar decisiones que afectan derechos humanos? Indique tipo de sistema, finalidad y contexto de su uso (por ejemplo: selección de beneficiarios de programas sociales, clasificación de riesgo en el sistema penal, priorización de pacientes en salud, migración, educación, vigilancia, moderación de contenidos, etc.

Punto B. ¿Cuáles son los beneficios identificados por los Estados en el uso de estos sistemas y cuáles son los principales riesgos para los derechos humanos?

El uso de la inteligencia artificial (IA)² en el sector público registra un ritmo acelerado en América Latina y el mundo. Bajo el discurso de optimizar y hacer más eficientes los procesos de las entidades estatales, los gobiernos han impulsado la adopción de sistemas basados en sistemas de IA que apoyan la gestión y toma de decisiones sobre sectores esenciales como el de la salud, justicia, educación, el empleo, la protección de las niñeces, el acceso a programas de bienestar, la seguridad ciudadana, entre otros.

En el 2019, en el marco del eje programático “Inteligencia Artificial e Inclusión en América Latina”³ Derechos Digitales desarrolló un conjunto de estudios de caso sobre el uso de sistemas de IA por parte del sector público en la región. Allí, se ahondó en las implicancias para el ejercicio de derechos fundamentales del uso de esa tecnología por parte de los Estados. A

¹ Para más información, ver: <https://www.derechosdigitales.org/>. Esta contribución fue preparada por Laura Mantilla, Lucía Camacho, y revisada por Paloma Lara-Castro. Para más información, contactar: paloma.lara.castro@derechosdigitales.org

² El término Inteligencia Artificial hará referencia a las diferentes técnicas específicas que abarca el paraguas conceptual de la IA, tales como los modelos de procesamiento de lenguaje natural (NLP), aprendizaje automático (ML), sistemas predictivos de riesgo, y procesos de decisión automatizada (ADM).

³ Para más información, consultar: <https://ia.derechosdigitales.org/publicaciones/>

continuación, se presentan los hallazgos por país, sector de implementación, tipo de sistema, su objetivo y los impactos que evidenciamos para el ejercicio de los derechos humanos.

País: México.

Sistema: Sistema de Actuación Temprana para la Permanencia Escolar (SATPE)⁴.

Institución pública responsable: Secretaría de Educación del Estado de Guanajuato⁵.

Derechos impactados: Derechos a la educación, la privacidad y protección de datos.

Objetivo/propósito de la herramienta: El SATPE busca identificar el riesgo de abandono escolar en la educación media a través del despliegue de un sistema de predicción, alerta e intervención temprana.

Beneficios identificados por el sector público: Frente a la problemática socio-educativa del abandono escolar del Sistema Educativo de México, STAPE hace parte de las políticas educativas dirigidas hacia la prevención del abandono en los niveles de secundaria y medio-superior. El resultado esperado de este algoritmo de aprendizaje es generar una lista de estudiantes clasificados a partir de su riesgo de abandono escolar, frente a los cuales las escuelas deberían desplegar estrategias de retención.

Impactos/resultados evidenciados: La investigación de Derechos Digitales⁶ encontró que, aunque el STAPE ha aumentado la conciencia entre personal educativo sobre el problema del abandono escolar, no ha contribuido a reducirlo.

De otra parte, El STAPE genera alertas en materia de protección de datos personales al integrar información del Sistema de Control Escolar⁷, el Catálogo de Escuelas Oficial del Estado de Guanajuato (CEO)⁸, y de la Recopilación de Información para la Mejora de los Aprendizajes (RIMA)⁹. Si bien la Secretaría de Educación de Guanajuato habilitó un aviso de privacidad simplificado que informa finalidades y obtiene consentimiento tácito o expreso para recabar información de la comunidad educativa, dicho aviso presenta un alcance limitado. Este se limita a señalar el uso de los datos para efectos operativos, estadísticos e integración de bases de datos, sin incorporar disposiciones específicas para el STAPE ni establecer salvaguardas explícitas frente a posibles usos secundarios o discriminatorios de la información recolectada.

⁴ Ricaurte, P., & Nájera, J. (2024). SATPE: análisis del sistema predictivo para la prevención del abandono escolar del estado de Guanajuato. Derechos Digitales. Disponible en: <https://ia.derechosdigitales.org/wp-content/uploads/2024/11/2024-MEX-SATPE-ESP.pdf>

⁵ En cooperación con el Banco Mundial, y con la participación de actores como Microsoft, USAID y la empresa privada Nutanix.

⁶ Ricaurte, P., & Nájera, J. (2024). SATPE: análisis del sistema predictivo para la prevención del abandono escolar del estado de Guanajuato. Derechos Digitales. Disponible en: <https://ia.derechosdigitales.org/wp-content/uploads/2024/11/2024-MEX-SATPE-ESP.pdf>

⁷ Sistema que recopila información sobre la matrícula, asistencia y rendimiento académico de los estudiantes en las escuelas públicas.

⁸ El cual proporciona información sobre las escuelas en el Estado de Guanajuato.

⁹ El cual contiene indicadores de aprendizaje.

País: Colombia.

Sistema: Software Watson Explorer (Fiscal Watson, de propiedad de IBM).¹⁰

Institución pública responsable: Fiscalía General de la Nación.

Derechos impactados: Derechos al debido proceso, un juicio justo, la privacidad y protección de datos.

Objetivo/propósito de la herramienta: El uso del sistema Fiscal Watson promete apoyar la gestión de información del Sistema Penal Oral Acusatorio (SPOA). Mediante Procesamiento de Lenguaje Natural avanzado y Machine Learning, el sistema permite asociar denuncias con características cuantitativas, cualitativas y territoriales similares agilizando la fase de indagación y facilitando la toma de decisiones de las personas funcionarias judiciales¹¹.

Beneficios identificados por el sector público: Este sistema de IA se compró y desplegó en el marco de estrategias nacionales tecno-optimistas orientadas hacia la eficiencia y mejoría de la gestión pública. Aunado a esto, la llamada crisis de la justicia en Colombia caracterizada por la congestión judicial, la impunidad, la tipificación errónea de los delitos, los reprocesos en las investigaciones, entre otros, hacen del sector judicial uno atractivo para el uso de la IA. Sin embargo, aun cuando el uso de Fiscal Watson apoye la fase de indagación, la labor de las personas investigadoras no puede agotarse en filtrar las búsquedas en la herramienta tecnológica, la información debe ser verificada y deben agotarse los demás medios investigativos¹².

Impactos/resultados evidenciados: El despliegue del Fiscal Watson ha evidenciado problemas de transparencia significativos. En primer lugar, se desconoce el funcionamiento interno del sistema, principalmente por motivos asociados a la protección de la propiedad intelectual de su propietario¹³. En ese sentido, no se conoce con exactitud cuáles son las fuentes de datos que indexa Watson, abriendo interrogantes por la calidad de la información con que funciona el sistema. La falta de transparencia aludida puede dar lugar a decisiones erróneas y sesgadas que produzcan consecuencias jurídicas o fácticas adversas para las personas¹⁴. En segundo lugar, el despliegue de esta tecnología no cuenta con mecanismos de monitoreo y evaluación de la eficiencia de la herramienta, así como tampoco se publica de manera pública, abierta y accesible información sobre su desempeño y efectividad, lo que dificulta establecer sus porcentajes de éxito en el análisis de información y si ha logrado el objetivo bajo el cual fue desplegado.

¹⁰ Palacios, L., Forero, V., & Labarthe, S. (2024). Fiscal Watson: estudio del uso de Inteligencia Artificial en la Fiscalía General de la Nación en Colombia. Derechos Digitales. Disponible en: https://ia.derechosdigitales.org/wp-content/uploads/2024/11/2024-COL-Fiscal_Watson-ESP.pdf

¹¹ Ibid.

¹² Para más información, consultar: <https://www.corteconstitucional.gov.co/relatoria/2024/T-323-24.htm>

¹³ Desarrollado por la empresa estadounidense International Business Machines (IBM).

¹⁴ Mantilla-León, L. (2025). El uso de IA en el sistema judicial colombiano: una investigación sobre el Fiscal Watson. Ámbito Jurídico. Disponible en: <https://www.ambitojuridico.com/noticias/penal/el-uso-de-ia-en-el-sistema-judicial-colombiano-una-investigacion-sobre-el-fiscal>

País: Colombia.

Sistema: PretorIA.¹⁵

Institución pública responsable: Corte Constitucional de Colombia.

Derechos impactados: Derecho de acceso a la información y transparencia, derecho de acceso a la justicia.

Objetivo/propósito de la herramienta: Su propósito es hacer más eficiente el proceso de selección de los casos de tutela judicial de los derechos fundamentales, a partir de los cuales esta Corte fija su jurisprudencia. El sistema trabaja con textos jurídicos en español y proporciona información sobre el contenido de las sentencias, así como datos estadísticos generales. PretorIA no tiene capacidad para tomar decisiones judiciales.

Beneficios identificados por el sector público: PretorIA se inserta en Colombia en el marco de las iniciativas dirigidas a incrementar el uso de tecnologías digitales en la Rama Judicial, tendencia incrementada por la pandemia del Covid-19. Este sistema permite presentar la información en fichas de resumen que indican la concurrencia o no de las categorías en un texto, gracias a un buscador que combina las palabras y categorías, y además permite la elaboración de estadísticas y la identificación de temas recurrentes. Todo ello permite hacer más eficiente el proceso de identificación y de preselección de los casos, y generar información sobre el comportamiento general de la jurisdicción constitucional de los derechos fundamentales en Colombia.

Impactos/resultados evidenciados: La investigación¹⁶ no encontró posibles afectaciones del sistema a derechos en términos de privacidad y protección de datos personales, dado que el tratamiento de información personal no constituye el núcleo de sus operaciones y el sistema tampoco tiene funciones predictivas aplicables a la fase de selección de tutelas. No obstante, la Corte Constitucional ha realizado una difusión limitada sobre el funcionamiento de PretorIA. Una mayor transparencia al respecto sería deseable desde la perspectiva del derecho de acceso a la información pública, en tanto permitiría acercar el proceso de selección a la ciudadanía, así como revelar el pensamiento jurídico de la Corte.

País: Brasil.

Sistema: Sistema Nacional de Empleo (SINE).¹⁷

Institución pública responsable: Secretaría Especial de Productividad, Empleo y Competitividad (SEPEC) del Ministerio de Economía.¹⁸

¹⁵ Saavedra, V.; Upegui, J.C. (2021). PretorIA y la automatización del procesamiento de causas de derechos humanos. Derechos Digitales. Disponible en: https://ia.derechosdigitales.org/wp-content/uploads/2021/03/CPC_informe_Colombia.pdf

¹⁶ Ibid.

¹⁷ Bruno, F.; Cardoso, P.; Faltay, P. (2021). Sistema Nacional de Empleo y la gestión automatizada de la desocupación laboral. Derechos Digitales. Disponible en: https://www.derechosdigitales.org/wp-content/uploads/CPC_informe_BRASIL_castellano.pdf

¹⁸ El despliegue de herramientas de IA en el SINE fue posible a través de un acuerdo de cooperación técnica entre el gobierno brasileño de Jair Bolsonaro y Microsoft en 2020.

Derechos impactados: derechos al trabajo, la igualdad, el acceso a la información, a la explicación de decisiones automatizadas.¹⁹

Objetivo/propósito de la herramienta: Con el SINE se busca facilitar el proceso de intermediación de la mano de obra a través de dos funciones. Primero, la utilización de IA para cruzar las vacantes registradas en el portal de empleo con los perfiles de las personas trabajadoras considerados como los más adecuados para desempeñarlas. Segundo, el perfilamiento de las personas trabajadoras registradas segmentándolas según sus capacidades y apoyando tanto el proceso interno de recomendación de vacantes, como el desarrollo de políticas públicas de empleo, en especial de cualificación profesional.

Beneficios identificados por el sector público: El SINE promete la reducción del tiempo de búsqueda de empleo con el fin de insertar nuevamente a las personas trabajadoras en el mercado laboral, reduciendo los costos y el tiempo de espera para trabajadores y empleadores. Además, abarca acciones y programas que tienen como objetivo la cualificación de las personas trabajadoras, priorizando a aquellas en situaciones de mayor vulnerabilidad social. Con el SINE se percibe mayor eficacia en la asignación de personas trabajadoras a actividades productivas, buscando la inclusión social a través del empleo.

Impactos/resultados evidenciados: La investigación²⁰ detalla impactos en derechos humanos asociados a: el riesgo de exclusión del público objetivo, bajo el cual el sistema puede terminar perjudicando a las personas trabajadoras que no logran ocuparse; la asimetría de poder y la opacidad de la herramienta, puesto que se evidencia una profunda desigualdad entre la ciudadanía y el sistema, ya que las personas trabajadoras no parecen comprender las reglas de funcionamiento del sistema, ni pueden impugnar los resultados de la intermediación o los perfiles asignados; y los riesgos asociados a la posibilidad de influencia de Microsoft sobre los criterios y diseños de las políticas públicas de empleo, así como la cesión de datos estratégicos sobre el mercado laboral brasileño, todo ello sin mecanismos sólidos de rendición de cuentas²¹.

País: Brasil.

Sistema: Auxilio de Emergencia (AE).²²

Institución pública responsable: El Ministerio de la Ciudadanía (MC), la Caja Económica Federal (CEF) y la empresa de Tecnología e Información de Pensiones (Dataprev).

Derechos impactados: Derechos a la alimentación, la igualdad, privacidad y protección de datos.

Objetivo/propósito de la herramienta: Su objetivo fue apaciguar los efectos económicos y sociales de la pandemia de COVID-19, y permitir a la población en situación de vulnerabilidad la

¹⁹ Tal como se incorporó a la Ley General de Protección de Datos en Brasil.

²⁰ Bruno, F.; Cardoso, P.; Faltay, P. (2021). Sistema Nacional de Empleo y la gestión automatizada de la desocupación laboral. Derechos Digitales. Disponible en: https://www.derechosdigitales.org/wp-content/uploads/CPC_informe_BRASIL_castellano.pdf

²¹ Ibid.

²² Tavares, C., Fonteles, J., Simao, B., Valente, M. (2022). El Auxilio de Emergencia en Brasil: Desafíos en la implementación de una política de protección social datificada. Derechos Digitales. Disponible en: <https://ia.derechosdigitales.org/wp-content/uploads/2022/06/IA-and-Inclusion-Brasil-2-ES.pdf>

permanencia del acceso a bienes de consumo, como la alimentación. La selección de personas beneficiarias es automatizada y ejecutada por Dataprev, mediante el cruce de múltiples bases de datos gubernamentales²³.

Beneficios identificados por el sector público: El AE se consolidó como el programa de transferencia de ingresos de mayor magnitud en la historia de Brasil, llegando a cubrir a 67,8 millones de personas. Su arquitectura institucional, basada en la digitalización y el uso intensivo de datos, facilitó la celeridad en la concesión de los beneficios, aunque simultáneamente generó barreras para sectores en situación de vulnerabilidad social y económica. Pese a estas limitaciones y las dificultades de acceso derivadas de la automatización, el modelo implementado permitió una cobertura masiva rápidamente.²⁴

Impactos/resultados evidenciados: El Auxilio de Emergencia está compuesto por un intenso flujo de datos, que atraviesa todas las etapas del programa. La selección de personas beneficiarias es automatizada, realizada por Dataprev, que cruza múltiples bases de datos, de diferentes órganos del Gobierno. Bajo estas condiciones, es preocupante que el algoritmo de AE automatizó el proceso de asignación de beneficios, sin existir supervisión humana. Ante la ausencia de mecanismos administrativos para apelar, la vía judicial sirvió a las personas beneficiarias excluidas para tramitar sus reclamos.

De otra parte, la exigencia del AE de confirmar la identidad mediante SMS u otro medio electrónico evidenció una profunda desconexión entre la herramienta y la brecha digital del país, al no considerar las limitaciones de conectividad y acceso a redes móviles que afectan a pueblos indígenas, comunidades quilombolas y tradicionales en determinados estados brasileños.

País: Paraguay.

Sistema: Portal EmpleaPY.²⁵

Institución pública responsable: Ministerio de Trabajo, Empleo y Seguridad Social (MTESS).²⁶

Derechos impactados: Derechos a la privacidad y la protección de datos.

Objetivo/propósito de la herramienta: EmpleaPy fue creado con el fin de aumentar la empleabilidad mediante la automatización de procesos. Su objetivo principal es fortalecer el vínculo entre el gobierno, las empresas y las personas interesadas en encontrar empleo, generando oportunidades laborales para la ciudadanía. El sistema facilita la gestión de

²³ Como la base de datos CadÚnico.

²⁴ Según la investigación de Derechos Digitales, la Caja Económica Federal lanzó la aplicación para el registro para el Auxilio de Emergencia el 7 de abril de 2020. *Después de 24 horas del lanzamiento de la aplicación, más de 22 millones de personas se habían registrado para la solicitud del beneficio. Los pagos comenzaron a realizarse el 9 de abril de 2020.* Ver: Tavares, C., Fonteles, J., Simao, B., Valente, M. (2022). El Auxilio de Emergencia en Brasil: Desafíos en la implementación de una política de protección social datificada. Derechos Digitales. Disponible en: <https://ia.derechosdigitales.org/wp-content/uploads/2022/06/IA-and-Inclusion-Brasil-2-ES.pdf>

²⁵ Sequera, M., Cuevas, M. (2024). EmpleaPY: análisis sobre la automatización de procesos para las políticas de empleo en el estado paraguayo. Derechos Digitales. Disponible en: <https://ia.derechosdigitales.org/wp-content/uploads/2024/11/2024-PAR-EmpleaPY-ESP.pdf>

²⁶ En cooperación técnica del Banco Interamericano de Desarrollo BID

consultas y verificación de perfiles de personas candidatas a vacantes de empleo, dejando en un proceso manual de revisión humana la vinculación de un candidato con una vacante, o las coincidencias entre ambos.

Beneficios identificados por el sector público: Al momento de la investigación, EmpleaPy había registrado a unas 64.000 personas interesadas, generando más de 16.097 intermediaciones laborales durante los primeros 6 meses.

Impactos/resultados evidenciados: Según el caso de estudio²⁷, la herramienta se desplegó sin análisis previos de protección a la privacidad de las personas trabajadoras, de hecho, la plataforma no cuenta con políticas de términos y condiciones sobre datos personales y privacidad para garantizar la protección de la información sensible de las personas usuarias. Además, el despliegue del sistema no ha previsto la realización de estudios de impacto que sean de carácter público.

País: Chile

Sistema: Sistema Alerta Niñez²⁸

Institución pública responsable: Subsecretaría de Evaluación Social y la Subsecretaría de Niñez, dependientes del Ministerio de Desarrollo Social y Familia de Chile

Derechos impactados: Derechos a la privacidad y la protección de datos, al acceso a la información

Objetivo/propósito de la herramienta: Este sistema, “piloto” al momento de la investigación, busca estimar y predecir el nivel de riesgo de las y los niñas, niños y adolescentes (NNA) de sufrir alguna vulneración en sus derechos. Mediante un análisis de datos con diferentes modelos algorítmicos, pretende anticipar de forma preventiva cada caso. El sistema genera un “índice de riesgo” para cada NNA que permite clasificar y priorizar los casos para las Oficinas Locales de Niñez.²⁹

Beneficios identificados por el sector público: Según la Subsecretaría de Niñez, el sistema se constituye en una plataforma de registro, gestión y monitoreo de los casos de NNA identificados con mayor riesgo. En ese escenario, promueve el ejercicio de derechos de NNA previniendo situaciones de vulneración, para favorecer su desarrollo integral, apoyando así la gestión del acceso de los NNA y sus familias a servicios territoriales. Como modelo predictivo, el sistema realiza una clasificación inicial según criterios de priorización para establecer un orden en el que se atenderán los casos.

Impactos/resultados evidenciados: La herramienta se entrenó con datos provenientes de diversas fuentes, incluyendo SENAME, Chile Crece Contigo, el Ministerio de Educación

²⁷ Sequera, M., Cuevas, M. (2024). EmpleaPY: análisis sobre la automatización de procesos para las políticas de empleo en el estado paraguayo. Derechos Digitales. Disponible en:

<https://ia.derechosdigitales.org/wp-content/uploads/2024/11/2024-PAR-EmpleaPY-ESP.pdf>

²⁸ Valderrama, M (2021). Sistema Alerta Niñez y la predicción del riesgo de vulneración de derechos de la infancia. Derechos Digitales. Disponible en:

https://ia.derechosdigitales.org/wp-content/uploads/2021/03/CPC_informe_Chile.pdf

²⁹ Ibid.

(matrícula y rendimiento escolar en colegios públicos y privados), el Registro Social de Hogares y datos censales sobre vulnerabilidad barrial y estadísticas de delitos por barrios. La investigación³⁰ encontró cómo en el proceso de licitación del sistema no se incluyeron criterios básicos de transparencia, justicia o gobernanza de datos, o rendición de cuentas. Por ejemplo, las autoridades argumentan que, en tanto se encuentra en una fase “piloto”, no pueden entregar información de la herramienta en cuestión. Esta característica dificulta la evaluación de su desempeño en un contexto donde preocupa la ausencia de documentación pública sobre el funcionamiento, rendimientos y efectos de los modelos predictivos del Sistema Alerta Niña. También se ha señalado la ausencia de procesos participativos para influenciar el diseño y despliegue del piloto.

País: Chile.

Sistema: Sistema Predictivo del Delito Urbano.³¹

Institución pública responsable: Departamento de Análisis Criminal (DAC) de Carabineros de Chile.

Derechos impactados: Derechos a la privacidad, la igualdad, al debido proceso, la libertad personal, acceso a la información.

Objetivo/propósito de la herramienta: El sistema se desplegó con el objetivo de predicción de ocurrencia de delitos y la finalidad de dirigir el patrullaje policial preventivo en las ciudades, definiendo áreas de mayor vigilancia y control. Al momento de la investigación, el sistema se había desplegado en cincuenta y ocho comisarías del país.

Beneficios identificados por el sector público: El sistema se despliega bajo el discurso de seguridad preventiva y busca contribuir a disminuir la sensación de inseguridad entre la población, a pesar de ser Chile uno de los países con bajas tasas de victimización y violencia en comparación con la región de América Latina. Evaluar los beneficios del sistema resulta una tarea difícil en tanto que solo existen auditorías internas que evalúan la funcionalidad técnica del sistema según niveles de predicción predefinidos. No hay mecanismos claros para detectar fallas en el sistema ni para evaluar la utilidad real de la tecnología.

Impactos/resultados evidenciados: El estudio³² identificó diversas falencias que afectan la calidad y representatividad de la información utilizada por el sistema. Por un lado, se detectaron deficiencias en aspectos clave de la gobernanza de datos, como su almacenamiento y protocolos de seguridad, mismas que fueron señaladas en un informe de la Contraloría General de la República. Por otro lado, se observaron problemas en los métodos de recolección de información, los cuales limitan la capacidad de los datos para reflejar la complejidad del fenómeno delictivo.

En algunos casos, estos sesgos contribuyen a sobredimensionar la incidencia delictiva en determinados sectores urbanos, lo que refuerza prácticas de vigilancia y estigmatiza a ciertas

³⁰ Ibid.

³¹ Buschman, J. (2021). Sistema Predictivo del Delito Urbano: Producción algorítmica de Zonas de Vigilancia y Control en la Ciudad. Derechos Digitales. Disponible en: <https://ia.derechosdigitales.org/wp-content/uploads/2022/05/CPC-IA-Chile-2-ES.pdf>

³² Ibid.

comunidades al asociarlas de manera prejuiciosa con la criminalidad. A ello se suma la ausencia de instancias externas de auditoría que aseguren la transparencia y el control público del sistema, así como la falta de mecanismos para difundir activamente información sobre su funcionamiento y resultados. Con todo, la herramienta ha sido desplegada bajo una noción de seguridad en la que prima la vigilancia por sobre el cuidado y respeto irrestricto a los derechos humanos

País: Uruguay

Sistema: CoronavirusUY.³³

Institución pública responsable: Ministerio de Salud Pública de Uruguay.

Derechos impactados: Derechos a la salud, la igualdad.

Objetivo/propósito de la herramienta: Mediante una aplicación de teléfono móvil, el sistema fue diseñado con el propósito de ofrecer información pública sobre estadísticas de contagio del COVID-19 y las medidas sanitarias consecuentes. CoronavirusUY se desplegó para reunir información de autodiagnóstico de contagio del virus, proveer asistencia médica remota, y alertar a las personas usuarias sobre posibles contactos con personas infectadas.

Beneficios identificados por el sector público: Este sistema recolectó y dispuso información de manera centralizada para dirigir acciones estatales tanto a nivel general como respecto de casos individuales en el contexto de la pandemia por COVID-19, facilitando desde la diseminación de recomendaciones de cuidado, hasta la atención vía telemedicina.

Impactos/resultados evidenciados: El impacto del sistema se vio marcado por la brecha digital existente en ese país, que restringe su alcance y excluye ciertos sectores del actuar estatal. Además, su implementación acelerada excluyó la participación ciudadana y cualquier evaluación formal de sus efectos. En esa línea, el sistema carece de un monitoreo continuo e institucional con participación pública. También es relevante que, paralelamente, se aprobaron leyes como la de Telemedicina y excepciones al consentimiento para datos de salud. Si bien la rápida respuesta del país ante la pandemia puede verse como un acierto, esto solo es válido si se fundamenta en protecciones previas sólidas en cuanto a los derechos humanos de la ciudadanía.

País: Argentina.

Sistema: Chatbot “Boti”.³⁴

Institución pública responsable: Gobierno de la Ciudad de Buenos Aires (GCBA).

Derechos impactados: Derecho a la privacidad, acceso a la información pública y protección de datos.

³³ Yael, D. (2021). CoronavirusUY y la tecnología como solución a la pandemia. Derechos Digitales. Disponible en: https://ia.derechosdigitales.org/wp-content/uploads/2021/03/CPC_informe_URUGUAY.pdf

³⁴ Ferreyra, E. (2024). Boti: estudio sobre el chatbot con procesamiento del lenguaje natural del Gobierno de la Ciudad de Buenos Aires. Derechos Digitales. Disponible en: <https://ia.derechosdigitales.org/wp-content/uploads/2024/11/2024-ARG-BOTI-ESP.pdf>

Objetivo/propósito de la herramienta: Se trata de un asistente virtual que permite a los ciudadanos interactuar y obtener información a través de WhatsApp. Ofrece respuestas sobre una amplia variedad de temas, incluyendo servicios gubernamentales, salud y movilidad urbana

Beneficios identificados por el sector público: Según el estudio, durante la pandemia, el Chatbot Boti se convirtió en la principal fuente de información oficial sobre síntomas, prevención, gestión de datos de casos de COVID-19, resultados de pruebas y posteriormente información sobre el plan de vacunación. Adicionalmente, se encontró que la Secretaría de Innovación y Transformación Digital desarrolló e implementó un sistema de inteligencia artificial denominado IATos, que complementó la detección de COVID-19 mediante el Chatbot Boti. La IA permitía clasificar sonidos de voz, respiración y tos a partir de audios enviados por WhatsApp, logrando un 86% de efectividad en la predicción de casos sospechosos.

Impactos/resultados evidenciados: Según la investigación³⁵, el despliegue de Boti estuvo acompañado de preocupaciones frente al tratamiento de datos personales de la herramienta. Por un lado, los fines que cumple la recolección de datos inicial son vagos y poco claros; no es claro si habrá usos secundarios de la información recolectada, y tampoco se explicitan los mecanismos de anonimización de la información. Asimismo, no hay mecanismos para aquellas personas que no deseen prestar su consentimiento para la entrega y recolección de sus datos. Por último, auditorías de la autoridad de protección de datos del GCBA detallaron la ausencia de responsables de los sistemas de información, así como la falta de evidencia de procedimientos dirigidos a implementar planes de contingencia y recuperación de la información.

Punto C. ¿Se han establecido marcos regulatorios que regulen el desarrollo, adquisición y uso de herramientas de inteligencia artificial por parte de los Estados?

En la región, ningún país ha avanzado en la creación de legislación (en sentido formal y material) que sea clara, suficiente y alineada a estándares interamericanos (de transparencia y acceso a la información), y que se aplicables al desarrollo, adquisición y uso de herramientas de inteligencia artificial por parte del Estado.

Sin embargo, en el monitoreo institucional que adelantamos sobre políticas públicas de IA en la región, identificamos por ahora la existencia de dos regulaciones que no constituyen legislación en sentido formal o material.

La primera, expedida en Chile, y vigente desde 2024. Se trata de la Resolución N° 371 Exenta, del 12 de agosto, emitida por el Consejo para la Transparencia de Chile, que regula la transparencia algorítmica³⁶. Dicha resolución contiene recomendaciones que promueven buenas prácticas de transparencia y publicidad sobre los sistemas de decisiones automatizadas y semiautomatizadas que utiliza el sector público en Chile, así como promueve la transparencia proactiva en el Estado.

Entre las recomendaciones, se sugiere el deber de publicar información relacionada con las contrataciones de sistemas de decisiones automatizadas y semiautomatizadas (artículo 5.2), así

³⁵ Ibid.

³⁶ Resolución 371 Exenta que “Aprueba texto de las recomendaciones del Consejo para la Transparencia sobre Transparencia Algorítmica”. Disponible en: <https://www.bccn.cl/leychile/navegar?idNorma=1206232>

como la creación, a cargo de las entidades públicas obligadas, a crear condiciones favorables a la transparencia en contratos celebrados con terceros para licenciar el uso de sistemas de este tipo (artículo 5.2). La Resolución prevé que, en la respuesta de solicitudes de acceso a la información sobre sistemas de decisiones automatizadas y semiautomatizadas, se ofrezca la información sin que la opacidad, la complejidad o las características técnicas de esos sistemas sirvan como excusa o causal para su reserva (artículo 5.4).

El Consejo para la Transparencia recomienda así mismo la publicación i) tipo de sistema de IA en uso, ii) canal de consultas, reclamaciones o quejas, iii) titularidad y derechos de propiedad sobre el SDA, iv) nombre y RUT del proveedor del sistema, v) nombre del servicio o trámite en que se está usando el sistema de IA y el beneficio de su aplicación en dicho servicio estatal, vi) información de la entidad y unidad estatal que usa dicho sistema de IA, vii) identificar el acto administrativo (según su tipo, fecha, denominación, número y enlace) que ordenó su implementación, xi) las especificaciones del sistema de IA (funcionamiento, impacto positivo, motivos o fundamentos de sus resultados, entre otras características técnicas, incluidos los tipos de datos que usa), x) información sobre evaluaciones de impacto y de derechos humanos que hayan sido efectuadas y xi) enlace a la política de privacidad del organismo responsable de su despliegue o uso (artículos 5 y 6).

Sin embargo, esta Resolución prevé recomendaciones y buenas prácticas cuya aplicabilidad o apropiación no ha sido monitoreada a la fecha por el propio Consejo Nacional para la Transparencia, por lo que es incierto el nivel de impacto de su apropiación por los sujetos obligados (ministerios, intendencias, gobernaciones, gobiernos regionales, municipalidades, Fuerzas Armadas, de Orden y Seguridad Pública, los órganos y los servicios públicos, según el punto 4.9 de la Resolución).

Por otra parte, en Colombia, la Procuraduría General de la Nación (autoridad de la transparencia y el acceso a la información pública) y la Defensoría del Pueblo, emitieron conjuntamente, y en atención a lo ordenado por la sentencia de tutela de la Corte Constitucional, T-067 de 2025, la Directiva Conjunta N° 007 de 2025 sobre “estándares sobre transparencia algorítmica para los sistemas algorítmicos utilizados por el Estado”.³⁷

La Directiva Conjunta fija estándares mínimos de transparencia algorítmica aplicables a las entidades públicas y otros sujetos obligados según la Ley Transparencia y del Derecho de Acceso a la Información Pública. La Directiva reconoce que “no crea derechos u obligaciones de manera directa como un acto administrativo individual, [pero que] la Directiva tiene efectos vinculantes dentro de la administración pública”³⁸ (artículo 2, literal d).

En su contenido, la Directiva señala que como parte del principio de transparencia los sujetos obligados deben publicar en sus sitios web información como i) el nombre del sistema de IA del cual están haciendo uso, ii) informar su objetivo, iii) informar sobre su función institucional, es

³⁷ Procuraduría General de la Nación; Defensoría del Pueblo (Septiembre 30, 2025). Estándares sobre transparencia algorítmica para los sistemas algorítmicos utilizados por el Estado. Disponible en: <https://www.defensoria.gov.co/documents/20123/3407303/300925DirectivaConjunta007.pdf/c47e1175-6f60-058a-3e0b-3dfaf82d5f23?t=1759261267112>

³⁸ Procuraduría General de la Nación; Defensoría del Pueblo (Septiembre 30, 2025). Estándares sobre transparencia algorítmica para los sistemas algorítmicos utilizados por el Estado. Disponible en: <https://www.defensoria.gov.co/documents/20123/3407303/300925DirectivaConjunta007.pdf/c47e1175-6f60-058a-3e0b-3dfaf82d5f23?t=1759261267112> ver artículo 2, literal D.

decir, a qué acción, programa o trámite está siendo integrada la IA, iv) el estado del sistema y su etapa de implementación, desarrollo o despliegue, v) cómo se lleva a cabo el tratamiento de datos y vi) cuál es el tipo de datos utilizados, vii) quién es su desarrollador, viii) ofrecer información de contacto para objetar, reclamar o solicitar revisiones sobre los resultados de dicha herramienta de IA, y ix) contacto para más información (artículo 12).

También ordena la realización de análisis de impacto algorítmico previo al despliegue de sistemas de IA (artículo 14), y ordena la publicación y actualización anual de la información sobre uso de sistemas de IA (artículo 20 al 22). En materia contractual, prevé que se instauren cláusulas contractuales con los proveedores de dichos sistemas, para que provean la información prescrita por la Directiva, así como cláusulas enfocadas en la adopción de buenas prácticas por parte de éstos (artículo 23).

La Directiva crea una Mesa de Seguimiento a los Estándares, que sesionará al menos dos veces al año para acompañar, monitorear, evaluar y fortalecer los estándares mínimos que ésta fija. La composición de dicha Mesa integran, entre otros, a las entidades públicas, y la sociedad civil (artículo 28).

Punto D. ¿Qué marcos regulatorios existentes (por ejemplo: leyes de protección de datos personales) son aplicables actualmente en el desarrollo y uso de sistemas de inteligencia artificial? ¿Han sido estos marcos suficientes para abordar los riesgos específicos que plantea la inteligencia artificial, o se han identificado limitaciones en su aplicación?

La arquitectura jurídica de base en los países de la región es aplicable a la IA, es decir, la legislación en materia de protección de datos y privacidad, acceso a la información, derecho de consumo, protección de la competencia, protección al derecho de autor, entre otros.³⁹

Esta arquitectura jurídica enfrenta, por una parte, el reto de actualizarse a las nuevas dinámicas del ecosistema digital, así como a los retos que traen consigo nuevas y emergentes tecnologías. Y, por otra, el reto de ampliar y defender su aplicación frente a las grandes empresas tecnológicas transnacionales que se domicilian en Estados Unidos (Big Tech), un reto que enfrentan por igual las regulaciones sobre protección de datos (y otras de la era preinternet), y las regulaciones para la IA.

En materia de protección de datos, hemos identificado que cada vez más se cierra la brecha regulatoria entre los países que cuentan con una ley en la materia, y los que no la tienen. A la fecha, países como Honduras, Guatemala, Bolivia y Venezuela siguen sin regular dicha materia, y próximamente entrarán en vigencia la ley de protección de datos en Chile y Paraguay. Pero países con leyes que imitan el estándar regulatorio europeo, como Brasil o Colombia, han enfrentado en su aplicación frente a las Big Tech respuestas diferenciadas que denotan, en el fondo, el impacto que tiene el tamaño del mercado de usuarios en la respuesta de dichas compañías frente al derecho local.

³⁹ Camacho, L (Febrero , 2026). Chile y el complejo camino para regular las plataformas digitales. Derechos Digitales. Disponible en: <https://www.derechosdigitales.org/recursos/chile-y-el-complejo-camino-para-regular-las-plataformas-digitales/>

Por ejemplo, en 2024 en Brasil, la Agencia Nacional de Protección de Datos (ANPD) emitió una medida cautelar para suspender⁴⁰ en el país la entrada en vigencia de la nueva política de privacidad de Meta, actualizada el 26 de junio⁴¹ y que autorizaba el despliegue de MetaAI en las plataformas de redes sociales como Facebook e Instagram. La compañía hizo dicho anuncio sin ofrecer al público de América Latina (a diferencia del de Europa) la posibilidad de consentir para rechazar dicha política.

La decisión de la autoridad se enfocó en proteger el consentimiento de las personas usuarias en ese país, así como del público menor de edad usuario de los servicios de dicha compañía. Sin embargo, la decisión luego fue retirada por la ANPD en agosto de ese año⁴², frente a la promesa de la compañía de ofrecer a los usuarios del país mecanismos para ejercer su consentimiento.

Sin embargo, países como Colombia y Brasil que tienen leyes de protección de datos que imitan el estándar del Reglamento Europeo de Protección de Datos no merecieron el mismo tratamiento por parte de Meta. En Colombia, por ejemplo, no se ofreció para usuarios del país la posibilidad de ejercer su consentimiento para rechazar los nuevos cambios que trajo consigo la política de junio de 2024.

También en Brasil, la suspensión judicial del servicio de la red social X en 2024, derivada del sucesivo incumplimiento de órdenes previas de la justicia por parte de la plataforma que se negó a la entrega de información solicitada para la investigación de una causa judicial, muestra cómo las autoridades enfrentan otros retos para que se reconozca la legitimidad de sus decisiones. Para entonces, el fallo desacatado por la compañía se acompañó de una campaña digital de desprestigio impulsada por el CEO de X, Elon Musk, contra el juez que suscribió dichas órdenes, Alexandre de Moraes.⁴³

También en Brasil, y de manera reciente, la plataforma digital X evadió una recomendación conjunta⁴⁴ suscrita por la Autoridad Nacional de Protección de Datos, el Ministerio Público Fiscal, y el Servicio Nacional del Consumidor, enfocada en ordenar a la plataforma corregir los defectos de su IA (Grok) que facilitaron a cientos de usuarios la generación y difusión masiva de *deepfakes* que afectaban a mujeres, niños, niñas y adolescentes.

⁴⁰ ANPD determina suspensão cautelar do tratamento de dados pessoais para treinamento da IA da Meta (Julho 02, 2024). Disponible en: <https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-determina-suspensao-cautelar-do-tratamento-de-dados-pessoais-para-treinamento-da-ia-da-meta>

⁴¹ Meta, Centro de Privacidad, Política de Privacidad vigente a partir del 26 de junio de 2024. Disponible en: <https://www.facebook.com/privacy/policy/version/25238980265745528/>

⁴² Meta cumpre exigências da ANPD e poderá retomar, com restrições, o uso de dados pessoais para treinamento de inteligência artificial (Agosto 30, 2024). Disponible en: <https://www.gov.br/anpd/pt-br/assuntos/noticias/meta-cumpre-exigencias-da-anpd-e-podera-retomar-com-restricoes-o-uso-de-dados-pessoais-para-treinamento-de-inteligencia-artificial>

⁴³ Camacho, L. (Septiembre, 2024). La suspensión de X en Brasil, y la vigencia del Estado de Derecho. Disponible en: <https://www.derechosdigitales.org/recursos/la-suspension-de-x-en-brasil-y-la-vigencia-del-estado-de-derecho/>

⁴⁴ MPF, ANPD e Senacon recomendam que X impeça geração e circulação de conteúdos sexualizados indevidos por meio do Grok (Janeiro 20, 2026). Disponible en: <https://www.mpf.mp.br/o-mpf/unidades/pr-sp/noticias/mpf-anpd-e-senac-recomendam-que-x-impeca-geracao-e-circulacao-de-conteudos-sexualizados-indevidos-por-meio-do-grok>

El argumento de X para evitar cumplir con la medida, apuntó en señalar⁴⁵ que X Brasil no tendría legalmente nada que ver con X en Estados Unidos, o sea, con la red social, y que por eso no cumpliría lo ordenado por las autoridades.⁴⁶ Un argumento de vieja data en la región, cuando las plataformas se enfrentan a decisiones judiciales en su contra.⁴⁷

La aplicación de marcos regulatorios sobre protección de datos en contra de las Big Tech sugiere dificultades de todo tipo que, por analogía, también serían extensibles a la regulación de la IA que emerja en países de la región. Entre los retos más comunes, emergen cuestionamientos de tipo procesal (notificación de decisiones judiciales sobre actores que no se domicilian en un país determinado), cuestionamientos sustanciales (como si una empresa domiciliada en un país hace parte o no del mismo conglomerado económico manejado por la red social en cuestión, por ejemplo); pero dejan entrever que, por encima, se superponen cuestionamientos continuos de las empresas al ejercicio del Estado de Derecho y la soberanía jurisdiccional de los países que actúan y emiten órdenes en su contra. Una realidad que, insistimos, es susceptible de permear la aplicación futura de las leyes de IA que van emergiendo en la región.

Punto F. ¿Cómo es el proceso de toma de decisiones para la adopción de políticas públicas que incluyan el uso de sistemas de inteligencia artificial y de inteligencia artificial generativa? ¿Se evalúa la necesidad y proporcionalidad de las mismas? ¿Se presta especial atención al impacto diferencial que puede tener para grupos históricamente discriminados? ¿Se documenta esta evaluación?

En la región identificamos dos momentos significativos de construcción de políticas pública sobre IA. El primero, comprendido entre 2018 y hasta mediados de 2023, cuando se avanzó en distintos países de la región en la construcción y actualización de las estrategias, planes o marcos éticos para la regulación de la IA. Y el segundo, caracterizado por la proliferación de proyectos de ley enfocados en replicar el modelo de regulación de la IA de Europa, y que se mantiene hasta ahora.

Esa primera faceta, países como Brasil, Colombia, Chile y Uruguay desplegaron la construcción de sus marcos éticos con procesos de participación heterogéneos en sus metodologías, contenido y alcance.

En la mayoría de procesos, se dio paso a la participación ciudadana, tal y como lo detallamos en nuestro informe “Inteligencia Artificial y participación en América Latina: las estrategias nacionales de IA”, publicado en 2022.⁴⁸ Sin embargo, la participación en la construcción de esas

⁴⁵ Dias, T. (February, 2026). X Tried to Sidestep Brazil’s Inquiry on AI Deepfakes- The Government Just Pushed Back. Tech Policy Press. Disponible en: <https://www.techpolicy.press/x-tried-to-sidestep-brazils-inquiry-on-ai-deepfakes-the-government-just-pushed-back/>

⁴⁶ MPF, ANPD e Senacon recomendam que X impeça geração e circulação de conteúdos sexualizados indevidos por meio do Grok (Janeiro 20, 2026). Disponible en: <https://www.mpf.mp.br/o-mpf/unidades/pr-sp/noticias/mpf-anpd-e-senacon-recomendam-que-x-impeca-geracao-e-circulacao-de-conteudos-sexualizados-indevidos-por-meio-do-grok>

⁴⁷ Corte Constitucional de Colombia, SU-429 de 2019. Disponible en: <https://www.corteconstitucional.gov.co/relatoria/2019/su420-19.htm>

⁴⁸ Hernández, L.; Canales, M.P.; De Souza, M. (2020). Inteligencia Artificial y participación en América Latina: Las estrategias nacionales de IA. Derechos Digitales. Disponible en: <https://www.derechosdigitales.org/wp-content/uploads/IA-Participacion-ES-2022.pdf>

políticas de *derecho blando*, no permeó todo el proceso de confección de la política en cuestión, y a veces se la relegaba a las fases finales de construcción de la política en cuestión.

En esta primera faceta, identificamos defectos asociados a la apertura de los procesos de participación, a las barreras tecnológicas –consultas exclusivamente en línea- para acceder y participar de cada consulta. En todas las experiencias analizadas identificamos defectos asociados al procesamiento de los aportes obtenidos de la participación ciudadana, y la incapacidad de los tomadores de decisiones de dar trazabilidad a dichos comentarios e identificar cómo se habría transformado la política pública en razón a dicha participación.⁴⁹

En materia de diversidad geográfica, así como de participación de grupos sexo-género diversos, grupos étnicos y racializados, entre otros, nuestra conclusión apunta a la ausencia de políticas afirmativas que, en los países objeto de estudio, se enfocaran en aumentar la participación de grupos sociales minoritarios, por lo que los procesos de participación y consultas que tuvieron lugar en estos países terminaron por reforzar su exclusión.⁵⁰

De hecho, en nuestro informe caracterizamos la participación según sectores (academia, instituciones públicas, individuos en general, sector público y sociedad civil) e identificamos que en los cuatro países las instituciones públicas tenían un muy bajo índice de participación en la construcción de dichas políticas, planes o estrategias de regulación ética de la IA. Además, identificamos que el sector privado tiene niveles de participación comparativamente altos en todos los países, superando la participación conjunta de actores como la academia y la sociedad civil.⁵¹

En cuanto a los procesos metodológicos que sirven para orientar la participación de la ciudadanía en estos procesos, identificamos también la ausencia de reglas claras más allá del envío de comentarios escritos a procesos de consulta, y que reducen el derecho a la participación ciudadana a un único proceso que no facilita la interacción dialógica entre las múltiples partes interesadas.

Concluimos dicho informe⁵², recomendando a los tomadores de políticas públicas que, entre otros, introdujeran mecanismos que fomentasen la participación de distintos grupos sociales subrepresentados en dichos procesos, así como aumentaran los mecanismos y oportunidades de participación (para que no fuese solo en línea, y por una única vez), incorporando etapas para el diálogo entre distintos actores, así como procesos para la evaluación del éxito de las políticas públicas una vez implementadas.

Sobre la segunda faceta de elaboración de políticas públicas de IA en la región, destacamos cómo la participación ciudadana en la construcción de proyectos de ley que regulan la IA se ve seriamente amenazada por el *lobby* legislativo y opaco de las Big Tech.

Al respecto, destacamos la experiencia de Brasil y el trabajo emprendido por la Comisión de Juristas de Brasil para la regulación de la IA (CJSUBIA) cuyo trabajo fue discutir y unificar una

⁴⁹ Ibid.

⁵⁰ Ibid.

⁵¹ Ibid.

⁵² Ibid.

propuesta legislativa para la IA, de cuyo trabajo se desprendería con posterioridad el actual proyecto de ley de IA en trámite N° 2338/2023.

El análisis de dicha experiencia está contenido en el artículo “Participação Cidadã e Regulação da Inteligência Artificial: Análise da Composição Setorial das Audiências Públicas na CJSUBIA”⁵³ que analizó el proceso de participación a propósito del trabajo deliberativo de dicha Comisión.

En el artículo concluimos que, aunque pueden existir procesos de construcción de políticas regulatorias con vocación abierta y participativa, para garantizar la participación ciudadana diversa y representativa, y no solo de los actores especializados en las discusiones de tecnologías, hace falta implementar medidas proactivas para propiciar la convocatoria, inclusión y visibilización de grupos sociales subrepresentados en políticas públicas digitales. También, en la experiencia del CJSUBIA evidenciamos cómo el ritmo acelerado de algunos procesos de discusión de lo que serán iniciativas regulatorias, la rápida transformación y discusión de documentos, así como los acortados tiempos para elaborar opiniones técnicas sobre asuntos complejos, puede expulsar del proceso de participación a organizaciones, grupos y personas interesadas sin experiencia previa en dichos espacios.

En la experiencia del CJSUBIA identificamos también la importancia de garantizar el acceso público, abierto y transparente a la información sobre los procesos de participación, los espacios de recogida y procesamiento de datos y comentarios de una política pública, así como de generación de información trazable que permita identificar cómo un borrador de política pública se transforma con el tiempo.⁵⁴

En este espacio, en concreto, también fue visible la facilidad con la que espacios participativos pueden verse capturados por pocos actores poderosos en la conversación, como sucedió con el rol protagónico y eclipsante del sector privado, en especial de las Big Tech cuyo papel fue predominante y desigual en comparación con la presencia de otros grupos de interés presentes en las discusiones. La participación significativa precisa de metodologías para equilibrar voces, y asegurar la representación mayoritaria de voces históricamente subrepresentadas de estos espacios (como grupos de personas sexo-género diversas, grupos étnicos y racializados, grupos de mujeres, grupos de interés enfocados en la protección de las niñeces en línea, entre otros).⁵⁵

También en Brasil, la experiencia del debate reciente del PL 2338/2023 que busca regular la IA es dicente sobre la transformación de la participación significativa en la construcción de políticas públicas digitales.

⁵³ García, J.M.; Venturini, J. (Julho, 2024). Participação Cidadã e Regulação da Inteligência Artificial: Análise da Composição Setorial das Audiências Públicas na CJSUBIA. Volume 5, número 1, Jornal Internet & Sociedade. Disponible en: <https://revista.internetlab.org.br/participacao-cidada-e-regulacao-da-inteligencia-artificial-analise-da-composicao-setorial-das-audiencias-publicas-na-cjsubia/>

⁵⁴ Ibid.

⁵⁵ Ibid.

Tal y como lo documentó en un reportaje de enero de 2026 la organización Aos Fatos⁵⁶, lobistas de los Big Tech (Amazon, Google, IBM, Meta, Microsoft, Open AI) visitaron al menos en 83 ocasiones a la comisión del congreso que discute en Brasil la regulación de la IA. Luego de sus visitas, diversos congresistas adoptaron posturas más flexibles sobre, por ejemplo, habilitar el uso de sistemas de IA para fines de vigilancia y seguridad pública; así también como visiones que se enfocan que establecer marcos regulatorios de la IA que no sean tan exigentes o inflexibles que no habiliten a la innovación, por lo que se debería adoptar enfoques legislativos más blandos que no ahuyenten la inversión de las empresas tecnológicas en el país, entre otros.

También, en 2024 relatamos⁵⁷ cómo dicho *lobby* habría sido efectivo para modular propuestas legislativas que se enfocaban en la garantía de derechos. Tal fue el caso de la transformación de las evaluaciones obligatorias de impacto en derechos, en evaluaciones meramente facultativas como resultado de presión de las Big Tech sobre los legisladores que, así, dieron media sanción al proyecto de ley en cuestión. Si bien la discusión del proyecto de ley N° 2338/2023 ha sido un proceso participativo, en éste la representación de grupos de interés es desigual y está atravesada por las herramientas y estrategias de *lobby* empresarial que, en el caso de Brasil, no está regulado con reglas claras, transparentes y que faciliten la rendición de cuentas.

PUNTO G. ¿Son adoptados sistemas de inteligencia artificial para facilitar mecanismos de vigilancia estatal?

Hemos identificado recientemente en la región el despliegue de sistemas de IA en el contexto de la vigilancia estatal en distintos escenarios. El primero, el despliegue de sistemas de reconocimiento facial en contextos de protesta en Argentina, Perú y Colombia. El segundo, el despliegue de sistemas de reconocimiento facial para la seguridad ciudadana, la gestión de la movilidad y el tránsito vehicular, así como el tránsito humano. Y el tercero, que ya fue documentado ante esta Relatoría en un informe sobre ciberpatrullaje y perfilamientos en línea⁵⁸ que elaboramos, enfocado en la vigilancia estatal del espacio cívico digital (un asunto ya abordado por esta Relatoría en el informe publicado en septiembre de 2025). Sobre este último contexto no nos referiremos, en tanto que no hemos documentado acciones recientes de ciberpatrullaje en donde se implementen tecnologías de IA.

⁵⁶ Aos Fatos (Janeiro 29, 2026). Deputados da comissão da IA atuaram em prol de 'big techs' após visitas de lobistas. Disponible en: <https://www.aosfatos.org/noticias/deputados-da-comissao-da-ia-atuaram-em-prol-de-big-techs-apos-visitas-de-lobistas/>

⁵⁷ García, J.M.; Camacho L. (Diciembre, 2024). Regulación de la IA en Brasil: media sanción para un proyecto de ley que seguirá en discusión. Derechos Digitales. Disponible en: <https://www.derechosdigitales.org/recursos/regulacion-de-la-ia-en-brasil-media-sancion-para-un-proyecto-de-ley-que-seguira-en-discusion/>

⁵⁸ Camacho, L. (Agosto, 2024). Perfilamiento en redes sociales y ciberpatrullaje como nuevas modalidades de la vigilancia masiva desplegada por los Estados: casos relevantes en América Latina. Contribución de Derechos Digitales para la consulta "sobre tecnologías de vigilancia digital y derechos humanos" emprendida por la Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos CIDH. Derechos Digitales. Disponible en: <https://www.derechosdigitales.org/recursos/perfilamiento-en-redes-sociales-y-ciberpatrullaje-como-nuevas-modalidades-de-la-vigilancia-masiva-desplegada-por-los-estados-de-la-region-casos-relevantes-en-america-latina/>

El primer escenario lo documentamos en 2025 en nuestra contribución dirigida a la ONU sobre “Impacto de la vigilancia digital asistida por IA en los derechos de reunión y de asociación, incluidos los efectos inhibitorios”⁵⁹, donde centramos la atención en Argentina, Colombia y Perú.

En Argentina,⁶⁰ se encuentra ahora mismo en despliegue el software de reconocimiento facial “LUNA”, de origen ruso, el cual es usado por la Policía Federal en contextos de protestas para identificar a manifestantes. La primera vez que se usó fue en 2022, en el contexto de marchas pacíficas en contra del nuevo acuerdo suscrito entre la Nación y el Fondo Monetario Internacional.

Este software, según notas de prensa que reportaron de su despliegue en el contexto de las protestas de los jubilados que son convocadas hace más de dos años todos los días miércoles frente al Congreso de la Nación, se usa con el objetivo identificar a los manifestantes para, según el gobierno nacional, retirarles subsidios a quienes marchan contra el Estado, -constituyendo una amenaza con el potencial de generar un efecto amedrentador o *chilling effect*. Sobre este sistema el Estado no desplegó mecanismos de consulta, transparencia sobre su origen o adquisición, ni evaluaciones de impacto en derechos.

En Perú,⁶¹ a partir de septiembre de 2025 se desplegó por la Policía Nacional un sistema de reconocimiento facial en contextos de protesta que se puso a prueba en el marco de las Marchas de la Generación Z, motivadas por el descontento social contra el gobierno de turno. El sistema se encuentra habilitado y articulado con el sistema de videovigilancia desplegado en ciudades como Lima, y acompañado de otras tecnologías de vigilancia como drones para habilitar la realización de búsquedas automatizadas entre las bases de datos de personas buscadas por la justicia, y quienes se encuentran en escenarios de protesta pacífica. Sobre este sistema el Estado no desplegó mecanismos de consulta, transparencia sobre su origen o adquisición, ni evaluaciones de impacto en derechos.

En Colombia,⁶² en el marco de las protestas que tuvieron lugar en noviembre de 2019, las autoridades de Policía amenazaron con desplegar en la ciudad de Bogotá sistemas de reconocimiento facial instalados en el helicóptero Halcón. Aunque la sociedad civil local identificó con posterioridad y a través de solicitudes de acceso a la información que dicho despliegue no habría sido materialmente efectuado, los anuncios públicos y las imágenes de

⁵⁹ Camacho Gutiérrez, L.; Lara-Castro, P. (Noviembre, 2025). Contribución a la consulta para el informe temático titulado “Impacto de la vigilancia digital y asistida por IA en los derechos de reunión y de asociación, incluidos los efectos inhibitorios”, convocado por la Relatoría Especial sobre el Derecho de Reunión y Asociación Pacífica. Uso de sistemas de reconocimiento facial y ciberpatrullaje en contextos de protesta en países de América Latina. Derechos Digitales. Disponible en: https://www.derechosdigitales.org/wp-content/uploads/2025/11/Derechos-Digitales_Submission-on-AI-and-FoAA.pdf

⁶⁰ Política Argentina (18 Marzo, 2025). En la previa a una nueva marcha, Seguridad pidió la detención de 28 personas que estuvieron en la movilización de la semana pasada. Disponible en: <https://www.politicargentina.com/notas/202503/64590-en-la-previa-a-una-nueva-marcha-seguridad-pidio-la-detencion-de-29-personas-que-estuvieron-en-la-movilizacion-de-la-semana-pasada.html>

⁶¹ Carrasco, M. (27 Septiembre, 2025). PNP desplegará cámaras con reconocimiento facial y drones para identificar a quienes cometan actos violentos en marchas. Infobae. Disponible en: <https://www.infobae.com/peru/2025/09/27/npn-desplegara-camaras-con-reconocimiento-facial-y-drones-para-identificar-a-quienes-cometan-actos-violentos-en-marchas/>

⁶² Semana (20 Noviembre, 2019). Así es el Halcón, el helicóptero que vigilará el paro en Bogotá usando reconocimiento facial. Disponible en: <https://www.semana.com/nacion/articulo/paro-21-de-noviembre-el-helicoptero-que-vigilara-en-bogota-usando-reconocimiento-facial/641014/>

prensa con la supuesta tecnología en manos de las autoridades de policía, habrían tenido un potencial efecto amedrentador sobre una protesta pacífica legítima.

En la segunda tendencia, el informe del Consorcio AlSur “Reconocimiento facial y tecnologías de vigilancia en América Latina”,⁶³ y del cual hacemos parte, identificó en 2025 tecnologías de reconocimiento actualmente activas en Argentina, Brasil, Chile, Colombia, México y Uruguay.

Entre los hallazgos, se identificaron 45 iniciativas nuevas desplegadas por los Estados desde 2021 y hasta 2025. En otro informe, habríamos identificado 38 iniciativas en curso⁶⁴ desplegadas entre 2017 y 2019) por lo que la tendencia es que en la región el interés por su adquisición va en claro aumento.

Identificamos en el informe de 2025 que las tecnologías de reconocimiento facial se despliegan, en general, para tres fines: de seguridad ciudadana (reconocimiento de prófugos), reconocimiento de patentes y otros fines asociados a la seguridad vehicular, y el control migratorio y de fronteras.

En materia de seguridad ciudadana, se destacan los siguientes países y casos⁶⁵:

- Argentina: El Sistema de Reconocimiento Facial de Prófugos de Mar del Plata (vigente desde 2022); el Sistema de Reconocimiento Facial en Provincia de Santa Fe para la videovigilancia pública (vigente desde 2024).
- Bolivia: El sistema de Seguridad BOL-110 (vigente desde 2020).
- Brasil⁶⁶: el Sistema SMART Sampa (vigente desde 2023); Sistema de Reconocimiento Facial en Copacabana, Río de Janeiro, (vigente desde 2023).
- Chile: Sistema Automatizado de Identificación Biométrica ABIS (vigente desde 2022); Sistema Integrado de Teleprotección con Inteligencia Artificial (vigente desde 2024).
- Ecuador: Sistema de Reconocimiento Facial en el ECU 911 (vigente desde 2025).

En materia de seguridad vehicular e identificación de patentes, destacamos las siguientes iniciativas:⁶⁷

⁶³ Ibarreche, X.; Martínez, C.; Lorenzo, C. (2025). Reconocimiento facial y tecnologías de vigilancia en América Latina: casos, proveedores y dinámicas comerciales. AlSur. Disponible en: <https://www.alsur.lat/sites/default/files/2025-08/Reconocimiento%20facial%20y%20tecnologias%20vigilancia.pdf>

⁶⁴ Venturini, J., Garay, V. (2021). Reconocimiento facial en América Latina. Tendencias en la implementación de una tecnología perversa. AlSur. Disponible en: <https://estudio.reconocimientofacial.info/reports/ALSUR->

⁶⁵ Ibarreche, X.; Martínez, C.; Lorenzo, C. (2025). Reconocimiento facial y tecnologías de vigilancia en América Latina: casos, proveedores y dinámicas comerciales. AlSur. Disponible en: <https://www.alsur.lat/sites/default/files/2025-08/Reconocimiento%20facial%20y%20tecnologias%20vigilancia.pdf>

⁶⁶ Por su parte, la organización de la sociedad civil de Brasil “O Panóptico”, ha monitoreado por su cuenta más de 516 iniciativas de reconocimiento facial desplegadas y activas en todo Brasil, que tienen un impacto potencial sobre más de 22 millones de personas en la actualidad. Sugerimos consultar: <https://www.opanoptico.com.br/#mapa>

⁶⁷ Ibarreche, X.; Martínez, C.; Lorenzo, C. (2025). Reconocimiento facial y tecnologías de vigilancia en América Latina: casos, proveedores y dinámicas comerciales. AlSur. Disponible en: <https://www.alsur.lat/sites/default/files/2025-08/Reconocimiento%20facial%20y%20tecnologias%20vigilancia.pdf>

- Argentina: El Sistema de Reconocimiento Facial en puestos de control en Provincia de Salta (vigente desde 2024); Puestos Móviles de Identificación Facial en Provincia de Mendoza (vigente desde 2024);
- Brasil: Sistema de Reconocimiento Facial en Paraíba (vigente desde 2023); Sistema de Reconocimiento Facial en Rondonia (vigente desde 2021); Sistema de Reconocimiento Facial en el Estado de Espírito Santo (vigente desde 2021).
- Chile: Sistema de Televigilancia con Inteligencia Artificial (vigente desde 2024)

En materia de control fronterizo y migratorio,⁶⁸ destaca por ejemplo el creciente protagonismo del Programa EUROFRONT, financiado por la Unión Europea, y que facilita a países como Ecuador, Bolivia, Perú, Argentina, Paraguay y Brasil, la adquisición y despliegue de tecnologías de identificación biométrica y de vigilancia en contextos donde el Estado de Derecho puede ser poroso y por el que transitan personas en condiciones de movilidad humana precarias.

También, en materia migratoria y de control fronterizo es crítico el despliegue del proyecto “Fronteira Tech –Muralha Inteligente” en la triple frontera Brasil, Argentina y Paraguay. El proyecto emplea cámaras de videovigilancia, sistemas de reconocimiento facial, GPS, sensores térmicos y de movimiento, tecnologías de identificación de patentes, entre otras.

El informe identifica también una presencia mayoritaria de empresas tecnológicas transnacionales como proveedoras principales de tecnologías para cualquiera de estos tres fines. Entre las más representativas se encuentran IDEMIA (Francia), NEC (Japón), Dahua (China), Hikvision (Canadá), Veridos (Alemania), Innovatics (República Checa), entre otras. El informe en cuestión resalta dos patrones en el relacionamiento de empresas del sector privado con los Estados en esta materia: (i) la opacidad de los términos y condiciones contractuales de los acuerdos sostenidos entre empresas y Estados, (ii) la opacidad sobre la naturaleza, diseño, características, arquitectura de las tecnologías y la rendición de cuenta de los proveedores frente a terceros, como la ciudadanía afectada por su despliegue.

En el informe se destaca también la ausencia generalizada de mecanismos de consulta previa al despliegue de las iniciativas de vigilancia desplegadas por países de la región, la falta de claridad sobre su eventual impacto en derechos humanos, ni transparencia proactiva asociada a los procesos contractuales en los que se soportó su adquisición.

Punto I. ¿Han prohibido los Estados el uso de algún sistema de inteligencia artificial en su territorio? Por ejemplo, tecnologías de puntuación social o de reconocimiento facial.

Destacamos dos precedentes en la región donde la sociedad civil organizada logró, mediante acciones judiciales, que los Estados prohibieran el despliegue de sistemas de IA de reconocimiento facial:

En Argentina, el Centro de Estudios Legales y Sociales (CELS), el Observatorio de Derecho Informático (ODIA), la Fundación Vía Libre y Democracia en Red, interpusieron en 2019 una acción contra el sistema de reconocimiento facial con IA en la Ciudad de Buenos Aires, que

⁶⁸ Ibid.

llevó a su suspensión por parte de la justicia en 2022⁶⁹. Las organizaciones argumentaron sobre los riesgos y violaciones de dicho sistema en materia de derechos humanos como la privacidad, protección de datos, libertad de expresión y autodeterminación informativa. Así mismo, señalaron que en los últimos tres años el Gobierno de la Ciudad de Buenos Aires, en convenio con el Registro Nacional de las Personas (ReNaPer), ha tratado con fines de vigilancia los datos biométricos de más de 7,5 millones de personas⁷⁰.

En este contexto, resulta particularmente preocupante el tratamiento de los datos personales mediante el cruce de la base de datos del ReNaPer con el registro de personales con orden de captura judicial (que incluye aproximadamente 40 mil personas), con el objetivo de ubicarlos mediante cámaras de videovigilancia equipadas con tecnología de identificación biométrica. Las organizaciones demandantes advirtieron una extralimitación en las tareas de identificación vulnerando el derecho a la intimidad de millones de personas. A partir de estas observaciones, se puso en tela de juicio la validez constitucional del Sistema de Reconocimiento Facial de Prófugos (SRFP), debido a las falencias detectadas en el manejo de la información personal. El gobierno actual ha buscado la reactivación de esta tecnología para identificar a manifestantes e iniciar procesos penales en su contra con el retiro de planes sociales⁷¹.

En Brasil, el Instituto de Defesa de Consumidores (idec) en conjunto con el Colectivo de Defesa de los Derechos Humanos (CADHu), Intervozes y Artículo 19 encaminaron en 2018 acciones judiciales contra la empresa “Via Quatro”, operador de la concesionaria responsable de la Línea 4 del Metro de San Pablo, que instaló puertas interactivas en algunas estaciones con el objetivo de proyectar anuncios personalizados a las personas usuarias, así como también para recabar información que pudiera servir en la medición de información demográfica y analizar la reacción emocional de los usuarios del servicio por intermedio de cámaras equipadas con tecnología de reconocimiento facial⁷².

Las organizaciones mencionadas argumentaron que la empresa no ha tomado las medidas necesarias para garantizar la seguridad de las personas usuarias del transporte público, por ejemplo, el Metro de São Paulo no produjo un informe de impacto para la contratación de tecnología de reconocimiento facial, ni tiene estudios que demuestren la seguridad de las bases de datos para la implementación del nuevo sistema de vigilancia. Además, no ha desarrollado ninguna política de protección de datos para niños, niñas y adolescentes, en concordancia con los principios establecidos por la Ley General de Protección de Datos (LGPD)⁷³.

En 2021, el Juzgado Civil N° 37 de São Paulo ordenó a la empresa concesionaria que suspendiera de inmediato la captura de imágenes, sonidos o cualquier tipo de dato personal de los usuarios, salvo que contara con su autorización explícita. Asimismo, el fallo estableció que,

⁶⁹ VíaLibre (2024). La Justicia define si vuelve el reconocimiento facial en las calles de Buenos Aires: ¿cómo auditar al vigilante artificial? Disponible en: <https://www.vialibre.org.ar/la-justicia-define-si-vuelve-el-reconocimiento-facial-en-las-calles-de-buenos-aires-como-auditar-al-vigilante-artificial/>

⁷⁰ CELS (2022). El Ministerio de Seguridad de la Ciudad buscó información biométrica de 7 millones de personas de manera ilegal. Disponible en: <https://www.cels.org.ar/web/2022/04/el-ministerio-de-seguridad-de-la-ciudad-busco-informacion-biometrica-de-7-millones-de-personas-de-manera-ilegal/>

⁷¹ VíaLibre (2023). NO al reconocimiento facial ¡No de nuevo! Disponible en: <https://www.vialibre.org.ar/noalreconocimientofacial2023/>

⁷² Idec (2020). Metrô não garante segurança de dados de reconhecimento facial em SP. Disponible en: <https://idec.org.br/noticia/metro-de-sp-nao-tem-garantia-de-seguranca-de-dados-de-reconhecimento-facial>

⁷³ Ibid.

en caso de querer implementar dicho sistema, la compañía debía informar de manera clara y detallada a los consumidores sobre el alcance de la recolección y el tratamiento de sus datos, además de implementar los mecanismos necesarios para garantizar un consentimiento informado. Por último, la sentencia impuso una multa por daño moral colectivo de 100.000 reales⁷⁴.

Punto J. ¿Han sido adoptados sistemas de inteligencia artificial generativa para ayudar en la escritura de documentos estatales tales como sentencias, actos administrativos, leyes? ¿Existe algún control o límite impuesto por los Estados para el uso de estos sistemas en la elaboración de documentos públicos? ¿Existe alguna medida que garantice la supervisión humana de documentos de especial relevancia, como, por ejemplo, las sentencias?

En el informe que preparamos⁷⁵ en 2025 para la Relatoría Especial de la ONU sobre independencia de jueces y abogados, identificamos el despliegue de innumerables sistemas de IA en los sistemas judiciales de la región.

En Argentina,⁷⁶ el despliegue en 2017 del sistema de IA denominado Pretoria, para automatizar procesos en la Fiscalía de la Ciudad de Buenos Aires, cuyo fin era predecir el resultado de la decisión judicial de manera rápida para asistir la resolución de casos. En 2024, con la masificación de ChatGPT a nivel global, la misma Fiscalía se inclinó a su adopción para los mismos fines iniciales, predecir el resultado de las decisiones judiciales, y diseñar los borradores de sentencias.

En Brasil, desde 2018, se desplegó por parte del Supremo Tribunal Federal (STF) el sistema VICTOR.⁷⁷ Un sistema de IA desarrollado por la Universidad de Brasilia que automatiza los procesos de apelación e identifica casos con un potencial impacto social para ser revisados por el Alto Tribunal. En Diciembre de 2024, el STF también desplegó el Módulo de Apoyo para elaborar borradores con IA, denominado “marIA”⁷⁸, enfocado en elaborar resúmenes de las opiniones de los jueces del STF, preparar reportes sobre apelaciones y llevar a cabo un análisis Inicial de los casos en trámite. Sus desarrolladores afirman que la herramienta podría ser auditable, aunque no existe información pública sobre cómo se ejecutaría dicho proceso en la práctica.

⁷⁴ Para más información, consultar:

<https://repositorio.mpd.gov.ar/jspui/bitstream/123456789/4661/1/Brasil.IDEC.Fallo%20completo.pdf>

⁷⁵ Camacho, L.; Lara-Castro, P. (2025). Responde call for inputs for the thematic report on “artificial intelligence and judicial systems”, United Nations Special Rapporteur in the independence of judges and lawyers. Derechos Digitales. Disponible en: https://www.derechosdigitales.org/wp-content/uploads/Derechos-Digitales_AI-and-justice_final.pdf

⁷⁶ Estevez, C.; Linares, S.; Fillottrani, P. (2020). Prometea, transformando la administración de justicia con herramientas de inteligencia artificial. Banco Interamericano de Desarrollo BID. Disponible en: <https://publications.iadb.org/es/publications/spanish/viewer/PROMETEA-Transformando-la-administracion-de-justicia-con-herramientas-de-inteligencia-artificial.pdf>

⁷⁷ Peixoto, F. (2020). Projeto Victor: Relato do Desenvolvimento da Inteligência Artificial na Repercussão Geral do Supremo Tribunal Federal. Revista Brasileira de Inteligência Artificial e Direito, Vol.1, N.1. Available at: <https://rbiad.com.br/index.php/rbiad/article/view/4/4>

⁷⁸ Prensa Latina (2025). Corte Suprema de Brasil implementa herramienta con IA. Available at: <https://www.prensa-latina.cu/2024/12/16/corte-suprema-de-brasil-implementa-herramienta-con-ia/>

También, en 2023, la Fiscalía General de la Unión⁷⁹ (AGU por sus siglas en portugués) suscribió un acuerdo con OpenAI para ser integrada en el sistema de información judicial de la Fiscalía, para resumir casos y sugerir borradores para asistir la tarea del ente investigador.

Sin embargo, el 2023 fue de público conocimiento el caso de uso de ChatGPT por un juez federal del Estado de Acre⁸⁰, en donde se habría usado dicha herramienta para elaborar apartados extensos de la decisión judicial. Cuando dicho uso trascendió a la opinión pública por las partes involucradas en el proceso, el juez del caso culpó a sus asistentes judiciales por el uso de la herramienta de manera irresponsable.

El Consejo Nacional de Justicia, organismo regulador del sistema judicial a nivel nacional en Brasil, fue el primero en América Latina en abordar el uso de sistemas de IA en la judicatura.

En 2020 emitió una resolución⁸¹ sobre el tema, la cual fue derogada en 2025 por una más amplia que autoriza explícitamente el uso de IA para redactar decisiones judiciales. La resolución de 2020, más breve, destacaba la protección de derechos y las obligaciones de publicidad y transparencia de los modelos de IA, como la necesidad de divulgar estos sistemas, evaluar su impacto en la protección de datos e identificar riesgos, aunque no precisaba responsables ni mecanismos de implementación, dejando la gobernanza únicamente en manos del Consejo.

La nueva resolución reafirma que los sistemas de IA en la judicatura deben respetar los derechos fundamentales; adopta una clasificación de riesgos (excesivo, alto y bajo), prohibiendo los sistemas de riesgo excesivo; y crea un ecosistema local de gobernanza de IA supervisado por el Comité Nacional de Inteligencia Artificial del Poder Judicial, encargado de la implementación y supervisión de la norma.⁸²

Sin embargo, aún no está claro cómo esta nueva regulación se alinearán con los procesos legislativos en curso para regular la IA en Brasil, que proponen una taxonomía de riesgos distinta y asignan la gobernanza a otros organismos, especialmente a aquellos liderados por la autoridad nacional de protección de datos (ANPD).⁸³

En Colombia, como ya hemos referido previamente, se desplegó en 2017 el Fiscal Watson por parte de la Fiscalía General de la Nación. Sugerimos en este sentido ver el ítem 1 A y B.

También, en 2020, la Corte Constitucional desplegó un sistema de IA propio denominado “PretorIA” para asistir el proceso de selección y clasificación de acciones de tutela o procesos de amparo. El despliegue inicial de dicha herramienta, que se efectuó bajo un piloto previo desplegado con otro nombre y por otro proveedor, fue criticado en su momento por organizaciones de la sociedad civil por la falta de transparencia en su adopción y despliegue,

⁷⁹ Advocacia-Geral da União (2023). AGU inova no uso de inteligência artificial para aprimorar eficiência e prestação de serviços à sociedade Disponible en: <https://www.gov.br/agu/pt-br/comunicacao/noticias/agu-inova-no-uso-de-inteligencia-artificial-para-aprimorar-eficiencia-e-prestacao-de-servicos-a-sociedade>

⁸⁰ G1 (2023). Juiz usa inteligência artificial para fazer decisão e cita jurisprudência falsa; CNJ investiga caso. Disponible en: <https://g1.globo.com/politica/blog/daniela-lima/post/2023/11/13/juiz-usa-inteligencia-artificial-para-fazer-decisao-e-cita-jurisprudencia-falsa-cnj-investiga-caso.ghtml>

⁸¹ Conselho Nacional de Justiça (2020). Resolução Nº 332 de 21/08/2020 “Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Poder Judiciário e dá outras providências”. Disponible en: <https://atos.cnj.jus.br/atos/detalhar/3429>

⁸² Ibid.

⁸³ Ibid.

así como la nula participación en el proceso de toma de decisión sobre la integración de la IA en un procedimiento jurídica y socialmente sensible como el de selección de los casos que serán revisados por el Alto Tribunal Constitucional –un proceso que, en el pasado, habría sido cuestionado por manipulación y corrupción, y donde solo un puñado de casos es seleccionado entre cientos de miles que llegan a diario a la Corte para ser revisados según criterios de interés, relevancia, novedad, entre otros-.

Sin embargo, y a partir de la emergencia de Largos Modelos de Lenguaje (LLM) de acceso público y gratuito, han trascendido a la opinión pública casos de uso de sistemas de IA por jueces que han empleado a ChatGPT para la redacción de decisiones judiciales. Entre esos casos se destacan los siguientes:

- Uso de ChatGPT en 2023, por un juzgado que decidió un caso sobre el acceso de un menor con autismo a una prestación determinada en salud. El juez usó dicho LLM para que, en su lugar, identificara la jurisprudencia aplicable al caso, y elaborase el raciocinio judicial a través del cual le dio la razón a la parte accionante.⁸⁴
- Uso de ChatGPT en 2024, por la Sala Penal del Tribunal Superior del Distrito de Pereira, para decidir un caso de homicidio por accidente de tránsito, la respuesta de ChatGPT fue copiada de manera integral en la decisión judicial en donde se usó un LLM de generación de texto para la estimación de un cálculo matemático.⁸⁵

En Colombia, la Corte Constitucional decidió a través de la sentencia T-323 de 2024⁸⁶ que el uso de ChatGPT por jueces debía cumplir los siguientes parámetros:

- A. Transparencia, entendida como la obligación de evidenciar con claridad y precisión el uso, alcances y ubicación en las actuaciones o decisiones de los resultados obtenidos por la utilización de tales herramientas, que permita a los usuarios e interesados su pleno conocimiento y la posibilidad efectiva de contradicción.
- B. Responsabilidad, comprendida como aquella obligación que existe de que el usuario de la herramienta de IA se encuentre capacitado y comprenda los impactos del uso de estas tecnologías, para a su vez dar cuenta del origen, idoneidad y necesidad del uso de la IA y la información suministrada por la misma, la cual debe ser verificada.
- C. Privacidad, es aquel deber de custodiar y proteger la reserva de los datos personales y sensibles que se ponen en conocimiento de la administración de justicia para cumplir con los fines propios de la Rama Judicial.
- D. No sustitución de la racionalidad humana, como expresión de la imposibilidad ética y jurídica de sustituir la acción y la responsabilidad del individuo de la especie humana en la gestión de las actuaciones y decisiones judiciales.

⁸⁴ Rama Judicial de Colombia, Juzgado Primero Laboral del Circuito de Cartagena, Sentencia No. 32, 30 de enero de 2023. Disponible en: <https://forogpp.com/wp-content/uploads/2023/01/sentencia-tutela-segunda-instancia-rad.-13001410500420220045901.pdf>

⁸⁵ Tribunal Superior del Distrito Judicial de Pereira, Sala de Decisión Penal, Magistrado Ponente Manuel Yarzagaray Bandera, Sentencia de Segunda Instancia, 3 de abril de 2024. Disponible en: <https://www.ambitojuridico.com/sites/default/files/2024-04/Sent-76001600019320138073401-24.pdf>

⁸⁶ Corte Constitucional de Colombia, T-323 de 2024. Disponible en: <https://www.corteconstitucional.gov.co/relatoria/2024/T-323-24.htm>

E. Seriedad y verificación, que implica la obligación de realizar un estricto escrutinio sobre las fuentes, alcances, restricciones, posibilidades, falencias y riesgos que presente la herramienta de cara a la actuación en curso o a la solución del problema jurídico correspondiente.

F. Prevención de riesgos, como mandato en cuanto aplicar los estándares adecuados de control sobre situaciones que generen riesgo por la aplicación de tecnologías tales, en aspectos como imprecisiones, desactualizaciones, alucinaciones, sesgos, inconsistencias y demás.

G. Igualdad y equidad, en cuanto se erradiquen todas las formas de discriminación relacionadas con la aplicación de sesgos derivada del uso de tales tecnologías y su impacto negativo en la eficacia de los derechos humanos.

H. Control humano, en tanto considerando los anteriores criterios, siempre se permita la realización efectiva de escrutinios sobre las actuaciones y decisiones en que se usen herramientas de IA, mediante el acceso a la debida información y el uso de recursos que deban ser resueltos por autoridades humanas. i

I. Regulación ética, que implica el desarrollo de estándares de comportamiento individual que se adecúen a los mandatos superiores y legales y a las pautas razonables para el uso de tales tecnologías por parte de los funcionarios y servidores de la Rama Judicial.

J. Adecuación a buenas prácticas y estándares colectivos, en tanto se apliquen los esquemas razonables que se definan para el funcionamiento de la Rama Judicial, desde su autonomía e independencia, a partir de las definiciones que adopten sus autoridades, tanto en sede de administración como de orientación jurisprudencial.

K. Seguimiento continuo y adaptación, a efecto que el uso de tales tecnologías consulte los avances jurídicos, sociológicos y tecnológicos que se vayan implementando, así como los esquemas de mejora y control que se construyan en forma progresiva.

L. Idoneidad. El uso de las tecnologías de la información y de las comunicaciones en la gestión y trámite de los procesos judiciales y asuntos en curso debe ser adecuado para facilitar y agilizar el acceso a la justicia.

A partir de dicha decisión, la Corte Constitucional ordenó al Consejo Superior de la Judicatura (que regula al poder judicial) a emitir directivas para regular el uso de la IA (incluidos los LLM) en la justicia. Sin embargo, se trata de directivas⁸⁷ que, en su contenido, delegan extensas responsabilidades en manos de cada juez para determinar el impacto del modelo de IA para los derechos de las partes, la protección del debido proceso, su impacto en privacidad, etc. Cargas que compiten con las propias de la tarea judicial en un escenario como el colombiano, donde los operadores judiciales enfrentan altos volúmenes de trabajo y poca capacidad humana para resolver casos en tiempo.

Por lo que, en la práctica, las directivas resultan impracticables, o al menos, no existe un incentivo para que jueces presionados por las mediciones de desempeño de su labor se vean apropiando su contenido y, por tanto, informando a las partes procesales sobre el uso de una

⁸⁷ Consejo Superior de la Judicatura (Diciembre, 2024). Acuerdo PCSJA24-122243, por medio de la cual se fijan lineamientos para el uso respetuoso, responsable, seguro y ético del uso de sistemas de inteligencia artificial en la Rama Judicial. Disponible en: <https://actosadministrativos.ramajudicial.gov.co/Home/GetFile?nombreArchivo=PCSJA24-12243.pdf>

determinada herramienta de IA o LLM en el proceso de confeccionar fallos y decidir causas judiciales.

En México,⁸⁸ en 2023, durante una audiencia del Tribunal Superior Electoral en donde uno de los jueces del Tribunal usó ChatGPT para obtener contexto histórico y social de la expresión “ya sabes quién”, con denotación política, en el marco de un caso en que se determinaría la legalidad de una campaña electoral.

En Paraguay,⁸⁹ en 2019, la sociedad civil elevó alertas por el proyecto piloto de integración de Prometea –una IA desarrollada para y por el contexto judicial argentino-, para ser desplegado por la Corte Suprema de Justicia con el fin de automatizar tareas como predecir el contenido de las decisiones judiciales. Sobre el resultado de dicho piloto no hay información pública disponible que transparente sus resultados o si la herramienta, en efecto, terminó siendo apropiada del todo por el Alto Tribunal.

En Perú,⁹⁰ en 2023, la Corte Superior de Justicia de Lima Sur, usó ChatGPT para estimar el porcentaje que le correspondía pagar de cuota alimenticia al padre obligado. Nuevamente, se aplicó un LLM diseñado para la generación de texto, en la elaboración de un cálculo matemático –es decir, para un fin para que el dicho LLM no fue entrenado inicialmente-.

Diversas preocupaciones en materia de derechos humanos surgen del uso de la IA en la justicia. Estos desafíos plantean riesgos serios en materia de protección de datos, acceso a la información, soberanía y gobernanza de datos, entre otros.

En materia de derecho a la privacidad y protección de datos, observamos que en la mayoría de los casos descritos, el uso de herramientas propietarias (de IBM, OpenAI u otros proveedores) implica políticas de privacidad que no necesariamente están alineadas con las leyes de protección de datos de los países donde se despliegan.

Esto resulta problemático no solo en términos del uso de sistemas de IA como ChatGPT, Fiscal Watson o VICTOR por parte de autoridades judiciales que desconocen el impacto de su interacción con estas herramientas y la privacidad de los demandantes y demás partes que participan en cada proceso judicial, sino también para los titulares de datos que desconocen cómo ejercer sus derechos si información personal o sensible es filtrada o expuesta por estos proveedores privados.

Este riesgo se agrava en países de América Latina donde el uso de IA en los sistemas judiciales tiene lugar sin políticas establecidas de anonimización de casos ni marcos sólidos de gobernanza de datos y ciberseguridad aplicables al poder judicial.

⁸⁸ Tribunal Electoral del Poder Judicial de la Federación (marzo de 2023). Sesión pública, miércoles 29 de marzo de 2023. Disponible en: <https://www.youtube.com/live/OwaZg3quyls?t=3679s>

⁸⁹ Sequera, M (2021). Inteligencia artificial en la sala constitucional en Paraguay. TEDIC. Disponible en: <https://www.tedic.org/inteligencia-artificial-en-la-sala-constitucional-en-paraguay/>

⁹⁰ Poder Judicial del Perú (2023). Juzgado Civil Transitorio de San Juan de Miraflores, sede Valle Riestra, resolución de segunda instancia, 27 de marzo. Disponible en: <https://img.lpderecho.pe/wp-content/uploads/2023/03/Expediente-00052-2022-18-3002-JP-FC-01-LPDerecho.pdf>

En materia de transparencia y acceso a la información, advertimos que el uso de sistemas de IA en la judicatura está mediado, en parte, por acuerdos firmados entre autoridades judiciales y empresas tecnológicas transnacionales como IBM u OpenAI, particularmente en los casos de Colombia y Brasil. La naturaleza comercial de estos acuerdos no es un factor diferenciador en términos de transparencia, ya que otros memorandos de entendimiento o colaboraciones con iniciativas no comerciales (como la que suscribió la Corte Constitucional en el primer piloto de IA, denominado Prometea, y que fue desarrollado por la Universidad de Buenos Aires) también son, en su mayoría, confidenciales.

La mayoría de estos acuerdos son secretos o no se divulgan públicamente, lo que los vuelve opacos para el público no solo respecto de sus costos y los términos de las licencias adquiridas, sino, de manera más amplia, respecto de especificaciones técnicas críticas como el código fuente del algoritmo (como en el caso de Fiscal Watson, propiedad de IBM), lo que limita la posibilidad de que actores de la sociedad civil auditen de forma independiente el desempeño de las herramientas en la asistencia o ejecución efectiva de tareas judiciales fundamentales.

Sin embargo, la transparencia es, sin duda, un asunto que va más allá de los acuerdos contractuales entre autoridades judiciales y proveedores de sistemas de IA. También se extiende al ámbito de la transparencia y la ética judicial, donde se espera que los jueces que utilicen tales herramientas de IA por iniciativa propia lo revelen de manera clara y abierta.

Casos como el identificado en Acre, Brasil, generan preocupación sobre cómo, en la práctica, los jueces pueden estar utilizando chat conversacional como ChatGPT sin informar claramente a las partes involucradas en el proceso y, cuando se les cuestiona, incluso pueden negar su uso, pese a que posteriormente pueda confirmarse.

Además, los sesgos de los jueces en el uso de herramientas de IA pueden afectar seriamente su capacidad para ejercer habilidades judiciales críticas, como la verificación y el contraste de fuentes humanas y documentales, el deber de veracidad y transparencia vinculado a sus actuaciones judiciales, entre otros. En los casos que involucran el uso de ChatGPT, los sesgos humanos en la interacción con la IA se vuelven más visibles. Por ejemplo:

- El derecho a la transparencia procesal: Este se vio particularmente afectado en el caso de Brasil, donde el juez de Acre desestimó las preocupaciones sobre el uso de ChatGPT y justificó un error grave, a saber, la alucinación de un precedente crucial en el caso que resolvió, como un error humano atribuible a un integrante de su despacho judicial.
- El derecho a la explicabilidad de las decisiones judiciales: En los casos de Perú, Colombia y México, fragmentos literales generados por ChatGPT fueron copiados y pegados en decisiones judiciales y presentados por los jueces como veraces y objetivos. No se adoptaron medidas para verificar o contrastar estos resultados.
- La identificación de la norma aplicable: Los casos de Brasil (Acre) y los dos de Colombia demuestran cómo, debido a sesgos de neutralidad y objetividad, los jueces confiaron ciegamente en que ChatGPT identificaría correctamente la norma aplicable en cada caso. En cambio, la herramienta produjo alucinaciones que pasaron inadvertidas para los jueces.

En este punto, hacemos un llamado a la RELE para que emita a los Estados una recomendación general de implementar moratorias sobre la adopción de sistemas de IA en el poder judicial hasta que su uso e implementación hayan sido debidamente regulados con una perspectiva de

derechos humanos. En los casos en que dicha regulación aún esté ausente, su despliegue debería satisfacer previamente evaluaciones de impacto en derechos humanos y evaluaciones de impacto en protección de datos; y si estas evaluaciones no son favorables, la herramienta de IA debería ser prohibida para su uso en la judicatura.

También, creemos que la RELE puede insistir en su futuro informe en la necesidad de que los Estados transparenten las herramientas de IA que actualmente se utilizan en el poder judicial y de aquellos sistemas de IA que se usaron pero fueron descartados en el pasado en pilotos poco exitosos. También, debe insistir a los Estados en la relevancia de informar al público sobre cada sistema de IA: (i) sus capacidades y especificaciones técnicas, así como su propósito inicial; (ii) su impacto en los derechos humanos y especialmente en el debido proceso; (iii) las medidas adoptadas para mantener a un humano en el circuito de decisión; y (iv) el impacto de su uso desde su adopción.

Y por último, urge que la RELE pueda hacer un llamado de atención sobre la adopción acelerada de sistemas de IA en tareas judiciales esenciales, y sobre cómo estas tecnologías pueden exacerbar problemas preexistentes dentro de las ramas judiciales, así como introducir nuevos desafíos. Entre ellos se incluyen el debilitamiento de los procesos de verificación y la erosión de la independencia del juicio humano, elemento esencial para el diseño de sistemas de justicia democráticos, entre otros.

PUNTO K. Los Estados utilizan actualmente *chatbots* basados en inteligencia artificial generativa para la atención y el atendimento de la ciudadanía en sus servicios públicos?

Los Estados de América Latina utilizan *chatbots* en el marco de sus narrativas de modernización de la administración pública y asistencia efectiva a la ciudadanía y sus solicitudes. El repositorio de investigación “Sistemas de Algoritmos Públicos”, desarrollado por la Universidad de los Andes en Colombia⁹¹ ha registrado un total de 157 sistemas de *chatbots* desplegados por entidades estatales en América Latina. Colombia, México, Brasil y Argentina registran el mayor número de estos con 42, 36, 15 y 12 sistemas respectivamente.

Los *chatbots* destacan por su capacidad de comunicarse a través del procesamiento del lenguaje natural en un formato conversacional, lo que facilita que se les atribuyan cualidades humanas y sean vistos como entes antropomórficos. Si bien dichas cualidades buscan optimizar la experiencia de las personas usuarias, su funcionamiento depende de que las personas compartan información personal con el sistema. Esta dinámica implica que, a través de los *chatbots*, los gobiernos también recolectan datos de la ciudadanía, muchas veces sin que esta sea plenamente consciente de ello, lo que intensifica las inquietudes relacionadas con la privacidad⁹². Estas cuestiones se abordarán con mayor detalle a continuación.

PUNTO L. Algunos de los problemas que puede generar el uso de *chatbots* basados en sistemas de inteligencia artificial generativa son la difusión de información incorrecta, la falta de opciones que permitan al ciudadano obtener respuesta a sus preguntas, la vulnerabilidad a los

⁹¹ Para más información, consultar: <https://algoritmos.uniandes.edu.co/>

⁹² Ferreyra, E. (2024). Boti: estudio sobre el chatbot con procesamiento del lenguaje natural del Gobierno de la Ciudad de Buenos Aires. Derechos Digitales. Disponible en: <https://ia.derechosdigitales.org/wp-content/uploads/2024/11/2024-ARG-BOTI-ESP.pdf>

ataques cibernéticos externos y el riesgo de filtración de datos confidenciales. ¿Qué medidas han adoptado los Estados que utilizan estas tecnologías para mitigar y evitar estos riesgos?

Siguiendo lo señalado por la Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (ACNUDH), en tanto los *chatbots* gubernamentales son tecnologías automatizadas que recolectan y procesan información, presentan ciertos riesgos críticos para la privacidad y la protección de los datos personales de la ciudadanía. Por ejemplo, existe el peligro de uso indebido o acceso no autorizado a los datos, lo que puede llevar a posibles daños a las personas afectadas. Otro ejemplo, la elaboración de perfiles detallados sin consentimiento explícito, exponiendo a los usuarios a discriminación o manipulación. En ese contexto, se recomienda a los Estados la implementación cuidadosa de los sistemas en cuestión conforme a los más altos estándares en protección de la privacidad y los datos personales, garantizando la protección contra abusos y la explotación de datos sensibles. La investigación de Derechos Digitales sobre el uso de *chatbots* en Argentina y Uruguay ha encontrado las siguientes prácticas desplegadas por los Estados para la mitigación de riesgos, revelando áreas cruciales de mejora en cada caso.

- Auditorías e investigaciones de organismos fiscalizadores: En el caso del Chat Boti en Argentina, la Auditoría General de la Ciudad de Buenos Aires y la Defensoría del Pueblo de la Ciudad desplegaron una evaluación exhaustiva del funcionamiento y la gestión del chatbot. Entre los hallazgos de la auditoría resaltaron la ausencia de una estructura administrativa claramente definida podría impactar en la eficiencia operativa y la transparencia del manejo de los recursos; así como información la falta de información detallada sobre el tipo de datos personales recolectados, su finalidad y las medidas de seguridad adoptadas, aspecto que fue integrado luego de la auditoría.
- Pruebas de continuidad basadas en la norma ISO 27001: Aunque en Argentina se despliegan estas pruebas y se realizan ejercicios de recuperación de manera periódica, la documentación que respalde dichas actividades no está disponible al público, lo que puede sugerir que estas prácticas fundamentales no se están ejecutando de forma completa o adecuada.
- Acceso al código fuente de los sistemas: En el caso del chatbot Coronavirus UY, la Agencia de Gobierno Electrónico y Sociedad de la Información y del Conocimiento (Agesic) ofrece acceso al código fuente de la aplicación como mecanismo capaz de facilitar eventuales auditorías independientes.

2. Participación social, transparencia y acceso a la información

Punto A. ¿Ha habido participación de la sociedad civil, el mundo académico u otros actores relevantes durante la evaluación, el desarrollo, la adquisición y la implementación de herramientas/sistemas de inteligencia artificial por parte de los Estados? ¿Han participado estos actores en la supervisión del uso de estas tecnologías por parte de los Estados? Describa los mecanismos utilizados y su impacto efectivo.

Para empezar, al tratarse de sistemas de IA con potencial de afectación a derechos fundamentales, la participación y el acceso a información sobre la implementación de

herramientas de IA por parte de los Estados, así como sobre las justificaciones que llevan a su adopción constituye un derecho⁹³.

Dicho esto, en los casos de estudio profundizados en la investigación de Derechos Digitales⁹⁴ encontramos como característica común la relativa opacidad de las implementaciones y un bajo nivel de participación pública para el diseño de las soluciones, así como para las fases de “testeo” y mejora de las herramientas de IA desplegadas.

En el caso del Sistema Nacional de Empleo (SINE) en Brasil⁹⁵ se realizó una investigación con usuarias y usuarios del sistema como la única instancia de participación ciudadana en el desarrollo de la iniciativa, sin participación de organizaciones de la sociedad civil. Posteriormente, representantes de sindicatos en el país alegaron por la transparencia activa de la política solicitando informes periódicos sobre los impactos de la iniciativa en la política pública de intermediación de mano de obra, exigiendo criterios de evaluación de tales informes, así como claridad en su periodicidad.

Para el caso chileno, durante el desarrollo de la licitación que dio pie al sistema Alerta Niñez⁹⁶, el único mecanismo de consulta conocido fue la entrevista a cinco expertas y expertos en gestión de datos y protección de la infancia. El estado de implementación del sistema, calificado en fase “piloto”, ha sido igualmente utilizado por la autoridad a cargo como justificación para el bajo nivel de información públicamente disponible sobre la iniciativa. Con respecto al Sistema Predictivo del Delito Urbano⁹⁷ en este país, cabe recordar la ausencia de instancias externas de auditoría que aseguren la transparencia y el control público del sistema, así como la falta de mecanismos para difundir activamente información sobre su funcionamiento y resultados.

En Colombia, el proceso de implementación de PretorIA⁹⁸ ha llevado al desarrollo de prácticas que favorecen el acceso a información sobre el proceso de selección de tutelas, particularmente a través de la puesta en disposición del registro de las audiencias seleccionadas en un canal de YouTube. No obstante, el proyecto carece de mayores iniciativas de transparencia activa. Por su parte, el despliegue del Software Watson Explorer (Fiscal Watson)⁹⁹ no cuenta con mecanismos de monitoreo y evaluación de la eficiencia de la

⁹³ Velasco, P. y Venturini, J. (2021). Decisiones automatizadas en la función pública en América Latina. Una aproximación comparada a su aplicación en Brasil, Chile, Colombia y Uruguay. Derechos Digitales. Disponible en: https://ia.derechosdigitales.org/wp-content/uploads/2021/03/CPC_informeComparado.pdf

⁹⁴ Para más información, consultar: <https://ia.derechosdigitales.org/publicaciones/>

⁹⁵ Bruno, F.; Cardoso, P.; Faltay, P. (2021). Sistema Nacional de Empleo y la gestión automatizada de la desocupación laboral. Derechos Digitales. Disponible en: https://www.derechosdigitales.org/wp-content/uploads/CPC_informe_BRASIL_castellano.pdf

⁹⁶ Valderrama, M (2021). Sistema Alerta Niñez y la predicción del riesgo de vulneración de derechos de la infancia. Derechos Digitales. Disponible en: https://ia.derechosdigitales.org/wp-content/uploads/2021/03/CPC_informe_Chile.pdf

⁹⁷ Buschman, J. (2021). Sistema Predictivo del Delito Urbano: Producción algorítmica de Zonas de Vigilancia y Control en la Ciudad. Derechos Digitales. Disponible en: <https://ia.derechosdigitales.org/wp-content/uploads/2022/05/CPC-IA-Chile-2-ES.pdf>

⁹⁸ Saavedra, V.; Upegui, J.C. (2021). PretorIA y la automatización del procesamiento de causas de derechos humanos. Derechos Digitales. Disponible en: https://ia.derechosdigitales.org/wp-content/uploads/2021/03/CPC_informe_Colombia.pdf

⁹⁹ Palacios, L., Forero, V., & Labarthe, S. (2024). Fiscal Watson: estudio del uso de Inteligencia Artificial en la Fiscalía General de la Nación en Colombia. Derechos Digitales. Disponible en: https://ia.derechosdigitales.org/wp-content/uploads/2024/11/2024-COL-Fiscal_Watson-ESP.pdf

herramienta, así como tampoco se publica de manera pública, abierta y accesible información sobre su desempeño y efectividad.

Para el caso uruguayo no se consideraron instancias activas de participación ciudadana en el desarrollo de la aplicación Coronavirus UY¹⁰⁰. Sin embargo, el código fuente de la aplicación se encuentra disponible ante la ciudadanía mediante un requerimiento directo a AGESIC. Como ha sido mencionado anteriormente, la apertura del código fuente tiene por objetivo brindar transparencia y ofrecer garantías a la población sobre el manejo de los datos recolectados por la aplicación.

En Argentina, el Chatbot “Boti”¹⁰¹ ha estado sujeto a auditorías de la autoridad de protección de datos del Gobierno de la Ciudad de Buenos Aires (GCBA) que detallaron la ausencia de responsables de los sistemas de información, así como la falta de evidencia de procedimientos dirigidos a implementar planes de contingencia y recuperación de la información.

Punto B. ¿Se han adoptado medidas específicas para garantizar la participación sustantiva de grupos históricamente discriminados/en situación de vulnerabilidad, incluidos pueblos indígenas y comunidades afrodescendientes?

La falta de mecanismos que fomenten la participación de géneros diversos, así como de grupos históricamente marginalizados y grupos alejados de los centros urbanos, sigue siendo una característica anclada en la realidad de la política pública latinoamericana. El despliegue de sistemas de IA no es la excepción.

En nuestra investigación¹⁰² evidenciamos la ausencia de mecanismos específicos para promover la participación de diferentes géneros, así como comunidades históricamente en situación de vulnerabilidad en las fases de la evaluación, el desarrollo, la adquisición y la implementación de herramientas de IA por parte de los Estados. Similar situación ocurre con los mecanismos para fomentar la participación efectiva de personas ciudadanas que se encuentran en condición de discapacidad; para este grupo poblacional no constatamos alternativas que les permitan enterarse y participar de la elaboración de políticas y sistemas de IA relacionados que pueden afectar directamente el goce de sus derechos.

Punto C. ¿Qué información hacen pública los Estados de forma proactiva sobre el uso de sistemas de inteligencia artificial, inclusión de inteligencia artificial generativa? ¿Se incluye información sobre el origen y representatividad de los datos con los que se entrenan los modelos de inteligencia artificial? (por ejemplo: información sobre sistemas utilizados y sus finalidades, proveedores o desarrolladores, presupuesto público involucrado, evaluaciones de impacto realizadas, procesos de tomas de decisiones, mecanismos de revisión, etc.)

En relación con el punto 1C de esta consulta, vale la pena señalar que, pese a que Colombia y Chile cuentan con una Directiva y una Recomendación sobre transparencia algorítmica, respectivamente, y que ordenan a las entidades públicas de cada país a la publicación de

¹⁰⁰ Yael, D. (2021). CoronavirusUY y la tecnología como solución a la pandemia. Derechos Digitales. Disponible en: https://ia.derechosdigitales.org/wp-content/uploads/2021/03/CPC_informe_URUGUAY.pdf

¹⁰¹ Ferreyra, E. (2024). Boti: estudio sobre el chatbot con procesamiento del lenguaje natural del Gobierno de la Ciudad de Buenos Aires. Derechos Digitales. Disponible en: <https://ia.derechosdigitales.org/wp-content/uploads/2024/11/2024-ARG-BOTI-ESP.pdf>

¹⁰² Consultar: <https://ia.derechosdigitales.org/publicaciones/>

información pública sobre el uso de sistemas de IA (sistemas automatizados, semiautomatizados, entre otros), a la fecha hemos identificado que no hay cumplimiento institucional de las autoridades de ninguno de esos dos países.

Por otro lado, en nuestra investigación¹⁰³ constatamos una carencia generalizada de esfuerzos sistemáticos de los Estados por desarrollar instancias de transparencia activa que integren la participación ciudadana, la evaluación de las diversas partes interesadas y la disponibilidad de información pública sobre los sistemas de IA adquiridos y desplegados. En ese contexto, la información pública disponible sobre las herramientas de IA es habitualmente pobre en cantidad y calidad, a menos que medien solicitudes de información a la autoridad (cuando hay una autoridad responsable).

A su vez, las instancias de acceso a la información sobre los sistemas, los datos con los que se entrenan y su auditoria no constituyen un eje central del despliegue de los sistemas de IA. En el punto E profundizaremos más sobre las obligaciones de respuesta ante solicitudes de acceso a la información por parte de la ciudadanía o “transparencia pasiva” de los casos analizados en la investigación.

Punto D. ¿Existen reportes que sugieran el uso de sistemas de inteligencia artificial por parte de instituciones estatales sin confirmación pública suficiente sobre su existencia, alcance o finalidad? Por ejemplo, el uso de tecnologías de espionaje y vigilancia contra disidentes políticos o periodistas, algoritmos para la selección de beneficiarios de programas sociales, la utilización de sistemas de IA en procesos electorales, etc.

En América Latina, el uso de sistemas de vigilancia digital y asistida por IA crece de forma acelerada y progresiva sin garantías suficientes para los derechos humanos¹⁰⁴. Así, la adquisición de sistemas de reconocimiento facial, por ejemplo, se encuentra amparada en políticas de opacidad y secreto que se excusan en la protección de la seguridad nacional.

En la práctica, estos sistemas han sido utilizadas para limitar el ejercicio del derecho a la protesta pacífica. En una contribución reciente a la Relatoría Especial sobre el Derecho de Reunión y Asociación Pacífica¹⁰⁵, documentamos algunos casos relevantes en la región que han sido detallados en puntos anteriores de este cuestionario de consulta.

En Argentina, el software de reconocimiento facial “LUNA”, de origen ruso, se utiliza para identificar a personas que participan en manifestaciones sociales. Se trata de una tecnología basada en inteligencia artificial que se encontraría operativa en la Policía Federal Argentina y en otras fuerzas de seguridad del Estado. Su implementación no estuvo precedida por procesos de consulta pública ni debate democrático. Frente a los pedidos de acceso a la información

¹⁰³ Para más información, consultar: <https://ia.derechosdigitales.org/publicaciones/>

¹⁰⁴ Venturini, J., Garay, V. (2021). Reconocimiento facial en América Latina. Tendencias en la implementación de una tecnología perversa. ALSur. Disponible en: https://estudio.reconocimientofacial.info/reports/ALSUR-Reconocimiento_facial_en_Latam-ES.pdf; Ibarreche, X.; Martínez, C.; Lorenzo, C. (2025). Reconocimiento facial y tecnologías de vigilancia en América Latina: casos, proveedores y dinámicas comerciales. ALSur. Disponible en: <https://www.alsur.lat/sites/default/files/2025-08/Reconocimiento%20facial%20y%20tecnologias%20vigilancia.pdf>

¹⁰⁵ Camacho Gutiérrez, L.; Lara-Castro, P. (Noviembre, 2025). Contribución a la consulta para el informe temático titulado “Impacto de la vigilancia digital y asistida por IA en los derechos de reunión y de asociación, incluidos los efectos inhibitorios”, convocado por la Relatoría Especial sobre el Derecho de Reunión y Asociación Pacífica. Uso de sistemas de reconocimiento facial y ciberpatrullaje en contextos de protesta en países de América Latina. Derechos Digitales. Disponible en: https://www.derechosdigitales.org/wp-content/uploads/2025/11/Derechos-Digitales_Submission-on-AI-and-FoAA.pdf

pública¹⁰⁶ presentados por organizaciones de la sociedad civil, las autoridades han rechazado sistemáticamente brindar información sobre las especificaciones técnicas del sistema, así como sobre los detalles de su proceso de adquisición y contratación.

En 2023, durante su primer discurso como presidente de Argentina, Javier Milei anunció que, frente a las protestas que tuvieran lugar durante su mandato, se implementaría el uso de software de reconocimiento facial para identificar a los manifestantes. La medida apuntaba especialmente a quienes fueran beneficiarios de subsidios, prestaciones económicas o programas sociales del Estado, con el objetivo de suspender dichos beneficios. Esta política quedó sintetizada en la consigna oficial: “el que corta, no cobra”¹⁰⁷.

Paralelamente al despliegue de esta tecnología, el gobierno habilitó la Línea 134 para que los titulares de planes sociales pudieran denunciar a los organizadores de protestas en caso de sentirse presionados para participar en las manifestaciones. En marzo de 2025, la justicia en lo contencioso administrativo dictó una medida cautelar que prohibió la suspensión de beneficios sociales a manifestantes, al considerar que dicha práctica generaba “un daño irreparable” a personas que ya se encuentran “en condición de vulnerabilidad”. Hasta la fecha, no se dispone de información pública oficial sobre la cantidad de personas a las que efectivamente se les habría suspendido los subsidios estatales en el marco de esta política.

En Perú, la Policía Nacional anunció la implementación de un sistema de reconocimiento facial en Lima, en el contexto de las protestas impulsadas por el descontento social hacia el gobierno de turno y protagonizadas por colectivos juveniles, estudiantes universitarios y gremios de transportistas. El despliegue de esta tecnología no fue precedido por mecanismos de consulta pública. Según informó la institución, el sistema, que integra herramientas de inteligencia artificial, estaría conectado a la red de videovigilancia del espacio público y a drones de monitoreo.

En Colombia, en noviembre de 2019, diversas protestas sociales estallaron en las principales ciudades del país en rechazo a las políticas económicas, políticas y de seguridad del gobierno de turno. Durante las manifestaciones, la Policía Nacional anunció el uso de un sistema de reconocimiento facial, con el fin de identificar a quienes estuvieran perturbando el orden público.

Ante este anuncio, organizaciones de la sociedad civil presentaron solicitudes de acceso a la información ante la Fiscalía General y la Policía Nacional, buscando conocer las especificaciones técnicas del sistema, los detalles de su implementación y su impacto real. En su respuesta, la Fiscalía señaló que no se había identificado a manifestantes ni se habían abierto procesos judiciales como resultado del uso de esta tecnología.

Como se evidencia en estos casos, el despliegue de sistemas de reconocimiento facial evidencia que no solo se trata de una tecnología con insuficiente regulación en la región, sino

¹⁰⁶ CELS (2022). Examen Periódico Universal - 4º Ciclo. Aporte al Examen de Argentina “Faltas de control en el uso de tecnologías para la prevención del delito y la investigación criminal”. Disponible en: <https://www.cels.org.ar/especiales/examenonu/wp-content/uploads/sites/13/2022/11/Reconocimiento-facial-EPU-2023.pdf>

¹⁰⁷ Chaves, F. (20 Diciembre, 2023). “El que corta no cobra”: una política de shock para erradicar los piquetes y garantizar la libre circulación que hoy será puesta a prueba. Disponible en: <https://www.infobae.com/politica/2023/12/20/el-que-corta-no-cobra-una-politica-de-shock-para-erradicar-los-piquetes-y-garantizar-la-libre-circulacion-que-hoy-sera-puesta-a-prueba/>

que su despliegue escapa a mecanismos de legitimidad social como los procesos de consulta pública.¹⁰⁸

Punto E. ¿Se han presentado solicitudes de acceso a la información pública relacionadas con el uso de inteligencia artificial por parte de los Estados? ¿Cuál fue la respuesta? En caso de denegación, ¿Cuál fue la justificación? ¿Ha sido mencionada alguna vez la seguridad pública o nacional como razón para denegar una solicitud de información?

Como hemos señalado en puntos anteriores, en el marco de nuestra investigación “Inteligencia Artificial e Inclusión en América Latina”¹⁰⁹ evidenciamos la baja disponibilidad de información respecto al estado y modo de ejecución de los sistemas de IA desplegados por los Estados en la región. Frente a esta característica, en la mayoría de los casos, la información sólo pudo ser obtenida a través de requerimientos formales de información pública, esto es, mecanismos de transparencia pasiva que, incluso, se mostraron limitados en algunos casos.

En Chile, la información sobre el Sistema Alerta Niñez¹¹⁰ fue restringida ya que la autoridad responsable utilizó el carácter piloto del programa como justificación para la escasa información pública disponible sobre esta iniciativa. En Argentina, el Gobierno de la Ciudad de Buenos Aires (GCBA) brindó una respuesta incompleta a una solicitud de información sobre el chatbot “Boti”¹¹¹. La vaguedad de las respuestas motivó un segundo requerimiento, el cual no obtuvo respuesta oportuna.

Otro caso relevante en Argentina es el de los pedidos de acceso a la información pública presentados por el Observatorio de Derecho Informático Argentino (O.D.I.A.) en relación con el uso de inteligencia artificial en cámaras de reconocimiento facial en la Ciudad de Buenos Aires. Según el Observatorio, ninguna de las dos solicitudes fue respondida de manera satisfactoria, lo que llevó a la organización a presentar una acción de amparo con el objetivo de suspender el uso del Sistema de Reconocimiento Facial de Prófugos. En una primera instancia la acción fue rechazada, pero luego apelaron y ampliaron la demanda para pedir por la inconstitucionalidad del uso de esta tecnología¹¹².

En ese contexto, es crucial recordar que el derecho de acceso a la información, previsto en el Sistema Interamericano de Derechos Humanos como parte integrante del derecho a la libertad de expresión, es reconocido, garantizado y regulado en los países analizados en nuestra investigación. Lo que encontramos es que, si bien el derecho de acceso a información puede

¹⁰⁸ Camacho Gutiérrez, L.; Lara-Castro, P. (Noviembre, 2025). Contribución a la consulta para el informe temático titulado “Impacto de la vigilancia digital y asistida por IA en los derechos de reunión y de asociación, incluidos los efectos inhibitorios”, convocado por la Relatoría Especial sobre el Derecho de Reunión y Asociación Pacífica. Uso de sistemas de reconocimiento facial y ciberpatrullaje en contextos de protesta en países de América Latina. Derechos Digitales. Disponible en: https://www.derechosdigitales.org/wp-content/uploads/2025/11/Derechos-Digitales_Submission-on-AI-and-FoAA.pdf

¹⁰⁹ Consultar: <https://ia.derechosdigitales.org/publicaciones/>

¹¹⁰ Valderrama, M (2021). Sistema Alerta Niñez y la predicción del riesgo de vulneración de derechos de la infancia. Derechos Digitales. Disponible en: https://ia.derechosdigitales.org/wp-content/uploads/2021/03/CPC_informe_Chile.pdf

¹¹¹ Ferreyra, E. (2024). Boti: estudio sobre el chatbot con procesamiento del lenguaje natural del Gobierno de la Ciudad de Buenos Aires. Derechos Digitales. Disponible en: <https://ia.derechosdigitales.org/wp-content/uploads/2024/11/2024-ARG-BOTI-ESP.pdf>

¹¹² García, JM (2024). Inteligencia artificial en el Estado: estudio colectivo sobre experiencias y riesgos para los derechos humanos. Derechos Digitales. Disponible en: https://ia.derechosdigitales.org/wp-content/uploads/2025/02/2024-LATAM-IA_en_el-Estado-ES.pdf

estar formalmente reconocido en los distintos países, todavía hay importantes desafíos en su implementación.

3. Supervisión, control y rendición de cuentas del uso de inteligencia artificial por parte de los Estados

Punto A. ¿Existen instancias de monitoreo y evaluación interna continua sobre la aplicación de tecnologías de inteligencia artificial por parte de los Estados? ¿Con qué frecuencia se llevan a cabo? ¿Son publicados?

A la fecha, no hemos documentado la existencia de este tipo de instancias que hayan sido desplegadas de manera efectiva en países de la región, pese a que hemos identificado la apropiación de sistemas de IA en la región que acumulan un despliegue de más de cinco años (Pretoria en Colombia; SINE en Brasil; STAPE en México; Portal EmpleaPY en Paraguay, entre otros) y en ocasiones, de casi diez años (Fiscal Watson en Colombia).

Punto B. ¿Existe alguna legislación, acto administrativo o política pública que establezca la necesidad de que las agencias estatales realicen evaluaciones de impacto sobre los derechos humanos cuando desarrollen o utilicen sistemas de IA, incluyendo la generativa?

En la actualidad, solo Chile y Colombia regulan sobre estos asuntos a través de la Recomendación del Consejo Nacional para la Transparencia, y la Directiva de la Procuraduría y la Defensoría del Pueblo, respectivamente, las cuales fueron abordadas en el punto 1C.

Sin embargo, la Recomendación en Chile¹¹³ prevé dicha obligación como uno de los ítems que deben ser publicados por las autoridades públicas, al señalar que deben publicar los enlaces “que dan cuenta de evaluaciones de impacto a la protección de datos” así como “antecedentes o enlaces que den cuenta de otras evaluaciones de impacto, como a los derechos humanos, junto con la indicación general de los riesgos detectados y las acciones de mitigación efectuadas” (artículo 1, C, IV, y V). Es decir, dicha Recomendación no especifica cómo llevar a cabo dicha evaluación de impacto, ni qué elementos mínimos debe contener.

En el caso de Colombia,¹¹⁴ la Directiva prevé lo que se denomina “análisis de impacto algorítmico” (artículo 13, C), es decir, una publicación que resume los principales hallazgos sobre los sesgos, riesgos o afectaciones diferenciadas de un sistema de IA, y su impacto específico en derechos humanos y auditorías sobre su desempeño.

Este análisis solo es exigible en los casos de sistemas algorítmicos usados por el estado para tareas críticas, como “(i) evaluar, priorizar y determinar la asignación o denegación de beneficios, subsidios, apoyos, empleos o servicios; (ii) predecir riesgos asociados a conductas como criminalidad, reincidencia, fraude, deserción, u otras situaciones que puedan afectar derechos; (iii) generar perfiles individuales o colectivos de personas; (iv) realizar inferencia de

¹¹³ Resolución 371 Exenta que “Aprueba texto de las recomendaciones del Consejo para la Transparencia sobre Transparencia Algorítmica”. Disponible en: <https://www.bccn.cl/leychile/navegar?idNorma=1206232>

¹¹⁴ Procuraduría General de la Nación; Defensoría del Pueblo (Septiembre 30, 2025). Estándares sobre transparencia algorítmica para los sistemas algorítmicos utilizados por el Estado. Disponible en: <https://www.defensoria.gov.co/documents/20123/3407303/300925DirectivaConjunta007.pdf/c47e1175-6f60-058a-3e0b-3dfaf82d5f23?t=1759261267112>

emociones; (v) intervenir o apoyar procesos relacionados con la prestación de servicios a niñas, niños y adolescentes (NNA), minorías étnicas, personas con discapacidad, migrantes u otras poblaciones en situación de vulnerabilidad; (vi) predecir o diagnosticar enfermedades, o (vii) realizar acompañamiento a personas con enfermedades mentales, los sujetos obligados en el artículo 5 de la Ley 1712 de 2014” (artículo 13, C).

También, prevé que las entidades públicas obligadas en razón a la Directiva, deben publicar información sobre “análisis de impacto o medidas de mitigación de riesgos”, pero esta obligación general no detalla si se trata de análisis de impacto enfocados en derechos humanos, o no (artículo 22).

Punto C. ¿Qué procesos internos y estructuras institucionales existen para garantizar la protección de los derechos humanos durante todo el ciclo de vida del sistema, incluida su revisión, modificación o interrupción?

A la fecha, no hemos documentado la existencia de procesos o procedimientos de esta naturaleza en el marco de las estructuras institucionales de las entidades públicas que en la actualidad despliegan sistemas de IA en los Estados.

Punto D. ¿Qué medidas han adoptado los Estados para fortalecer las capacidades técnicas, jurídicas y éticas de las personas funcionarias que desarrollan, adquieren o utilizan sistemas de inteligencia artificial y de inteligencia artificial generativa, en particular en materia de derechos humanos, protección de datos y transparencia? ¿Existen programas de capacitación?

A la fecha, no hemos encontrado evidencia sobre el despliegue de esfuerzos de capacitación técnica, jurídica o ética liderada por los Estados, y cuyo contenido y resultados hayan sido informados de manera proactiva al público.

Punto E. ¿Existen instancias de auditoría externa por parte de organizaciones independientes, con capacidad técnica y acceso a la información necesaria, para supervisar el uso estatal de sistemas de inteligencia artificial, incluyendo la inteligencia artificial generativa? ¿Con qué frecuencia se llevan a cabo? ¿Cuáles han sido las consecuencias de las auditorías?

A la fecha, no hemos encontrado información pública que dé cuenta de la existencia de procesos de auditoría externa aplicada a sistemas de IA desplegada por los Estados en la región.

Como sea, vale la pena recordar que el Sistema de Reconocimiento Facial de Prófugos de la Ciudad de Buenos Aires, y que fue suspendido en 2023 en razón al litigio interpuesto por O.D.I.A y otras organizaciones de la sociedad civil, estaría suspendido hasta tanto no se lleve a cabo una auditoría sobre el funcionamiento de dicho sistema. Sin embargo, la auditoría en cuestión no ha sido ejecutada, entre otros, por problemas asociados a la protección de la propiedad intelectual de dicho sistema. Situación que ejemplifica uno de los principales obstáculos que enfrenta en la práctica la realización de auditorías de este tipo.¹¹⁵

¹¹⁵ Ver: <https://odia.ar/casos/reconocimiento-facial>

Punto F. ¿Qué mecanismos existen para investigar violaciones de derechos humanos derivadas del uso de sistemas de inteligencia artificial y inteligencia artificial generativa? ¿Han existido investigaciones? ¿Cuáles han sido los resultados?

Por ahora, nos hemos identificado la existencia de mecanismos específicos enfocados en la investigación de violaciones a los derechos humanos derivadas del uso de sistemas de inteligencia artificial e inteligencia artificial generativa.

Sin embargo, la Directiva sobre transparencia algorítmica vigente en Colombia, prevé que en caso de que alguna persona o grupo de personas “considere que un sistema algorítmico utilizado por el Estado ha vulnerado sus derechos, pueden acudir a la Procuraduría General de la Nación y la Defensoría del Pueblo. Estas entidades, en ejercicio de sus funciones, podrán acompañarla o acompañarlos en el proceso de defensa y protección de sus derechos” (artículo 25). Sin que, para tales efectos, se hayan creado vías, acciones o mecanismos especiales de protección adicionales a los existentes.¹¹⁶

Punto G. ¿Existen mecanismos efectivos de apelación y protección judicial para las personas cuyos derechos se vean afectados por decisiones basadas en inteligencia artificial, incluso la generativa?

Hemos identificado y participado en calidad de *amicus curiae* ante distintos litigios que muestran que las herramientas jurídicas preexistentes para la satisfacción y reparación de derechos pueden, con sus dinámicas propias, dar algún tipo de respuesta de la institucionalidad a los accionantes que las usan. Sin embargo, dicha respuesta puede ser dilatada en el tiempo, y con ello, afectar la satisfacción en el acceso a la justicia de las personas afectadas.

Por ejemplo, en el caso promovido por O.D.I.A ante la justicia argentina en contra del sistema de reconocimiento facial, y en el que participamos en calidad de *amicus curiae*¹¹⁷, la justicia local resolvió por la suspensión de dicha tecnología, en aras de proteger los derechos de la ciudadanía. Sin embargo, aunque la respuesta de la justicia fue en su momento favorable a la protección de derechos de la ciudadanía, ciertas órdenes de la sentencia aún se encuentran incumplidas en tanto que, por ejemplo, la orden enfocada en autorizar la auditoría del sistema no ha sido cumplida por motivaciones asociadas a la protección del secreto comercial del que es titular la empresa desarrolladora de dicho sistema.¹¹⁸

Otro caso legal que identificamos en relación con este ítem, fue el promovido por Dejusticia¹¹⁹ en Colombia para exigir el acceso al código fuente de la aplicación CoronApp. La acción de tutela fue decidida por la Corte Constitucional casi cinco años después de haber sido impulsada dicha acción legal. Si bien su contenido es favorable para el acceso a la información pública frente al desarrollo y despliegue de una tecnología con impacto en la salud pública, su

¹¹⁶ Procuraduría General de la Nación; Defensoría del Pueblo (Septiembre 30, 2025). Estándares sobre transparencia algorítmica para los sistemas algorítmicos utilizados por el Estado. Disponible en: <https://www.defensoria.gov.co/documents/20123/3407303/300925DirectivaConjunta007.pdf/c47e1175-6f60-058a-3e0b-3dfaf82d5f23?t=1759261267112>

¹¹⁷ Ver: <https://srfp.odia.legal/dd.pdf>

¹¹⁸ Ver: <https://odia.ar/casos/reconocimiento-facial>

¹¹⁹ Dejusticia (Octubre, 2021). Intervinimos en dos casos ante la Corte Constitucional relacionados con la aplicación CoronApp. Disponible en: <https://www.dejusticia.org/litigation/intervinimos-en-dos-casos-ante-la-corte-constitucional-relacionados-con-la-aplicacion-coronapp/>

resultado no impactó necesariamente el contexto social que motivó la interposición de dicha acción.

4. Desarrollo, provisión y uso de inteligencia artificial por parte de actores privados

Punto F. ¿Ha habido negativas por parte de las empresas que desarrollan sistemas de inteligencia artificial, incluyendo la generativa, a proporcionar información sobre el desarrollo y el uso de dichos sistemas? ¿Cuáles fueron las justificaciones? ¿Se han mencionado alguna vez la propiedad intelectual o el secreto industrial como razones para no divulgar información sobre los sistemas de inteligencia artificial?

En este sentido, remitimos a las respuestas del punto 1C, así como a las investigaciones allí descritas que enfrentaron en general problemas de opacidad por deficiencias en la aplicación de las regulaciones sobre acceso a la información pública. Llamamos la atención sobre la opacidad argüida no solo por empresas sino por los Estados que las protegen, ambos amparados en razones que favorecen el secreto por motivos de propiedad intelectual, o simplemente por la entrega de información insuficiente, parcial, deficiente o incompleta.

En dicho informe de balance “Inteligencia artificial en el Estado: Estudio colectivo sobre experiencias y riesgos para los derechos humanos”¹²⁰, destacamos la experiencia de Colombia, a propósito del despliegue de Fiscal Watson, de propiedad de IBM, pues los investigadores se enfrentaron a problemas de acceso a la información pública pues la Fiscalía General de la Nación, al responder a los pedidos de acceso a la información, señalaba la necesidad de reservar la información requerida por la protección del secreto comercial y la propiedad intelectual de IBM, razón por la cual no era posible conceder detalles a los investigadores sobre el funcionamiento de la herramienta.

La experiencia de Colombia, en el litigio por el acceso al código fuente de la aplicación para el rastreo del Covid-19 también ejemplifica este escenario¹²¹. La sentencia de la Corte Constitucional, que favoreció el acceso a la información requerida (código fuente de CoronApp) por investigadores de Dejusticia, y que habría sido negada por el Estado por razones de reserva amparada en la protección del secreto comercial (secreto sobre una herramienta que era de propiedad del Estado), condujo en dicho contexto a la afirmación del deber de transparencia algorítmica a cargo de las autoridades públicas que desarrollen tecnologías.

Sin embargo, la negativa inicial de las entidades accionadas (Agencia Nacional Digital, Ministerio de Salud, Instituto Nacional de Salud) condujo a un litigio extenuante de más de cinco años de duración, donde la reserva y la opacidad se sobrepuso a la transparencia proactiva según lo dictan estándares interamericanos en esta materia. Al respecto, llamamos la atención de la RELE para que en el abordaje de la reserva de la información que protege los intereses del sector privado, llame la atención sobre los deberes de las entidades públicas en la generación de condiciones contractuales favorables a la transparencia proactiva.

¹²⁰ García, JM (2024). Inteligencia artificial en el Estado: estudio colectivo sobre experiencias y riesgos para los derechos humanos. Derechos Digitales. Disponible en: https://ia.derechosdigitales.org/wp-content/uploads/2025/02/2024-LATAM-IA_en_el-Estado-ES.pdf

¹²¹ Camacho, OL (Marzo, 2025). Coronapp: Datos públicos y lección que deja la Corte. La Silla Vacía. Disponible en: <https://www.lasillavacia.com/red-de-expertos/red-de-democracia-y-tecnologia/coronapp-datos-publicos-y-la-leccion-que-deja-la-corte/>

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

Punto 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?

En el aporte que remitimos a una consulta elevada por las Relatorías Conjuntas sobre Libertad de Expresión en agosto de 2025 “Artificial Intelligence and Freedom of Expression: Contribution to the Joint Declaration Consultation”¹²², abordamos en mayor extensión varios elementos asociados a esta pregunta.

En dicha contribución desagregamos la naturaleza de los impactos y retos para la libertad de expresión y de prensa en tres ítems: (i) retos en la capa de recopilación y procesamiento de datos para entrenar sistemas de IA e IA generativa, (ii) retos en la capa de generación de contenidos (texto, imágenes, audio), (iii) la capa predictiva y de reconocimiento de patrones.

En la primera capa, el impacto inmediato y directo para la libertad de prensa tiene que ver con el uso no autorizado, no remunerado y no reconocido de contenidos periodísticos para entrenar sistemas de IA e IA generativa, especialmente los de fines comerciales. Esta práctica agrava la ya precaria situación económica de los medios, cuyos modelos de negocio han sido debilitados por las grandes tecnológicas que, en muchos casos, también desarrollan los modelos generativos más utilizados.

En tanto que, para la libertad de expresión, el reto inmediato se asocia a la explotación intensiva de datos de usuarios -a menudo considerados “públicos”- se ha utilizado para ampliar de forma indebida los conjuntos de entrenamiento de IA/IA generativa. Esto genera riesgos graves como vigilancia estatal o corporativa, erosión de la privacidad en línea (clave para la libertad de expresión) y una disminución general de la expectativa de privacidad en el espacio público digital.¹²³

En la segunda capa, identificamos que existe el riesgo creciente de reemplazo del trabajo humano en las redacciones. Algunos medios ya utilizan GenAI para asistir o sustituir periodistas en la producción de noticias y análisis, lo que profundiza la crisis laboral del sector; también, que la adopción acrítica de GenAI puede debilitar estándares esenciales del periodismo como la veracidad, precisión, contextualización, oportunidad y calidad del contenido; además, que el empleo de GenAI para fact-checking es riesgoso debido a problemas como alucinaciones, reproducción de sesgos y pérdida de precisión ante consultas complejas. Además, la sustitución del consumo de noticias tradicionales por contenidos generados por IA amenaza la percepción pública de la prensa como pilar de la democracia y pone en riesgo el pluralismo mediático.¹²⁴

Y por último, en la tercera capa, identificamos como un riesgo actual y presente, el uso de IA y GenAI por parte de gobiernos para clasificar, monitorear, predecir y perfilar el discurso público

¹²² Derechos Digitales (August, 2025). Artificial Intelligence and Freedom of Expression: Contribution to the Joint Declaration Consultation. Disponible en: <https://www.derechosdigitales.org/wp-content/uploads/2025/08/Artificial-Intelligence-and-Freedom-of-Expression-Contribution-to-the-Joint-Declaration-Consultation.pdf>

¹²³ Ibid.

¹²⁴ Ibid.

-especialmente de voces críticas- representa una amenaza creciente, observada en partes de América Latina. Estas prácticas pueden venir acompañadas de medidas coercitivas o policiales y derivar en censura, autocensura, restricción de ciertas expresiones y estigmatización de grupos o movimientos políticos.¹²⁵

Punto 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?

En el informe que presentamos conjuntamente con APC ante el Mecanismo Experto Conjunto sobre el Derecho al Desarrollo de ONU, enfocamos la atención en el impacto de la IA en los derechos culturales.¹²⁶ En dicho informe expresamos que el potencial de las tecnologías digitales, especialmente la inteligencia artificial (IA), para promover los derechos económicos, sociales y culturales ha sido ampliamente debatido.

Para que las sociedades se beneficien plenamente, estas tecnologías deben diseñarse y aplicarse conforme a principios de derechos humanos como igualdad, no discriminación, participación, transparencia, accesibilidad, asequibilidad y protección de la privacidad.

El impacto de la IA en los derechos culturales es amplio y se manifiesta en cuatro áreas principales:

- **Preservación de lenguas:** La IA puede apoyar la preservación y revitalización lingüística si se basa en enfoques inclusivos y de derechos humanos. Sin embargo, modelos de lenguaje y herramientas de IA pueden documentar idiomas en peligro, generar contenidos en lenguas locales y facilitar su transmisión a nuevas generaciones.
- **Creación y expresión artística:** La IA influye directamente en la creatividad, la libertad artística y el derecho a participar en la vida cultural. No obstante, la creatividad humana sigue siendo irreproducible por la IA, que solo recombina estilos existentes. Esto genera preocupaciones en un contexto donde el trabajo artístico ya es precario y donde muchas herramientas de IA están impulsadas por fines comerciales.
- **Acceso al patrimonio cultural:** La IA abre nuevas formas de interactuar con el patrimonio cultural mediante experiencias digitales inmersivas, reconstrucciones virtuales y mejor catalogación. Sin embargo, estos beneficios dependen de la participación significativa de las comunidades en el diseño de las herramientas y en la gestión de los datos patrimoniales, para evitar la homogeneización cultural y proteger sus intereses.
- **Participación en sistemas de conocimiento:** Internet y la IA pueden democratizar el acceso al conocimiento, facilitando aprender, investigar y crear. Iniciativas previas ya han promovido bibliotecas digitales abiertas. No obstante, la IA reproduce los sesgos de sus datos de entrenamiento, lo que puede generar contenidos parciales o discriminatorios. Dado que produce información a gran escala, es crucial garantizar que represente la diversidad cultural y contribuya a un acceso equitativo al conocimiento y al bienestar social.

¹²⁵ Ibid.

¹²⁶ EMRTD study “Artificial Intelligence, Cultural Rights, and the Rights to Development”, Association for Progressive Communications APC and Derechos Digitales - Joint Submission for the Expert Mechanism on the Right to Development (November, 2025). Disponible en: https://www.derechosdigitales.org/wp-content/uploads/2025/11/Joint-Submission_Cultural-Rights-and-AI-Submission-APC-and-DD.pdf

Adicionalmente, en nuestro informe conjunto identificamos cómo el sesgo algorítmico representa un riesgo significativo para los derechos culturales porque los sistemas que determinan qué se ve, se escucha y se valora en línea están lejos de ser culturalmente neutrales o equitativos. Cuando los algoritmos de recomendación, búsqueda, traducción o etiquetado de contenidos se construyen sobre conjuntos de datos y modelos que reflejan predominantemente normas occidentales, masculinas y anglófonas, incorporan desde el inicio una jerarquía estructural de culturas e identidades.¹²⁷

Además, cuando los datos de entrenamiento omiten conocimientos locales o tradiciones visuales, y cuando los modelos promueven representaciones sesgadas o ignoran por completo las lenguas minoritarias, se pone en riesgo la continuidad de la transmisión cultural. Con el tiempo, esto puede afectar la autodeterminación de los pueblos, la participación significativa en el desarrollo y la representación cultural equitativa.¹²⁸

Llamamos la atención de la RELE para abordar en un informe especializado en la materia el impacto de la IA en la protección de los derechos culturales que incluya la consulta a comunidades y grupos étnicos y racializados que incluya, no solo una perspectiva sobre el impacto visto desde la generación de contenidos y las representaciones culturales que éstas implican, sino sobre el impacto visto desde el cambio climático en tanto que tecnologías digitales con potencial para agravar la crisis climática.

Punto E. 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?

Este es un aspecto ampliamente documentado por organizaciones de la sociedad civil en la región. En el informe “La moderación de contenidos”¹²⁹ del Consorcio AlSur, del cual hacemos parte (y que ya ha sido socializado en el pasado con la RELE), describimos los impactos de la moderación automatizada y el impacto de la moderación en los discursos protegidos, no protegidos, la responsabilidad de intermediarios en internet, y principios y recomendaciones deseables en materia de transparencia que han sido desarrollados por la sociedad civil hasta ahora.

Punto F. ¿Cuál ha sido el impacto del uso de sistemas de inteligencia artificial en el acceso a la información por parte de la ciudadanía?

Al respecto, sugerimos consultar el ítem 1K sobre *chatbots*, que son por ahora, el instrumento más empleado por los Estados para facilitar el ejercicio del derecho de acceso a la información a través de la implementación de sistemas de IA.

Punto G. 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*)?

¹²⁷ Ibid.

¹²⁸ Ibid.

¹²⁹ Chorny, V.; García, L.; Macías, G. (2022). La moderación de contenidos desde una perspectiva interamericana. AlSur. Disponible en: https://www.alsur.lat/sites/default/files/2025-10/moderacion_contenidos_alsur_.pdf

Por ahora, observamos muy limitados avances en la materia en la región de esfuerzos estatales y del sector privado enfocados en contrarrestar la generación y difusión de desinformación y, más reciente, de *deepfakes*.

En Chile, por ejemplo, los resultados de la Comisión sobre la Desinformación que condujeron a la producción de un informe estudiando el fenómeno a nivel local y global, y otro enfocado en las recomendaciones para contrarrestar la desinformación, generaron entusiasmo en su momento. El Estado habría enfocado esfuerzos de política pública para el análisis de un fenómeno complejo para que fuesen delimitados objetivos accionables a nivel regulatorio. Sin embargo, a la fecha, las setenta y dos recomendaciones elevadas por la Comisión sobre la Desinformación reporta muy limitados avances, por ejemplo, en lo que tiene que ver con la implementación de regulación de las plataformas digitales con un enfoque de derechos humanos para, por ejemplo, transparentar el impacto de las decisiones automatizadas en la moderación de contenidos.

En materia electoral, identificamos en un informe que remitimos a la Relatoría Especial sobre Libertad de Expresión y Opinión de ONU, titulado “Threats and Challenges to freedom of expresison during elections in Latin American Countries”¹³⁰ que, tratándose de contener los efectos de la generación de *deepfakes* que afectan a mujeres candidatas, los Estados siguen sin avanzar significativamente en la materia.

En ese informe identificamos casos de afectación a mujeres candidatas en México, Paraguay y Brasil que, frente a la generación de *deepfakes* que afectaban sus candidaturas, se vieron sin apoyo de las autoridades electorales ni de las plataformas digitales en las cuales circulaban contenidos que afectaba su privacidad, autonomía y libertad sexual.¹³¹

Así como identificamos eventos sistemáticos de violencia de género facilitada por las tecnologías en contextos electorales en países como Panamá, Costa Rica, Honduras y Colombia. En Colombia se detalla cómo incluso las mujeres candidatas no cuentan ni siquiera con el propio apoyo de sus partidos políticos para hacer frente a dicho fenómeno; y en los casos de Panamá, Honduras y Costa Rica se da cuenta de la ausencia de articulación entre los Estados y las plataformas, dejando a las mujeres candidatas en escenarios que las forzaron a retirar sus candidaturas.¹³²

La actuación de los Estados para contrarrestar los *deepfakes* se enfrenta también a la inacción o resistencia de las Big Tech. Como mencionamos en el ítem 1D, países como Brasil actuaron recientemente frente a la generación y difusión en masa de falsificaciones profundas que circularon en la red social X, y que afectaron a mujeres, niños, niñas y adolescentes, a través de una recomendación conjunta¹³³ emitida por la ANPD, Senacon y el Ministerio Público Fiscal. La

¹³⁰ Derechos Digitales, TEDIC, IPANDETEC, Karisma Foundation (January, 2025). Submission to the UN Special Rapporteur on Freedom of Opinion and Expression: Thematic report on “Freedom of Expression and Elections in the Digital Age”, *Threats and challenges to freedom of expression during elections in Latin American Countries*. Disponible en: https://www.derechosdigitales.org/wp-content/uploads/Submission_Foe-and-Elections_Derechos-Digitales-Karisma-TEDIC-IPANDETEC.pdf

¹³¹ Ibid.

¹³² Ibid.

¹³³ MPF, ANPD e Senacon recomendam que X impeça geração e circulação de conteúdos sexualizados indevidos por meio do Grok (Janeiro 20, 2026). Disponible en: <https://www.mpf.mp.br/o-mpf/unidades/pr-sp/noticias/mpf->

plataforma respondió al Estado que X Brasil no tendría conexión legal con X, la red social, por lo que no actuarían¹³⁴ ante el pedido de las autoridades que requería, entre otros, limitar acceso a dicha herramienta, así como supervisar su uso por parte de la plataforma.

Autoridades como la Superintendencia de Industria y Comercio en Colombia, también reaccionaron frente al mismo evento de difusión y generación masiva de *deepfakes* en X con una aproximación mucho más tímida desde el punto de vista regulatorio: con su adhesión a una alerta internacional¹³⁵ a la que adhirieron también 60 autoridades de protección de datos de distintos países. La alerta destaca “varias expectativas dirigidas a las organizaciones que desarrollan o utilizan IA generativa”, que incluyen por ejemplo, acciones enfocadas en “Abordar los riesgos específicos para niñas y niños, implementando salvaguardas reforzadas e incluyendo información clara para niños, padres, tutores y educadores”¹³⁶, entre otros.

En materia de integridad de la información, también observamos¹³⁷ las iniciativas estatales para, por ejemplo, abordar la desinformación climática. En la región, la iniciativa liderada en 2024 por Brasil¹³⁸ en el marco del G20 y con el apoyo de UNESCO y la ONU sientan enfoques favorables en la región para alentar la acción multilateral. También, y en el marco de la COP30, que tuvo lugar en Brasil¹³⁹ en 2025 y que condujo a la Declaración sobre la Integridad de la Información sobre Cambio Climático apuntan, nuevamente, a llamar la atención en la dirección correcta al convocar a los Estados y las plataformas digitales a la acción conjunta. Sin embargo, a la fecha, identificamos un muy limitado impacto de estas iniciativas en la generación de acciones locales de política pública en los países que han suscrito o adherido a dichas iniciativas.

[anpd-e-senacon-recomendam-que-x-impeca-geracao-e-circulacao-de-conteudos-sexualizados-indevidos-por-meio-do-grok](#)

¹³⁴ Dias, T. (February, 2026). X Tried to Sidestep Brazil’s Inquiry on AI Deepfakes- The Government Just Pushed Back. Tech Policy Press. Disponible en: <https://www.techpolicy.press/x-tried-to-sidestep-brazils-inquiry-on-ai-deepfakes-the-government-just-pushed-back/>

¹³⁵ Superintendencia de Industria y Comercio (Febrero, 2026). SIC lidera alerta internacional por riesgos de la IA generativa en la creación de imágenes íntimas manipuladas y no consentidas. Disponible en: <https://sedeelectronica.sic.gov.co/noticias/sic-lidera-alerta-internacional-por-riesgos-de-la-ia-generativa-en-la-creacion-de-imagenes-intimas-manipuladas-y-no-consentidas>

¹³⁶ Ibid.

¹³⁷ Camacho, OL. (Noviembre, 2025). Integridad de la información en tiempos de emergencia climática. Razón Pública. Disponible en: <https://razonpublica.com/integridad-la-informacion-tiempos-emergencia-climatica/>

¹³⁸ UNESCO (November, 2024). G20: Leader’s Summit: Brazil, UNESCO and UN launch Global Initiative for Information Integrity on Climate Change. Disponible en: <https://www.unesco.org/en/articles/g20-leaders-summit-brazil-unesco-and-un-launch-global-initiative-information-integrity-climate>

¹³⁹ Ver: https://www.unesco.org/sites/default/files/medias/fichiers/2025/11/2025.11.12%20-%20COP%2030%20-%20%20Declaration%20on%20Information%20Integrity%20on%20Climate%20Change_final_ok.pdf?hub=780

6. Lagunas, buenas prácticas y recomendaciones

Punto 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?

Por ahora, observamos cambios a nivel regulatorio que son mayormente fragmentados. En la región, por ahora, solo Perú cuenta con una Ley generalista de IA, el resto de países continúan tramitando proyectos de naturaleza similar.

Ese país, así como el resto, según nuestro monitoreo normativo, han introducido reformas específicas a cuerpos normativos preexistentes al menos con dos enfoques (i) uno, mayoritario, autorizando y promoviendo el uso de IA en el sector público, para ciertos fines y objetivos de política pública¹⁴⁰, o (ii) o castigando su uso —enfoque minoritario— con una perspectiva exclusivamente punitivista, como es el caso de las reformas que castigan en Perú (Ley 32.314), Colombia (Ley 2502/2025) y Brasil (Ley 15.123) el uso de *deepfakes* como un agravante en los delitos de suplantación personal.

Sin embargo, las reformas regulatorias introducidas por ahora, no abordan a profundidad los problemas que el impacto de dicho uso significa para el ejercicio en derechos (ver punto 6B, y 6C).

Punto B. ¿Cuáles considera que son las mejores prácticas, por parte de los Estados, del sector privado y de la sociedad civil para garantizar la transparencia, la participación y la rendición de cuentas en el desarrollo y uso de sistemas de inteligencia artificial, incluso de inteligencia artificial generativa?

En el marco de procesos de consulta impulsados previamente por organismos internacionales y regionales de derechos humanos¹⁴¹, hemos formulado una serie de recomendaciones para garantizar la transparencia, la participación y la rendición de cuentas en el desarrollo y uso de sistemas de IA por parte de los Estados. Dichas recomendaciones se presentan a continuación:

Las mejores prácticas por parte de los Estados:

- Implementar moratorias o prohibiciones frente al uso de herramientas de IA incompatibles con el ejercicio de derechos y libertades fundamentales, o que no puedan ser utilizadas en cumplimiento con el derecho internacional de derechos humanos. En

¹⁴⁰ Un registro público de este monitoreo lo ofrece la Universidad de los Andes en: https://algoritmos.uniandes.edu.co/regulacion-sobre-ia-en-america-latina/?_riaalc_subregion_de_america=america-del-sur&_riaalc_rama_del_poder_publico=legislativo&_riaalc_tipo_de_instrumento_regulatorio=ley

¹⁴¹ Camacho Gutiérrez, L.; Lara-Castro, P. (Noviembre, 2025). Contribución a la consulta para el informe temático titulado “Impacto de la vigilancia digital y asistida por IA en los derechos de reunión y de asociación, incluidos los efectos inhibitorios”, convocado por la Relatoría Especial sobre el Derecho de Reunión y Asociación Pacífica. Uso de sistemas de reconocimiento facial y ciberpatrullaje en contextos de protesta en países de América Latina. Derechos Digitales. Disponible en: <https://www.derechosdigitales.org/wp-content/uploads/2025/11/Derechos-Digitales-Submission-on-AI-and-FoAA.pdf> y CELS, Derechos Digitales, Fundación Vía Libre (2025). Aporte conjunto a la ‘Convocatoria de Aportaciones sobre los Impactos en los Derechos Humanos del Uso de la Inteligencia Artificial en la Lucha contra el Terrorismo’, de la Relatoría Especial sobre la lucha contra el terrorismo y los derechos humanos. Disponible en: https://www.derechosdigitales.org/wp-content/uploads/2025/08/Informe_OHCHR-contra-terrorismo-e-IA_CELS-Derechos-Digitales-Fundacion-Via-Libre.pdf

especial, de acuerdo con el informe de la Alta Comisionada de las Naciones Unidas para los Derechos Humanos sobre el derecho a la privacidad en la era digital¹⁴², urge introducir moratorias sobre el uso de tecnologías de reconocimiento facial en espacios públicos hasta que las autoridades responsables demuestren públicamente que la tecnología respeta la privacidad y la protección de los datos de las personas ciudadanas.

- Desarrollar y publicar evaluaciones de impacto en derechos humanos comprehensivos y regulares al diseño, desarrollo, compra, despliegue y operación de sistemas de IA utilizados en la función pública, considerando sus potenciales efectos en el ejercicio de derechos fundamentales de las personas ciudadanas.
- Incorporar procesos participativos activos, abiertos y diversos en el diseño de las políticas públicas que buscan implementar sistemas de IA en el Estado para que distintos grupos poblacionales, en especial los más impactados y aquellos grupos históricamente en condición de vulnerabilidad, puedan involucrarse en los procesos de toma de decisiones públicas.
- Adoptar medidas de transparencia activa sobre los sistemas de IA que los Estados se encuentran adoptando, ahondando en los propósitos que dichos sistemas buscan satisfacer, así como identificar a los desarrolladores y operadores de estas tecnologías. Además, informar a las personas impactadas por la toma de decisiones automatizadas detallando sobre el uso de sus datos por sistemas de IA.
- Adoptar el contenido de la Resolución de la Asamblea de las Naciones Unidas A/HRC/50/56¹⁴³, que establece obligaciones específicas para los Estados que adquieren, compran, licencian o tercerizan actividades de la función pública en empresas privadas del sector tecnológico.

Las mejores prácticas por parte del sector privado:

- Transparentar los acuerdos que sostienen con los Estados para la oferta de sistemas de IA a ser utilizados en el ejercicio de la función pública. En ese contexto, es crucial informar sobre el valor de dichos acuerdos comerciales, el objetivo de su suscripción, el tiempo de duración del contrato, así como describir la tecnología proveída
- En concordancia con lo establecido en el Protocolo de la Esperanza¹⁴⁴, abstenerse de diseñar, desarrollar, producir y vender tecnologías que puedan ser utilizadas por gobiernos o entidades privadas para reprimir derechos y libertades fundamentales, así como el ejercicio de su defensa. En ese sentido, las empresas deben actuar con la debida diligencia en materia de derechos humanos, lo que implica no sólo la transparencia en la venta de sus tecnologías, sino también la aplicación de estudios exhaustivos que midan las posibles consecuencias de sus tecnologías sobre las personas.
- Inhibirse de suscribir acuerdos comerciales con Estados que no cuenten con regulaciones adecuadas en materia de derechos humanos, por ejemplo, en los campos de privacidad y protección de datos, protección del derecho a la protesta pacífica,

¹⁴² Alta Comisionada de las Naciones Unidas para los Derechos Humanos (2021). El derecho a la privacidad en la era digital. A/HRC/48/31. Disponible en: <https://docs.un.org/es/A/HRC/48/31>

¹⁴³ Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (2022). La aplicación práctica de los Principios Rectores sobre las Empresas y los Derechos Humanos a las actividades de las empresas tecnológicas. A/HRC/50/56. Disponible en: <https://docs.un.org/es/A/HRC/50/56>

¹⁴⁴ Para más información, ver: <https://esperanzaprotocol.net/wp-content/uploads/2025/05/PLE-Digital-Espanol.pdf>

protección de la libertad de expresión; o con Estados que no cumplan con el derecho internacional de derechos humanos¹⁴⁵.

- Desplegar y publicar evaluaciones de impacto en privacidad y derechos humanos de sus tecnologías, en especial las que son potencialmente empleadas, adquiridas o licenciadas por Estados.
- Adoptar el contenido de la Resolución de la Asamblea de las Naciones Unidas A/HRC/50/56¹⁴⁶, que consagra acciones específicas para las empresas que desarrollan, diseñan, venden o licencian sistemas de IA apropiados por los Estados.
- Divulgar información clave sobre el funcionamiento de los sistemas de IA que desarrollan, proporcionando explicaciones accesibles sobre cómo influyen en la toma de decisiones que afectan a las personas. Estas medidas deben complementarse con auditorías independientes que permitan evaluar y supervisar estos sistemas.

Las mejores prácticas por parte de la sociedad civil:

- Establecer grupos de trabajo que permitan avanzar en el monitoreo, estudio, documentación, evaluación e intercambio de información en torno a casos de despliegue de sistemas de IA por los Estados.
- Desplegar sesiones informativas y de intercambio de conocimiento para ampliar capacidades y conocimiento técnico mutuo sobre el diseño, desarrollo, compra, despliegue y operación de sistemas de IA utilizados en la función pública.
- Apropiar los mecanismos del Sistema Interamericano de Derechos Humanos (como el involucramiento con el SIMORE, las audiencias temáticas, el envío de casos ante el Sistema Interamericano de Protección de los Derechos Humanos) para fortalecer las capacidades de incidencia a nivel regional de cara al uso de sistemas de IA por los Estados.

Punto C. ¿Qué lagunas normativas, institucionales o técnicas persisten y qué recomendaciones propondría para abordarlas?

Sugerimos ver el punto inmediatamente anterior. Adicionalmente, evidenciamos en nuestro monitoreo regulatorio y de políticas públicas:

- Ausencias normativas frente al impacto medio ambiental de la infraestructura digital que hace posible el funcionamiento de la IA, en concreto, de los centros de datos (y la flexibilización de las normativas sobre licenciamiento ambiental en la región que facilitan su instalación por encima de la consulta previa con comunidades, etc.).
- Ausencias normativas en lo que tiene que ver con el balance entre tecnologías de vigilancia articuladas con sistemas de IA, su inserción en el ámbito de la seguridad ciudadana, la lucha contra el terrorismo –entre otros- y la protección de derechos.
- La articulación coordinada institucional a nivel interamericano para responder a retos como los que enfrenta la integridad de la información, en especial en tiempos de elecciones; para enfocar esfuerzos en la protección de niños, niñas y adolescentes; para

¹⁴⁵ De acuerdo con el el Protocolo de la Esperanza, disponible en:

<https://esperanzaprotocol.net/wp-content/uploads/2025/05/PLE-Digital-Espanol.pdf>

¹⁴⁶ Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (2022). La aplicación práctica de los Principios Rectores sobre las Empresas y los Derechos Humanos a las actividades de las empresas tecnológicas. A/HRC/50/56. Disponible en: <https://docs.un.org/es/A/HRC/50/56>

enfocar esfuerzos concernientes a abordar los impactos de la IA en la reproducción y amplificación de formas de violencia de género facilitadas por la tecnología; para enfocar esfuerzos enfocados en la protección de la privacidad de la ciudadanía y de grupos especialmente vulnerables (grupos de migrantes, grupos étnicos y racializados, grupos y comunidades genero-sexo disidentes, entre otros).

- La debilidad regulatoria de las autoridades de protección de datos que no cuentan con la capacidad técnica ni humana para emitir decisiones frente al tratamiento de datos de las Big Tech y el continuo cambio de sus políticas que afecta el ejercicio de derechos como el consentimiento informado.
- La debilidad técnica y de capacidades humanas de las instituciones de investigación judicial (jueces, fiscales) para avanzar casos de violencia de género facilitada por tecnologías como los *deepfakes*. Ausencia, así mismo, de marcos regulatorios con perspectiva de género que aborden dicho fenómeno, más allá de la visión exclusivamente criminalizante.
- La ausencia de esfuerzos enfocados en la alfabetización sobre IA de funcionarios públicos, la ausencia de incentivos para la transparencia —funcionarios que usan sistemas de IA sin informar al público de ello, por temor la escarnio público—, entre otros.

Punto D. ¿Puede citar ejemplos de sistemas de inteligencia artificial e inteligencia artificial generativa que se estén utilizando para promover los derechos humanos, en particular en el ámbito de la libertad de expresión? Por ejemplo, en la audiencia de oficio celebrada por la CIDH en marzo de 2025 se mencionó la existencia de sistemas de inteligencia artificial para apoyar mujeres víctimas de violencia.

En la investigación de Derechos Digitales titulada “Latin America in a Glimpse 2024: Reflexiones para una IA comunitaria y feminista”¹⁴⁷, documentamos ejemplos de sistemas de inteligencia artificial que no sólo potencian el ejercicio de los derechos humanos, sino que están diseñados desde metodologías críticas que priorizan la transparencia, participación, inclusión, equidad y privacidad.

Destacamos tres proyectos que, desde los feminismos, buscan responder a tiempo ante la violencia y la desinformación basada en género que atraviesa los entornos digitales, así como recolectar información sobre violencias de género apoyando a activistas y organizaciones en sus propias prácticas de recolección de datos.

Sof-IA (Sistema de Oída Feminista)¹⁴⁸, es un prototipo de agente conversacional (tipo *chatbot*)¹⁴⁹ que responde y aborda situaciones de violencia de género facilitada por la tecnología (VGFT) particularmente en Chile. Sof-IA nació para apoyar y potenciar la capacidad de respuesta inmediata frente a situaciones de violencia digital de género. Con base en principios feministas y protegiendo la libertad de expresión y opinión, esta IA asiste activistas, personas con voz pública, mujeres periodistas, personas de la comunidad LGTBQIA+ que enfrenten violencias en línea. Los datos obtenidos son sistematizados para la incidencia en políticas públicas y acciones colectivas que aborden la VGFT en dicho país.

¹⁴⁷ Consultar en: https://www.derechosdigitales.org/wp-content/uploads/Glimpse_2024_ESP.pdf

¹⁴⁸ Desarrollado por la Fundación Datos Protegidos, la Escuela de Diseño de la Universidad Católica de Chile y el Observatorio de Estadísticas de Género e Interseccionalidad.

¹⁴⁹ Para más información sobre SOF+IA, ver: <https://sofiachat.cl/>

AymurAI¹⁵⁰ es un proyecto de código abierto destinado a fortalecer las bases de datos judiciales en temas de violencia de género, con herramientas para anonimizar sentencias, facilitando así el acceso a datos que informen políticas y acciones en pro de la equidad de género”. Este modelo emplea la información encontrada en sentencias judiciales en América Latina con el fin de reconocer los patrones de violencia que eventualmente pueden desembocar en feminicidios. Actualmente, AymurAI se encuentra en fase de implementación en el Juzgado Penal Contravencional y de Faltas de la Ciudad de Buenos Aires, así como en el Juzgado Federal de Primera Instancia de Oberá. Asimismo, la herramienta tiene como objetivo robustecer los registros judiciales relacionados con violencia de género y contribuir al diseño de políticas públicas orientadas a promover la equidad de género.

E.D.I.A (Estereotipos y Discriminación en Inteligencia Artificial)¹⁵¹, es una herramienta diseñada para que cualquier persona, sin necesidad de conocimientos en programación, pueda comparar oraciones y palabras que contienen estereotipos de género. Su principal objetivo es ofrecer una primera alfabetización crítica sobre el uso de la IA generativa, creando de manera colaborativa un conjunto de datos para evaluar estas tecnologías, contribuyendo a construir una ciudadanía consciente en estas temáticas. El sistema busca que las comunidades locales registren sus experiencias de discriminación como antecedentes para auditar las tecnologías del lenguaje automatizado, detectando y caracterizando los comportamientos discriminatorios y los discursos de odio presentes en los entornos digitales.

De otra parte, hemos documentado herramientas de IA que buscan la preservación de lenguas indígenas de América Latina, permitiendo a los hablantes nativos la libertad de expresión en la propia lengua. Como caso particular, resaltamos la herramienta Alkuaa¹⁵², una interfaz de programación de aplicaciones (API) abierta creada por el medio independiente El Surti, que transcribe audio en guaraní a texto. La iniciativa surge porque los grandes modelos de lenguaje (LLM) amplían la brecha digital al excluir lenguas orales como el guaraní por falta de datos. A través de Mingas (hackatones comunitarios de trabajo colectivo), se alimenta un repositorio abierto con audios en guaraní para entrenar el modelo de IA. El objetivo es que organizaciones y medios comprendan y respondan mejor a las audiencias guaraní hablantes.

Punto E. ¿Qué medidas han adoptado los Estados para promover el acceso equitativo a los beneficios de la inteligencia artificial en zonas rurales, con baja conectividad o entre grupos tradicionalmente excluidos? ¿Qué métricas utilizan para medir las brechas en el acceso digital y en el uso de sistemas de inteligencia artificial y de inteligencia artificial generativa entre distintos grupos poblacionales?

En cuanto al acceso equitativo a las tecnologías digitales, destacamos los persistentes desafíos que enfrenta América Latina para promover una conectividad significativa y garantizar el acceso universal a Internet. Tal como lo hemos manifestado en instancias de consulta impulsadas por organismos como la Unión Internacional de Telecomunicaciones (UIT)¹⁵³, la creciente atención que los Estados prestan a la inteligencia artificial podría desviar el foco de

¹⁵⁰ Para más información sobre AymurAI, ver: <https://www.aymurai.info/>

¹⁵¹ Para más información sobre E.D.I.A, ver: <https://huggingface.co/spaces/vialibre/edia>

¹⁵² Para más información, ver: <https://elsurti.com/pt/aikuaa/>

¹⁵³ Derechos Digitales (2024). Submission to the ITU Council Working Group on International Internet-related Public Policy Issues (CWG-internet) Online Open Consultation on “The developmental aspects to strengthen the Internet”. Disponible en: https://www.derechosdigitales.org/wp-content/uploads/Submission_Online-Open-Consultation_Derechos-Digitales.pdf

los esfuerzos esenciales destinados a reducir la brecha digital en la región. Por ello, resulta fundamental que las políticas públicas prioricen el desarrollo de estándares técnicos orientados a medir esta brecha, de modo que el avance y la adopción de la IA no terminen postergando, minimizando o invisibilizando la urgencia de implementar medidas concretas para lograr un acceso verdaderamente inclusivo.

Lo anterior cobra especial relevancia si se considera que, paralelamente al impulso estatal de herramientas de IA, persisten fenómenos como los cortes de internet, cada vez más frecuentes en América Latina¹⁵⁴. Ante esta realidad, resulta apremiante contar con directrices claras que permitan analizar las causas y el impacto de estas interrupciones en la conectividad, así como promover estándares para su documentación y monitoreo. Dichos esfuerzos deben involucrar no solo a los tomadores de decisiones, sino también a otros actores clave, incluyendo la sociedad civil, la comunidad técnica, y la academia.

En nuestro trabajo hemos identificado cómo la persistencia de la brecha digital, frente a la progresiva digitalización de los servicios públicos hoy potenciada por la inteligencia artificial, tiende a reforzar las desigualdades estructurales y a excluir a quienes aún carecen de un acceso significativo a Internet. Esta situación se agrava en regiones con territorios aislados, como las zonas selváticas, que permanecen profundamente desconectadas¹⁵⁵. Por ello, resulta imperativo prestar especial atención a las necesidades de acceso equitativo a tecnologías digitales en estas áreas.

Adicionalmente, consideramos oportuno incorporar un enfoque étnico y multicultural en la concepción del acceso tecnologías, incluida acá la IA, atendiendo las particularidades de su implementación en comunidades donde puede incidir directamente en las prácticas culturales y sociales locales. Un acceso que no contemple estas dimensiones corre el riesgo de reproducir lógicas de exclusión, en lugar de contribuir al desarrollo digital inclusivo y sostenible de la región.

Fin

¹⁵⁴ Derechos Digitales (2023). Una panorámica de los apagones de internet en América Latina. Derechos Digitales. En: <https://www.derechosdigitales.org/wp-content/uploads/Shutdowns-panorama.pdf>

¹⁵⁵ Lara Castro, P.; Souza, M. (2023). Latin America in a Glimpse. Acceso a internet en la región amazónica. Tendencias en los casos de estudio desde Brasil, Bolivia, Colombia y Ecuador. Derechos Digitales. En: https://www.derechosdigitales.org/wp-content/uploads/DD_Amazonia_5_General-3.pdf