



Asamblea General

Distr. general
13 de mayo de 2026
Español
Original: inglés

Consejo de Derechos Humanos

62º período de sesiones

15 de junio a 10 de julio de 2026

Temas 3 y 5 de la agenda

**Promoción y protección de todos los derechos humanos,
civiles, políticos, económicos, sociales y culturales,
incluido el derecho al desarrollo**

Órganos y mecanismos de derechos humanos

Repercusiones de los sistemas de inteligencia artificial en la buena gobernanza

Estudio del Comité Asesor del Consejo de Derechos Humanos*

Resumen

En el presente estudio, presentado de conformidad con la resolución 57/5 del Consejo de Derechos Humanos, el Comité Asesor del Consejo de Derechos Humanos examina las repercusiones de los sistemas de inteligencia artificial (IA) en la buena gobernanza.

El Comité Asesor analiza el modo en que la IA puede mejorar la prestación de servicios públicos, reforzar la transparencia y la lucha contra la corrupción y fomentar una participación inclusiva. Examina los riesgos para los derechos humanos que plantea la implantación de la IA en la gobernanza pública, en particular la vigilancia masiva, la discriminación algorítmica, la opacidad de la toma automatizada de decisiones, la manipulación del entorno informativo, el desplazamiento de puestos de trabajo y los daños ambientales. Examina asimismo el riesgo de que la concentración del desarrollo de la IA en un número limitado de actores afiance los patrones de dependencia digital y subraya la necesidad imperiosa de que todos los Estados participen de forma inclusiva y equitativa en la formulación de las normas y criterios aplicables.

El Comité Asesor formula recomendaciones encaminadas a que la implantación de la IA en la gobernanza pública se ajuste a las obligaciones internacionales en materia de derechos humanos y a los principios de la buena gobernanza.

* Se acordó publicar este documento tras la fecha prevista debido a circunstancias que escapan al control de quien lo presenta.



I. Introducción

A. Mandato

1. El Consejo de Derechos Humanos, en su resolución 57/5, solicitó al Comité Asesor del Consejo de Derechos Humanos que preparara un estudio sobre las repercusiones de los sistemas de inteligencia artificial (IA) en la buena gobernanza, indicando en particular los ámbitos en que esos sistemas podían contribuir a promover y proteger los derechos humanos mediante la buena gobernanza y aquellos en que suponían un reto para la buena gobernanza y los derechos humanos, destacando las buenas prácticas a nivel mundial de desarrollo, despliegue, uso y gestión de sistemas de IA para promover y proteger los derechos humanos mediante la buena gobernanza y determinando las garantías necesarias.

2. En su 32º período de sesiones, el Comité Asesor estableció un grupo de redacción integrado por Miznah Omair Alomair, Nicole Ameline, Joseph Gérard Angoh, Edward Amoako Asante, Buhm-Suk Baek (Relator), Rabah Boudache, Ignacio de Casas, Riva Ganguly Das, Alessandra Devulsky, Veronika Fikfak, Jewel Major, Afsaneh Nadipour, Ren Yisheng, Patrycja Sasnal, Vassilis Tzevelekos (Presidente) y Frans Viljoen.

B. Objeto del estudio

3. El presente estudio examina la integración de la IA en la gobernanza pública y sus repercusiones para los derechos humanos. Por consiguiente, no se centra en la gobernanza de la IA como tal, sino en el modo en que los sistemas de IA, al ser utilizados por las autoridades públicas, pueden promover u obstaculizar los derechos humanos. El estudio analiza las condiciones que deben darse para que la IA contribuya a que la gobernanza cumpla su promesa de eficiencia y transparencia. Advierte que, si no se fundamenta en las obligaciones de derechos humanos, la implantación de la IA puede crear nuevas vías para la comisión de violaciones sistemáticas de los derechos humanos que queden fuera del alcance de los mecanismos de reparación tradicionales.

4. El análisis repasa la evolución experimentada por el Consejo de Derechos Humanos, desde su resolución 7/11 (2008) hasta su resolución 57/5 (2024), en cuanto al marco que vincula la buena gobernanza a los derechos humanos, evolución que pone de manifiesto el creciente reconocimiento de que no se trata de dos ámbitos meramente complementarios, sino mutuamente constitutivos. Examina también el modo en que la IA pone a prueba la adaptabilidad de este marco consolidado. Si bien la IA y los sistemas algorítmicos pueden mejorar la capacidad de los Estados para cumplir las obligaciones que les incumben en virtud del derecho internacional de los derechos humanos, en particular optimizando la prestación de asistencia sanitaria, ampliando el acceso a la educación y reforzando los mecanismos de protección social, dichos sistemas también permiten desarrollar capacidades de vigilancia hasta el momento inéditas y pueden afianzar patrones discriminatorios. Además, la sustitución de procedimientos administrativos transparentes por procesos computacionales opacos puede mermar la eficacia del control administrativo y judicial y socavar el derecho a un recurso efectivo.

5. El presente estudio tiene en cuenta las deliberaciones mantenidas en el marco de las Naciones Unidas y las aportaciones de las partes interesadas, entre las que se cuentan las 32 respuestas al cuestionario del Comité Asesor recibidas de Estados, organizaciones regionales, instituciones nacionales de derechos humanos, la sociedad civil e instituciones académicas¹. El análisis desarrolla un marco integral basado en los derechos humanos y fundamentado en los principios del recurso efectivo y la rendición de cuentas, y subraya en todo momento la necesidad imperiosa de integrar la diligencia debida en materia de derechos humanos en la adquisición, el diseño y la implantación de la IA en la administración pública, garantizar la transparencia algorítmica para que las personas afectadas puedan comprender y cuestionar las decisiones adoptadas mediante automatización y establecer sólidas

¹ Pueden consultarse en <https://www.ohchr.org/en/hr-bodies/hrc/advisory-committee/impact-artificial-intelligence-systems-good-governance>.

salvaguardias institucionales que puedan detectar y combatir los sesgos discriminatorios, las violaciones de la privacidad y las irregularidades de procedimiento. Examina asimismo el impacto desproporcionado de la IA en las poblaciones vulnerables y analiza el riesgo de que la falta de medidas proactivas para garantizar un diseño inclusivo y una participación efectiva agrave las desigualdades existentes. En términos globales, el estudio reflexiona sobre la manera de aprovechar el potencial transformador de la IA en favor de la buena gobernanza sobre la base de los derechos humanos, la dignidad humana y la rendición de cuentas.

C. Buena gobernanza y derechos humanos

6. La Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (ACNUDH) define la buena gobernanza como el proceso por el cual las instituciones públicas dirigen los asuntos públicos, gestionan los recursos comunes y garantizan la realización de los derechos humanos mediante una actuación libre de abuso y corrupción². Esta concepción tomó forma cuando la Comisión de Derechos Humanos, en su resolución 2000/64, reconoció formalmente el vínculo entre la gobernanza, el desarrollo y la dignidad humana, y estableció que la promoción de los derechos humanos exige una gobernanza transparente, sujeta a rendición de cuentas y que responda a las necesidades de la población. Por consiguiente, como ha afirmado el Consejo de Derechos Humanos, la buena gobernanza debe ser participativa, estar sujeta a rendición de cuentas, ser transparente, receptiva y eficaz y regirse por el estado de derecho y los principios de derechos humanos³.

7. Desde 2008, el Consejo de Derechos Humanos ha desarrollado progresivamente el marco normativo que vincula la buena gobernanza y los derechos humanos, en el que se reconoce que un gobierno transparente, sujeto a rendición de cuentas y participativo constituye la base para la promoción de los derechos humanos, se hace hincapié en la lucha contra la corrupción y se insta a lograr una función pública profesional y basada en el mérito que respete el más alto grado de integridad⁴. Más recientemente, el Consejo ha analizado los retos que plantea la era digital, reconociendo que el acceso a Internet y a las tecnologías digitales es esencial para acceder a la información, los servicios públicos, la educación y la justicia, al tiempo que ha acogido con satisfacción el uso de innovaciones tecnológicas para prevenir, detectar y combatir la corrupción, siempre que se salvaguarden la protección de los datos personales y el derecho a la privacidad⁵.

8. En su resolución 57/5, el Consejo de Derechos Humanos dio un paso fundamental al incluir formalmente la IA en la agenda de la buena gobernanza y los derechos humanos. Reconoció que la IA presentaba tanto oportunidades como riesgos, puesto que podía brindar beneficios, como la mejora de la lucha contra la corrupción y de la prestación de servicios públicos, pero también planteaba retos, como la desinformación, que podían menoscabar los procesos democráticos. El Consejo instó a los Estados, a la sociedad civil, a los medios de comunicación y a otras partes interesadas a que elaboraran estrategias para mitigar los riesgos y promover al mismo tiempo el pleno disfrute de los derechos humanos y las libertades fundamentales. En particular, instó a los Estados a que promovieran sistemas seguros y fiables de IA de manera inclusiva y equitativa, y en beneficio de todos, reflejando el principio de que el desarrollo de la IA debe regirse por los principios de la buena gobernanza y los derechos humanos, entre los que figuran la transparencia, la equidad y la rendición de cuentas. Paralelamente, en 2023, el Secretario General convocó el Órgano Asesor de Alto Nivel sobre Inteligencia Artificial⁶. El Pacto Digital Global refleja la necesidad urgente de que la IA refuerce, y no menoscabe, los derechos humanos, la igualdad y el estado de derecho.

² Véase <https://www.ohchr.org/es/good-governance>.

³ Véanse, por ejemplo, las resoluciones del Consejo de Derechos Humanos 7/11, 19/20, 25/8, 31/14, 37/6, 45/9, 51/5 y 57/5.

⁴ Véanse las resoluciones del Consejo de Derechos Humanos 7/11, 19/20 y 25/8.

⁵ Véanse las resoluciones del Consejo de Derechos Humanos 45/9 y 51/5.

⁶ *Gobernanza de la IA en beneficio de la Humanidad: Informe final* (Nueva York, Naciones Unidas, 2024).

II. La buena gobernanza facilitada por la inteligencia artificial: oportunidades y riesgos para los derechos humanos

A. Contribuciones positivas de la inteligencia artificial a los derechos humanos y la buena gobernanza

1. Mejora de la prestación de servicios públicos y efectividad de los derechos económicos, sociales y culturales

9. Una de las ventajas que se suele reconocer a la IA en el ámbito de la gobernanza es su capacidad para transformar la prestación de servicios públicos y contribuir así a hacer efectivos todos los derechos, incluidos los económicos, sociales y culturales. Los sistemas de IA pueden procesar grandes volúmenes de datos y automatizar tareas complejas, lo que puede hacer que los servicios administrativos sean más eficientes, accesibles y adaptados a las necesidades individuales. Varios Estados y observadores internacionales han expresado cierto optimismo en que la IA, diseñada, implantada y regulada de forma adecuada, pueda apoyar de manera sustancial la capacidad de los Gobiernos en esferas como la atención de la salud, la educación y la protección social⁷.

10. El derecho a la salud se puede promover gracias a innovaciones basadas en la IA en el ámbito del diagnóstico médico y la gestión de la salud pública⁸, lo que pone de manifiesto la capacidad de esta tecnología para fomentar la disponibilidad y la calidad de los servicios de salud, que es una obligación fundamental en el marco del derecho al más alto nivel posible de salud⁹. La IA puede mejorar la precisión de los diagnósticos y optimizar el funcionamiento de los hospitales y el triaje de los pacientes, y con ello contribuir a que los sistemas de atención de la salud sean más eficaces y equitativos¹⁰.

11. En el ámbito de la educación, la IA puede reducir las brechas educativas y mejorar la gestión institucional¹¹. Las plataformas educativas basadas en la IA pueden reducir las desigualdades adaptando en tiempo real la enseñanza a los estilos y necesidades de aprendizaje individuales, lo que facilita una mejor inclusión de las personas que experimentan dificultades en los sistemas ordinarios, con independencia de su edad, su idioma o su discapacidad¹². Las aplicaciones que convierten los materiales didácticos a formatos alternativos constituyen herramientas de apoyo para las personas con discapacidad. La IA también puede asistir en la administración educativa detectando el riesgo de abandono de los estudios a fin de poner en marcha intervenciones personalizadas, al tiempo que permite a los docentes dedicar más atención a las necesidades individuales de los alumnos. Si se usan aplicando un enfoque centrado en las personas que dé prioridad al acceso equitativo, estas herramientas pueden reducir la disparidad en los resultados educativos y promover el principio, propio de la buena gobernanza, de la prestación de servicios inclusiva y centrada en las personas. En cambio, abrazar la IA de forma acrítica puede agravar la brecha digital y socavar la capacidad de acción humana, lo que, en última instancia, atenta contra el derecho a una educación inclusiva y equitativa¹³.

12. La IA puede mejorar la gestión de los programas de bienestar social y apoyar así el derecho a la seguridad social y a un nivel de vida adecuado¹⁴. Los algoritmos de aprendizaje automático pueden analizar con rapidez los datos de los hogares para determinar los más

⁷ Contribuciones del Brasil, Singapur, Uzbekistán y People's Solidarity for Participatory Democracy.

⁸ Contribuciones de Chile, Cuba y el Ecuador.

⁹ Contribución de la Comisión de Derechos Humanos de la Ciudad de México.

¹⁰ Contribuciones de Chile y el Ecuador.

¹¹ Véase [A/79/520](#).

¹² Contribuciones de Venezuela (República Bolivariana de) y la Comisión Nacional de Derechos Humanos de Myanmar.

¹³ Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), *Informe de seguimiento de la educación en el mundo, 2023: tecnología en la educación: ¿una herramienta en los términos de quién?* (París, 2023).

¹⁴ Contribución de España.

necesitados de asistencia o detectar errores en la distribución de las prestaciones. La automatización de la traducción y la interpretación puede facilitar el acceso de las personas a la justicia y los servicios públicos. Aplicada con transparencia, esta automatización contribuye a que unos recursos que son limitados lleguen rápidamente a los beneficiarios apropiados, lo que abre vías más eficaces para el ejercicio de los derechos socioeconómicos.

13. La tramitación automatizada de solicitudes puede mejorar la eficiencia y la capacidad de respuesta de las instituciones, elementos esenciales para una buena gobernanza. La integración de la IA en procesos administrativos como las declaraciones de impuestos o el cálculo de prestaciones puede traducirse en una mayor eficiencia y facilitar que los funcionarios se dediquen a tareas más complejas, que exigen un mayor grado de conocimientos y criterios¹⁵.

14. Más allá de los servicios concretos, la capacidad de análisis predictivo de la IA puede facilitar la transición de una gobernanza reactiva a otra proactiva. Por ejemplo, su integración en la gestión de desastres puede ayudar a predecir crisis como inundaciones o brotes epidémicos antes de que se produzcan, de modo que las autoridades puedan adoptar medidas preventivas¹⁶.

15. Pueden aprovecharse tecnologías basadas en la IA para facilitar determinadas funciones de aplicación de la ley, seguridad nacional y control fronterizo. Dichos sistemas pueden asistir en las investigaciones penales, por ejemplo mediante el reconocimiento de matrículas, y contribuir a las tareas de detección de actividades ilícitas. No obstante, cabe señalar que la implantación de la IA en estos ámbitos plantea algunas de las mayores preocupaciones en materia de derechos humanos, por lo que es necesario contar con marcos jurídicos sólidos que garanticen la transparencia, la rendición de cuentas y el pleno cumplimiento de las normas internacionales de derechos humanos¹⁷.

16. En la práctica, las aplicaciones mencionadas de la IA contribuyen a la buena gobernanza porque refuerzan la formulación de políticas basadas en datos empíricos. Dichas tecnologías mejoran la capacidad de las autoridades públicas para tomar rápidamente decisiones fundamentadas y asignar recursos en función de las necesidades previstas, lo que encarna el principio de capacidad de respuesta, fundamental para lograr una gobernanza centrada en las personas. Los modelos predictivos, que deben ser diseñados, implantados y gestionados teniendo debidamente en cuenta el derecho a la privacidad, el principio de no discriminación y otras consideraciones de derechos humanos, albergan un importante potencial para transformar la administración pública de modo que se promueva el ejercicio de los derechos económicos, sociales y culturales.

2. Fomento de la transparencia, la rendición de cuentas y la lucha contra la corrupción

17. La IA puede ser importante como herramienta para impulsar la transparencia en la administración pública y fomentar con ello la confianza de los ciudadanos y la rendición de cuentas por las instituciones gubernamentales. Al reducir el margen de discrecionalidad individual en la toma de decisiones rutinarias y generar registros digitales auditables, los sistemas de IA bien diseñados pueden reducir las oportunidades de corrupción y facilitar la detección de irregularidades¹⁸. Las innovaciones en materia de gobierno digital pueden contribuir a prevenir la corrupción y promover la rendición de cuentas, siempre que se respeten debidamente los datos personales y la privacidad¹⁹. En términos más generales, las tecnologías digitales y emergentes, incluida la IA, pueden facilitar la cooperación

¹⁵ Banco Mundial, *Artificial Intelligence in the Public Sector: Maximizing Opportunities, Managing Risks* (Washington D. C., 2020).

¹⁶ Contribuciones de Cuba y la Plataforma de Derechos Humanos de Ginebra.

¹⁷ Véase [A/HRC/60/45](#).

¹⁸ Business at OECD, "Harnessing AI for integrity: opportunities, challenges, and the business case against corruption" (Organización de Cooperación y Desarrollo Económicos (OCDE), París, 2024).

¹⁹ Véase la resolución 51/5 del Consejo de Derechos Humanos.

internacional para prevenir y combatir la corrupción, siempre que su desarrollo y aplicación se ajusten al derecho internacional de los derechos humanos aplicable²⁰.

18. La IA también puede reforzar la rendición de cuentas mejorando el mantenimiento de registros y el análisis de datos. Los sistemas de IA del sector público registran sistemáticamente cada acción y los datos en que se basan las decisiones, con lo que se crea un registro de auditoría exhaustivo. Esto puede ser un elemento disuasorio frente a conductas indebidas, puesto que las decisiones se pueden rastrear y revisar en caso de recurso o reclamación. Actualmente, en algunas jurisdicciones, se llevan a cabo iniciativas destinadas a diseñar sistemas de IA transparentes que puedan explicar el razonamiento de sus algoritmos, de modo que la gobernanza algorítmica, en lugar de socavar el estado de derecho, lo complementa²¹.

19. La IA también puede potenciar la participación ciudadana y la vigilancia independiente, lo que refuerza la rendición de cuentas. Las políticas de datos abiertos combinadas con análisis basados en la IA permiten a la sociedad civil, a los medios de comunicación y a otras partes interesadas analizar grandes volúmenes de información pública en busca de indicios de malas prácticas o vulneraciones de derechos. El Programa de las Naciones Unidas para el Desarrollo ha resaltado el valor de este uso de la IA, por ejemplo para vigilar las adquisiciones públicas y los flujos financieros ilícitos en favor de una gobernanza más amplia²².

20. La IA no es una solución a todos los retos a que hace frente la gobernanza, pero ofrece nuevas vías para promover los principios de transparencia y rendición de cuentas. Gracias a sus aplicaciones, las autoridades públicas pueden realizar avances importantes en el refuerzo de la integridad institucional, la lucha contra la corrupción y la defensa del estado de derecho.

3. Una herramienta para la inclusión, la participación y la igualdad

21. Otro aspecto positivo de la IA es su capacidad para promover la inclusión y la igualdad en la gobernanza. La buena gobernanza exige que las instituciones sean inclusivas y den voz y acceso a todos los sectores de la sociedad, incluidos los grupos marginados. La IA puede ayudar a superar las barreras a que se enfrentan estos grupos. Por ejemplo, las tecnologías de apoyo basadas en ella pueden permitir a las personas con discapacidad acceder a la información y los servicios públicos en igualdad de condiciones con los demás, en cumplimiento de las obligaciones establecidas en la Convención sobre los Derechos de las Personas con Discapacidad de garantizar la accesibilidad de la información y los servicios. Estas herramientas también pueden promover el derecho de acceso a la información de las comunidades que tienen dificultades para interactuar con las autoridades debido a barreras lingüísticas o de alfabetización.

22. La IA puede fomentar la participación en los procesos democráticos. Sus aplicaciones pueden ampliar el alcance de la democracia participativa creando nuevos canales para que las personas, incluidos los grupos históricamente infrarrepresentados en la formulación de políticas, contribuyan a la gobernanza. El Consejo de Derechos Humanos ha alentado las innovaciones de este tipo, señalando que las tecnologías digitales pueden facilitar la participación inclusiva en la vida pública y exhortando a los Estados a reducir las brechas digitales para que las poblaciones marginadas puedan participar en igualdad de condiciones con los demás²³. Además, las herramientas de traducción y procesamiento del lenguaje basadas en la IA pueden favorecer la participación política de las personas pertenecientes a las minorías lingüísticas en los Estados multilingües, por ejemplo facilitándoles el acceso a los procesos electorales. Sin embargo, lograr una participación efectiva a través de estas herramientas depende de que el acceso sea equitativo.

²⁰ Conferencia de los Estados Partes en la Convención de las Naciones Unidas contra la Corrupción, resolución 11/1.

²¹ Véase <https://e-estonia.com/estonia-and-automated-decision-making-challenges-for-public-administration>.

²² *New Technologies for Sustainable Development: Perspectives on Integrity, Trust and Anti-Corruption* (Nueva York, 2021).

²³ Resolución 47/23 del Consejo de Derechos Humanos.

23. La IA también puede ayudar a detectar y resolver desigualdades en la prestación de servicios. Mediante el análisis de los datos administrativos y otros conjuntos de datos pertinentes, las autoridades públicas pueden reconocer los casos en que la prestación es sistemáticamente deficiente para determinados grupos o zonas, como los barrios que adolecen de menor acceso a la atención de la salud o de mayores tiempos de respuesta ante las emergencias, lo que supone aplicar los principios de igualdad y no discriminación²⁴. Utilizadas con transparencia y con la participación de la comunidad, dichas aplicaciones favorecen el derecho al desarrollo y el compromiso común de no dejar a nadie atrás²⁵. Las comunidades afectadas no deberían limitarse a ser meras destinatarias pasivas de las políticas, sino participar activamente en la creación y la gobernanza de la IA. Cabe señalar, sin embargo, que para que se den estos beneficios, que no son en absoluto automáticos, es necesario que los propios sistemas de IA estén diseñados y entrenados de modo que no reproduzcan los sesgos existentes.

B. Repercusiones negativas de la inteligencia artificial en los derechos humanos y la buena gobernanza

24. Pese a su potencial, la integración de la IA en la gobernanza pública plantea dudas importantes en relación con la protección de los derechos humanos. Sin marcos jurídicos sólidos, salvaguardias en materia de derechos humanos y una supervisión humana, dichos sistemas pueden facilitar la vigilancia masiva, afianzar los sesgos sistémicos, mermar las debidas garantías y distorsionar el discurso público. Si no se toman medidas contra esos riesgos, pueden verse afectados los cimientos de la buena gobernanza y el estado de derecho. Calificar la IA de tecnología neutra sería una simplificación inadmisible, ya que su diseño y su aplicación pueden reflejar y agravar los desequilibrios de poder existentes.

1. Erosión de la privacidad a través de la vigilancia masiva y la explotación de los datos

25. La primera duda en materia de derechos humanos que plantea la IA en el ámbito de la gobernanza son sus efectos en la privacidad²⁶. En general, estos sistemas se basan en la recopilación y el tratamiento de grandes volúmenes de datos personales, como imágenes faciales, registros de ubicación, contenidos publicados en los medios sociales o identificadores biométricos, de modo que, utilizados por las autoridades públicas, pueden ampliar notablemente la capacidad de vigilancia. Las técnicas de aprendizaje automático son particularmente eficaces a la hora de encontrar patrones en grandes conjuntos de datos sin relación aparente entre sí, lo que da lugar a lo que se ha denominado “efecto mosaico”, es decir, la elaboración de perfiles completos de los individuos combinando información procedente de fuentes diversas²⁷. Un aspecto que caracteriza a estos sistemas es su capacidad para deducir información personal sensible de una persona sin que esta la haya facilitado explícitamente. Los órganos judiciales de diversas jurisdicciones han concluido que algunos sistemas de elaboración de perfiles en masa vulneran derechos fundamentales porque permiten una vigilancia encubierta y desproporcionada de las comunidades vulnerables sin la transparencia ni la justificación adecuadas, lo que pone de relieve la necesidad de establecer salvaguardias estrictas para que el perfilado automatizado no vulnere la privacidad ni perpetúe la discriminación.

26. Esta evolución suscita graves preocupaciones en relación con el derecho a la privacidad y su condición de derecho cuyo disfrute es imprescindible para ejercer las libertades de expresión, asociación y religión o creencias. La vigilancia asistida por la IA puede exponer la esfera de la vida privada a intrusiones arbitrarias, ya que el reconocimiento de patrones y la extracción de datos a gran escala pueden dar lugar a una vigilancia

²⁴ Véase <https://stories.undp.org/bending-the-ai-arc-towards-equity>.

²⁵ Órgano Asesor de Alto Nivel sobre Inteligencia Artificial (constituido en 2023), *Gobernanza de la IA en beneficio de la humanidad*.

²⁶ Véase A/HRC/60/45.

²⁷ Véase A/HRC/48/31; y David E. Pozen, “The mosaic theory, national security, and the Freedom of Information Act”, *Yale Law Journal*, vol. 115, núm. 3 (diciembre de 2005).

omnipresente que no respete los requisitos tradicionales de necesidad y proporcionalidad²⁸. Por ejemplo, combinar la videovigilancia interconectada en red con las tecnologías de reconocimiento facial permite el seguimiento de personas en espacios públicos en tiempo real sin la intervención permanente de un ser humano y resulta particularmente intrusiva, por lo que exige el máximo grado de justificación por motivos de privacidad y protección de datos.

27. Los riesgos que plantea la IA para la privacidad van más allá de las cámaras. Las autoridades recurren cada vez más al aprendizaje automático para analizar conjuntos de datos combinados, por ejemplo datos fiscales, datos relativos a la protección social e información de acceso público en Internet, con el fin de identificar a personas que puedan ser consideradas de riesgo. Varias investigaciones administrativas han concluido que la combinación automatizada a gran escala de datos personales puede dar lugar a decisiones erróneas con repercusiones importantes, lo que demuestra que la automatización puede vulnerar la privacidad y las debidas garantías²⁹.

28. En términos más generales, la vigilancia facilitada por la IA supone un cambio cualitativo por el que se pasa de la interceptación selectiva de comunicaciones fundamentada en las sospechas sobre una persona concreta a un filtrado algorítmico a gran escala que podría someter a poblaciones enteras a un escrutinio permanente. El ACNUDH ha advertido de que la vigilancia indiscriminada atenta contra el ejercicio de las libertades de expresión y de asociación³⁰. El efecto disuasorio de la vigilancia generalizada, que hace que las personas renuncien a ejercer sus derechos legítimos porque perciben un riesgo de sufrir sanciones o consecuencias negativas, está ampliamente documentado, especialmente en lo que respecta a la libertad de expresión³¹. Cuando saben que se las puede estar vigilando, a las personas les cuesta más expresarse con libertad o reunirse. El análisis predictivo puede ahondar estas preocupaciones al centrar la vigilancia en determinados barrios o comunidades basándose en indicadores asociados a la raza o la situación socioeconómica, lo que refuerza las prácticas discriminatorias, a la vez que erosiona la privacidad.

29. Los grupos marginados y vulnerables corren especial riesgo. Cada vez es más frecuente que en las fronteras se recojan los datos biométricos de los refugiados y migrantes o se los someta a evaluaciones algorítmicas de los riesgos³², lo que puede dar lugar a discriminación y atentar contra la privacidad y las debidas garantías. Los defensores de los derechos humanos, los periodistas y los activistas políticos se ven igualmente expuestos cuando se utilizan herramientas de vigilancia a gran escala. En este contexto, la privacidad se convierte en una preocupación colectiva. Cuando la vigilancia es omnipresente, el espacio democrático para la disidencia se reduce para todos. El Consejo de Derechos Humanos ha pedido que se apliquen estrictos criterios de legalidad, necesidad y proporcionalidad a cualquier medida de vigilancia digital³³, pero muchas prácticas basadas en la IA se llevan a cabo en marcos jurídicos obsoletos, en los que no se previeron las capacidades actuales.

30. En términos globales, la IA ha hecho aún más perentoria la protección de la privacidad³⁴. Los mismos usos que la hacen valiosa para la gobernanza son los que impulsan una recopilación y un tratamiento cada vez más generalizados de los datos personales, lo que constituye una vulneración directa del derecho a la privacidad y obstaculiza el ejercicio de otros derechos fundamentales. A falta de garantías adecuadas, la automatización de la vigilancia puede convertirse en la automatización de la censura y el control social, algo esencialmente incompatible con la buena gobernanza democrática.

²⁸ Véase [A/HRC/41/35](#).

²⁹ Contribución del Institute for Digital Rights.

³⁰ Véase [A/HRC/27/37](#).

³¹ Anita R. Gohdes, *Repression in the Digital Age: Surveillance, Censorship, and the Dynamics of State Violence* (Oxford, Oxford University Press, 2024).

³² Véase [A/75/590](#).

³³ Resolución 48/4 del Consejo de Derechos Humanos.

³⁴ Véase [A/HRC/48/31](#). Véase también [A/HRC/61/68](#), párrs. 25 y 26.

2. Sesgos algorítmicos y discriminación: afianzamiento de las desigualdades

31. La siguiente gran preocupación es el sesgo algorítmico, por el cual los sistemas de IA reproducen o amplifican los sesgos sociales y sus repercusiones en el derecho a la igualdad y a la no discriminación. La buena gobernanza exige que la prestación de servicios públicos y la adopción de decisiones se lleven a cabo con imparcialidad y un trato equitativo para todas las personas, con independencia de la raza, el género, la religión, la discapacidad u otros motivos de discriminación prohibidos. Sin embargo, los sistemas de IA no son intrínsecamente neutros y suelen reflejar pautas de discriminación arraigadas en los datos con que se entrenan o derivadas de su diseño. Por consiguiente, su implantación en la gobernanza puede dar lugar a una discriminación sistémica que resulte más difícil de detectar y combatir con los recursos disponibles que la discriminación derivada de la toma de decisiones tradicional por seres humanos³⁵.

32. Se han observado indicios de sesgo en las aplicaciones de la IA en todos los ámbitos del sector público. En el caso de la justicia penal, las herramientas de evaluación de riesgos utilizadas para apoyar las decisiones relativas a la libertad provisional, la imposición de penas y la libertad condicional han suscitado inquietud por la disparidad de sus resultados. Se han observado problemas similares en los sistemas de selección para el empleo, en los que pueden reproducirse sesgos de género o relacionados con la discapacidad, así como en los algoritmos financieros o administrativos, cuyos indicadores pueden penalizar inadvertidamente las interrupciones en la trayectoria laboral o a determinadas categorías de contribuyentes. En el contexto de la gobernanza, los modelos entrenados con datos que reflejan desigualdades socioeconómicas históricas pueden tener tendencia a señalar con más frecuencia la necesidad de someter a escrutinio a las empresas propiedad de minorías o a las comunidades de bajos ingresos³⁶. En sus contribuciones para el presente estudio, las partes interesadas señalaron de forma reiterada el riesgo de que la IA agrave la discriminación y las brechas de accesibilidad y mencionaron numerosos casos en que las desigualdades preexistentes, relacionadas con la raza, la clase social, el sexo, el género, la discapacidad o la situación migratoria, no solo no se habían reducido, sino que se habían amplificado, lo que contraviene la obligación de los Estados de prevenir las prácticas discriminatorias y promover la igualdad sustantiva³⁷.

33. Una de las principales dificultades radica en la opacidad. Muchos sistemas de IA exponen sus razonamientos de forma poco o nada inteligible. Además, a una máquina no puede suponerse intención. Por consiguiente, en ocasiones es difícil para las personas afectadas siquiera distinguir que se ha producido un sesgo, y aún más exigir responsabilidades u obtener reparación. Los marcos tradicionales contra la discriminación, que suelen estar orientados a demostrar la existencia de una discriminación intencionada, resultan inadecuados para los complejos perjuicios que se derivan del análisis de datos. Por ello, muchos análisis jurídicos y de derechos humanos abogan por un enfoque basado en los resultados, que se fije en la disparidad de los efectos³⁸. La observación de resultados adversos estadísticamente significativos para un grupo protegido debería dar paso al escrutinio del sistema y a la obligación de modificarlo o prescindir de él con independencia de la intención, medidas que responden a lo que se entiende comúnmente por discriminación indirecta. En términos globales, el sesgo algorítmico puede atentar gravemente contra el principio de no discriminación, ya que, tras la apariencia de la objetividad tecnológica, puede reforzar las desigualdades estructurales.

3. El problema de la caja negra: erosión de la transparencia, las debidas garantías y el derecho a recursos

34. Otro grave problema que plantea el uso de la IA en el ámbito de la gobernanza es la opacidad de la toma de decisiones algorítmica, lo que se conoce como el problema de la caja negra. En muchos casos, los sistemas de aprendizaje automático y basados en redes

³⁵ Véase [A/HRC/48/31](#); y [A/HRC/54/70](#), párr. 24.

³⁶ Contribución de la Plataforma de Derechos Humanos de Ginebra.

³⁷ Contribuciones de Irlanda, México, Singapur, España, Association pour la Défense des Droits de l'Homme et des Revendications Démocratiques/Culturelles du Peuple Azerbaïdjanais, Iran – ARC, Red Progresista Coreana - Jinbonet y Transparency International.

³⁸ Véase [A/HRC/56/68](#).

neuronales no ofrecen una fundamentación inteligible de sus decisiones para las personas afectadas. En el ámbito de la gobernanza, esa opacidad puede entrar en conflicto con los principios de transparencia, rendición de cuentas y estado de derecho y mermar garantías procesales como el derecho a comprender e impugnar las decisiones que afectan a los derechos y obligaciones, el derecho a un juicio imparcial y el derecho a un recurso efectivo. El ejercicio de los derechos en la era digital exige que la aplicación de sistemas automatizados sea transparente y esté sujeta a rendición de cuentas, puesto que la falta de explicabilidad puede obstaculizar el acceso a la justicia³⁹. La transparencia es esencial para la rendición de cuentas, como lo es la explicabilidad para el respeto de las debidas garantías y la atribución clara de responsabilidades para el acceso a un recurso efectivo⁴⁰.

35. Los tribunales tanto nacionales como internacionales se ven enfrentados a estas cuestiones. Los marcos regionales de protección de datos reconocen el derecho a no ser objeto de decisiones exclusivamente automatizadas cuando tengan efectos importantes y, si es el caso, el derecho a recibir información sustancial sobre la lógica en que se basan⁴¹. Sigue siendo objeto de debate la definición de “explicación sustancial” en el caso de los modelos complejos. Mientras algunos abogan por las explicaciones simplificadas o las medidas de transparencia, otros advierten de que, en el caso de los sistemas de aprendizaje profundo, las explicaciones *a posteriori* pueden resultar excesivamente técnicas o limitarse a meras inferencias⁴². Más allá de los debates, despunta la tendencia a considerar que la explicabilidad es una garantía básica.

36. La falta de transparencia también genera un déficit de rendición de cuentas. Cuando un sistema de IA causa perjuicios, la responsabilidad puede diluirse entre sus desarrolladores, sus proveedores y las autoridades públicas que lo adquieren e implantan. Esta incertidumbre menoscaba el derecho a un recurso efectivo, previsto en el artículo 2, párrafo 3, del Pacto Internacional de Derechos Civiles y Políticos, puesto que las personas afectadas no pueden señalar fácilmente a una parte responsable para exigirle cuentas. Esto plantea problemas en el ámbito de la rendición de cuentas, como los relativos a la atribución de responsabilidad por el incumplimiento de las obligaciones de derechos humanos o la distribución de esa responsabilidad entre los actores implicados, como las autoridades públicas que implantan los sistemas de IA y las empresas que los diseñan y suministran. El problema se agrava cuando las funciones públicas se externalizan de hecho a sistemas propietarios protegidos por el secreto comercial o por derechos de propiedad intelectual, lo que impide tanto a los funcionarios públicos como al público en general el acceso a una explicación clara sobre la base de las decisiones adoptadas⁴³. En el contexto del sector público, dicha falta de transparencia puede menoscabar los principios establecidos de responsabilidad administrativa y la obligación de las autoridades públicas de motivar las decisiones que afecten a los derechos individuales. La Alta Comisionada de las Naciones Unidas para los Derechos Humanos ha señalado la necesidad de que las empresas actúen con transparencia en lo relativo al diseño y el funcionamiento de las tecnologías de vigilancia y otras tecnologías de alto riesgo que suministran a los Gobiernos⁴⁴. El Grupo de Trabajo sobre la cuestión de los derechos humanos y las empresas transnacionales y otras empresas y el ACNUDH han subrayado que es necesario contar con una IA explicable, realizar evaluaciones de las repercusiones de los algoritmos e integrar la diligencia debida en materia de derechos humanos a lo largo de todo el ciclo de vida de la IA⁴⁵.

37. Tan importante cuestión exige la adopción no solo de medidas de recurso *a posteriori*, sino también de medidas preventivas sistémicas, entre las que deben figurar la realización obligatoria de evaluaciones del impacto en los derechos humanos, auditorías periódicas y registros públicos de los sistemas de toma automatizada de decisiones, y el establecimiento

³⁹ Véase A/HRC/48/31.

⁴⁰ Resolución 54/18 del Consejo de Derechos Humanos; y Contribución de People’s Solidarity for Participatory Democracy.

⁴¹ Reglamento (UE) 2016/679 (Reglamento general de protección de datos), art. 22.

⁴² OCDE, *AI Principles*. Puede consultarse en <https://www.oecd.org/en/topics/sub-issues/ai-principles.html>.

⁴³ Contribución de Red Progresista Coreana – Jinbonet.

⁴⁴ Véase A/HRC/48/31.

⁴⁵ Véanse A/HRC/59/32 y A/HRC/59/53.

de autoridades de supervisión independientes. Dicho enfoque preventivo responde al consenso de que los mecanismos correctivos convencionales pueden no llegar a tiempo en los ultraveloces entornos de las tecnologías basadas en los datos⁴⁶. Al mismo tiempo, en los principios de la OCDE sobre la IA se establece que la transparencia, la explicabilidad y la supervisión humana son elementos fundamentales y se insta a la creación de mecanismos que permitan a las personas afectadas impugnar los resultados de la IA (principio 1.3).

4. Amenazas a la libertad de expresión, el acceso a la información y la participación democrática

38. Las repercusiones de la IA en la libertad de expresión y la integridad de la esfera pública merecen atención urgente, ya que la buena gobernanza exige la existencia de un público informado y un entorno plural para el intercambio de ideas y para que los procesos políticos sean equitativos y transparentes. Una cuestión esencial es la proliferación de la desinformación generada por IA, que pone en peligro el acceso a información fidedigna mediante la producción barata y a gran escala de textos, imágenes y vídeos realistas con información falsa o engañosa⁴⁷. Por ejemplo, los observadores en los períodos preelectorales de varios países han documentado la existencia de redes coordinadas y la difusión de vídeos ultrafalsificados con el objetivo de desacreditar a determinados candidatos o reducir la participación electoral⁴⁸. Ello se complica cuando la desinformación se difunde a través de aplicaciones de mensajería cifrada, lo que dificulta la difusión de información fidedigna para contrarrestarla⁴⁹. El uso malintencionado de estas tecnologías causa la pérdida de confianza en todo el entorno informativo: cuando las personas comprueban que se puede crear contenido falso pero muy convincente, pueden desconfiar de los hechos genuinos por considerar que también pueden ser falsos, lo que dificulta la elaboración de políticas basadas en datos y merma la confianza pública en las comunicaciones gubernamentales. Además, los actores políticos y otras partes interesadas pueden analizar datos personales para ofrecer contenidos personalizados convincentes y con ello influir en la intención de voto o reducir la participación electoral. Tales prácticas, especialmente cuando se basan en la elaboración intrusiva de perfiles, pueden vulnerar el derecho a la privacidad y obstaculizar la libre formación de opiniones, elementos fundamentales de la libertad de pensamiento y expresión y necesarios para la participación efectiva en los asuntos públicos⁵⁰. Además, los algoritmos de recomendación de las plataformas basadas en la IA, que amplifican de forma sistemática los sesgos para lograr la máxima atención del usuario, pueden influir progresivamente en las creencias de este sin que él se dé cuenta, lo que suscita inquietud con respecto a la autonomía mental. Estas dinámicas erosionan la gobernanza porque fragmentan la base factual común, exacerbando la polarización e impiden los debates políticos razonados.

39. Hay perjuicios que afectan no solo a la libertad de expresión, sino también a la igualdad y a la seguridad de la persona. El Consejo de Derechos Humanos ha expresado preocupación por la posibilidad de que la manipulación de la información, en particular a través de las tecnologías digitales, alimente la violencia y los conflictos, y ha instado a los Estados a que exhorten a los actores privados a hacer frente a la desinformación y defiendan los derechos humanos⁵¹. La IA también puede amplificar los contenidos de odio o discriminatorios generando numerosas variaciones que eluden los filtros básicos, lo que supone una amenaza para la dignidad, la seguridad y la igualdad de las comunidades afectadas.

40. Otra dimensión de la cuestión es la censura y la moderación de contenidos. Los Estados pueden utilizar la IA para supervisar y restringir contenidos en línea a gran escala. Si bien puede ser legítimo eliminar contenidos ilícitos, los filtros automáticos suelen hacerlo

⁴⁶ UNESCO, “Recomendación sobre la Ética de la Inteligencia Artificial” (París, 2022); y Matthew U. Scherer, “Regulating Artificial intelligence systems: risks, challenges, competencies, and strategies”, *Harvard Journal of Law & Technology*, vol. 29, núm. 2 (primavera de 2016).

⁴⁷ Contribución de la American University.

⁴⁸ Freedom House, “Freedom on the Net 2019: the crisis of social media” (Washington D. C., 2019).

⁴⁹ A/HRC/61/68, párr. 74.

⁵⁰ Véanse A/73/348 y A/HRC/47/25. Véase también Comité de Derechos Humanos, observación general núm. 34 (2011), párrs. 9 y 10.

⁵¹ Resolución 55/10 del Consejo de Derechos Humanos. Véase también A/HRC/61/68, párr. 70.

en exceso, a menudo sin tener en cuenta elementos como el contexto, la intención satírica o el discurso de interés público, por lo que pueden interferir sin justificación en la libertad de expresión legítima y generar un efecto disuasorio. Toda restricción que se imponga a la libertad de expresión, incluidas las medidas destinadas a combatir la desinformación, debe ajustarse estrictamente a los requisitos establecidos en el artículo 19, párrafo 3, del Pacto Internacional de Derechos Civiles y Políticos, y cumplir los principios de legalidad, necesidad y proporcionalidad. Las medidas que se adopten para combatir la desinformación no deben servir de pretexto para restringir indebidamente la legítima libertad de expresión⁵².

41. En su conjunto, todos estos fenómenos minan la confianza en el entorno informativo. Cuando las personas no pueden confiar en la veracidad de lo que leen o ven ni en la autenticidad de sus interacciones en línea, se deteriora la cohesión social y se socava un pilar fundamental de la buena gobernanza. Además, esta desconfianza tiende a contagiarse a todas las formas de información, lo que plantea importantes dificultades a los medios de comunicación, las organizaciones que se dedican a la investigación y los órganos gubernamentales que deben recurrir a la comunicación digital en el desempeño de su labor. Para hacer frente a esas dificultades es necesario adoptar enfoques multipartitos cooperativos⁵³. Los Estados, las plataformas privadas y la sociedad civil deben colaborar para establecer las medidas de protección adecuadas, como garantizar la transparencia en la publicidad política y el etiquetado de los contenidos generados por IA, incorporar elementos que atestigüen la participación humana, invertir en alfabetización mediática y digital para reforzar la resiliencia social⁵⁴ y adoptar medidas que promuevan la equidad y el pluralismo en los sistemas de recomendación. Toda restricción a la libertad de expresión que se imponga para alcanzar estos objetivos debe ajustarse a los principios de legalidad, necesidad y proporcionalidad del derecho internacional.

42. Las posibles repercusiones negativas de la IA en la libertad de expresión y la participación democrática son amplias y afectan a todo el sistema. Atentan contra la capacidad del público para participar en un debate informado y exigir cuentas a las autoridades, lo que mina el tejido de la gobernanza democrática. Pese a ello, dichas repercusiones no son inevitables y pueden combatirse mediante medidas sostenidas multipartitas que preserven la capacidad de acción de las personas, el pluralismo y la confianza, en consonancia con los derechos humanos y el estado de derecho. Entre otras cosas, es fundamental incorporar salvaguardias en forma de elementos que atestigüen la participación humana a fin de preservar la confianza en los espacios públicos digitales y lograr que la participación en los procesos democráticos refleje una aportación humana genuina⁵⁵. Estas medidas pueden contribuir a salvaguardar la autenticidad de la participación cívica y mitigar los riesgos relacionados con la desinformación, las conductas coordinadas no auténticas y la influencia indebida en los procesos decisorios.

5. Automatización y derecho al trabajo: desplazamiento y aumento de la desigualdad

43. La automatización basada en la IA puede afectar de forma importante al derecho al trabajo, consagrado en el artículo 23 de la Declaración Universal de Derechos Humanos y el artículo 6 del Pacto Internacional de Derechos Económicos, Sociales y Culturales. A lo largo de la historia, los avances tecnológicos han desplazado algunas formas de empleo y han creado otras, pero la velocidad y el alcance de la implantación actual de la IA son inéditos, puesto que no solo afectan al trabajo manual rutinario, sino también a los puestos administrativos, creativos y decisorios. El impacto del desplazamiento causado por la IA se distribuye de forma desigual. Los sectores que presentan un riesgo más elevado de automatización suelen emplear a un número desproporcionado de personas pertenecientes a grupos marginados. Las mujeres, a menudo sobrerrepresentadas en los puestos de oficina y de apoyo administrativo, se enfrentan a un riesgo notablemente mayor de perder el empleo

⁵² [A/HRC/61/68](#), párr. 21.

⁵³ Véase [A/HRC/47/52](#).

⁵⁴ Resolución 57/5 del Consejo de Derechos Humanos.

⁵⁵ Bryan Ford, "Identity and personhood in digital democracy: evaluating inclusion, equality, security, and privacy in pseudonym parties and other proofs of personhood" (2020).

debido a la automatización⁵⁶. Además, las consecuencias van más allá de las economías avanzadas. Los países en desarrollo, que tradicionalmente han recurrido a las industrias intensivas en mano de obra para su desarrollo económico, se enfrentan a riesgos particulares, ya que la automatización puede erosionar las ventajas en cuanto a costos laborales en que se basan esas estrategias. Los análisis realizados por las Naciones Unidas han advertido de que, en los países en desarrollo, una parte significativa del empleo podría ser susceptible de automatización, lo que hace temer el aumento de la pobreza, la inestabilidad económica y la desigualdad a nivel mundial⁵⁷.

44. Más allá de las consecuencias económicas, el desplazamiento de empleo masivo causado por la IA afecta a la dignidad humana y la cohesión social. Para la mayoría de las personas, el trabajo no es una mera fuente de ingresos, sino lo que fundamenta su identidad, su propósito en la vida y su participación en la sociedad. El despido repentino de personas con voluntad de trabajar podría acarrear profundos costos psicológicos y sociales, por ejemplo en forma de alienación y pérdida de autoestima. Dicha posibilidad exige respuestas proactivas en materia de gobernanza. Para garantizar que nadie se quede atrás, es necesario adoptar medidas concertadas que defiendan el derecho al trabajo, que abarca el acceso a oportunidades de trabajo decente, como elemento fundamental de la buena gobernanza en la era de la IA.

6. Consumo energético y derechos ambientales: amenazas para la sostenibilidad

45. La implantación de sistemas de IA por las autoridades públicas da lugar a una huella material notable, si bien a menudo pasada por alto, lo que repercute en el derecho a un medio ambiente limpio, saludable y sostenible⁵⁸. Pese a la percepción de que la IA es una tecnología totalmente intangible, su infraestructura ejerce un consumo intensivo de recursos. El crecimiento exponencial de la demanda de capacidad computacional para el entrenamiento de modelos a gran escala y la ejecución de análisis de datos en tiempo real ha incrementado considerablemente el consumo de energía. Las previsiones indican que las necesidades energéticas de los centros de datos se dispararán de aquí a 2030, lo que puede causar una mayor dependencia de los combustibles fósiles para satisfacer la demanda de energía de base⁵⁹. Dicha posibilidad supone una amenaza para los compromisos contraídos a nivel mundial en relación con el clima y contribuye a agravar la contaminación atmosférica local, lo que representa un riesgo para la salud pública y los ecosistemas.

46. Las infraestructuras que requiere la IA generan emisiones de carbono, pero, además, su gestión térmica ejerce una presión considerable sobre los recursos de agua dulce⁶⁰. El importante volumen de agua que consume la refrigeración de los centros de datos compite en muchos casos con las necesidades de las comunidades locales, sobre todo en regiones con escasez de agua. En este contexto, un obstáculo esencial para la buena gobernanza es la opacidad de los datos ambientales. En general, los grandes proveedores de tecnología evitan divulgar datos precisos sobre la captación y el consumo de agua, lo que dificulta la rendición de cuentas por las autoridades públicas y la sociedad civil y vulnera el derecho al acceso a la información, que es esencial para una gobernanza participativa eficaz en la gestión de los recursos.

47. Además, el ciclo de vida de la IA perpetúa los patrones de injusticia ambiental y desigualdad a escala mundial. El *hardware* necesario para el funcionamiento de la IA fomenta la extracción de minerales críticos, en muchos casos en países en desarrollo y en condiciones que pueden vulnerar los derechos laborales y las normas ambientales. Además,

⁵⁶ Paweł Gmyrek y otros, “Generative AI and jobs: a global analysis of potential effects on job quantity and quality”, documento de trabajo núm. 96 de la OIT (Ginebra, 2023).

⁵⁷ *World Social Report 2020: Inequality in a Rapidly Changing World* (publicación de las Naciones Unidas, 2020).

⁵⁸ Contribuciones de Chile, Marruecos, la Arabia Saudita, la Association pour la Défense des Droits de l’Homme et des Revendications Démocratiques/Culturelles du Peuple Azerbaidjanais, Iran – ARC, el Institute for Digital Rights, el Centre for Advanced Study in International Humanitarian Law y la Plataforma de Derechos Humanos de Ginebra.

⁵⁹ Agencia Internacional de la Energía, *Energy and AI: World Energy Outlook Special Report* (2025).

⁶⁰ Joyeeta Gupta, Hilmer Bosch y Luc van Vliet, “AI’s excessive water consumption threatens to drown out its environmental contributions”, *The Conversation*, 21 de marzo de 2024.

su rápida obsolescencia acelera la generación de desechos electrónicos, con los que cargan mayoritariamente los países en desarrollo. Esta dinámica refleja una forma de brecha digital por la que los beneficios económicos derivados de la IA se concentran en los países desarrollados, mientras que las externalidades ambientales y la degradación de los recursos se desplazan a los países en desarrollo, lo que ahonda las disparidades existentes a nivel mundial. Por consiguiente, los Estados que deseen aprovechar la IA para promover la buena gobernanza deben hacer frente a las consecuencias ambientales y distributivas mencionadas en el marco de sus obligaciones de derechos humanos.

7. Inteligencia artificial en el sistema judicial y privación de libertad: proteger la equidad y la dignidad humana

48. En el ámbito de la administración de justicia, las preocupaciones relativas al uso de la IA son especialmente acusadas. Esta tecnología se integra cada vez más en los sistemas judiciales para realizar tareas administrativas, como la gestión de expedientes, la traducción, la anonimización automática de sentencias o la presentación electrónica de documentos, como también en los motores de búsqueda jurídicos y como herramienta de apoyo para jueces y abogados a la hora de redactar resúmenes de casos o analizar jurisprudencia⁶¹. Los tribunales y las autoridades judiciales también han empezado a considerar la posibilidad de utilizar herramientas de evaluación de riesgos basadas en la IA para facilitar la toma de decisiones relativas a las fianzas, la libertad condicional y la imposición de penas⁶². Dichas aplicaciones pueden mejorar la eficiencia frente al enorme volumen de casos, pero su uso obliga a plantearse hasta qué punto se pueden confiar funciones judiciales esenciales a sistemas algorítmicos. La independencia judicial puede verse afectada cuando son los poderes ejecutivo o legislativo, y no el propio poder judicial, quienes se encargan de la selección de los datos para entrenar al sistema, del diseño de los algoritmos o de la supervisión de los sistemas de IA utilizados en los tribunales, y todavía más si dichas funciones se subcontratan a empresas privadas⁶³. La esencia de un juicio imparcial reside en la deliberación humana y la rendición de cuentas. Si bien la IA puede prestar asistencia en los procesos judiciales, los jueces desempeñan una función irremplazable: ejercer su facultad discrecional y tener en cuenta las particularidades de cada caso. Por ello, autoridades del poder judicial han subrayado que, en el ámbito de la justicia, las personas, y no las máquinas, deben seguir siendo quienes tomen las decisiones en último término⁶⁴. En todos estos casos son fundamentales la estricta supervisión humana, la transparencia y la posibilidad de que las partes entiendan y cuestionen la función de la IA. La opacidad de los sistemas de IA puede afectar negativamente al derecho a la libertad y a la seguridad de la persona y al derecho a un juicio imparcial, consagrados en los artículos 9 y 14 del Pacto Internacional de Derechos Civiles y Políticos, cuando las personas acusadas no pueden impugnar las decisiones que los afectan. Conscientes de estos riesgos, algunas jurisdicciones han empezado a calificar el uso de sistemas de IA por las autoridades judiciales como de alto riesgo y a prohibir las evaluaciones de riesgos asistidas por la IA que se basen exclusivamente en la elaboración de perfiles o en rasgos de la personalidad para predecir el riesgo de comisión de delitos⁶⁵. Del mismo modo, la integración de la IA en los procedimientos judiciales debe realizarse con la debida cautela para preservar la independencia judicial, la equidad procesal y la confianza pública en la administración de justicia.

49. En los entornos penitenciarios, la introducción de sistemas basados en la IA ha suscitado nuevas preocupaciones en materia de derechos humanos. Algunos centros de nueva construcción, que suelen denominarse “prisiones inteligentes”, recurren al reconocimiento facial, la vigilancia biométrica y el análisis predictivo para mejorar la seguridad y la

⁶¹ Véase [A/79/296](#).

⁶² Consejo de Europa, *European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment* (Estrasburgo, 2019).

⁶³ Véase [A/79/296](#); y UNESCO, *Kit de herramientas global sobre IA y el estado de derecho para el poder judicial* (París, 2024).

⁶⁴ UNESCO, “Recomendación sobre la Ética de la Inteligencia Artificial”. Véase también UNESCO, *Directrices para el uso de sistemas de IA en cortes y tribunales* (París, 2025).

⁶⁵ Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial.

eficiencia operacional. Dichas herramientas pueden contribuir a reducir los niveles de violencia y contrabando, pero también restar humanidad al entorno penitenciario al minimizar el contacto interpersonal y el ejercicio de la discrecionalidad profesional. La vigilancia algorítmica permanente de las personas privadas de libertad puede intensificar la sensación de aislamiento. Además, recurrir en exceso a la automatización para gestionar la privación de libertad puede malbaratar los esfuerzos de rehabilitación, ya que elimina el factor humano, esencial para prestar una atención adecuada y realizar evaluaciones individualizadas. En las Reglas Mínimas de las Naciones Unidas para el Tratamiento de los Reclusos (Reglas Nelson Mandela) se hace hincapié en la importancia del contacto humano apreciable, puesto que, sin él, el aislamiento prolongado constituye una forma de trato cruel, inhumano o degradante (reglas 43 y 44). Por consiguiente, si no se somete a control alguno, la automatización en los lugares de reclusión puede atentar contra la obligación, establecida en el derecho internacional de los derechos humanos, de tratar a todas las personas privadas de libertad con humanidad y con el respeto debido a la dignidad inherente al ser humano.

III. Construir una inteligencia artificial respetuosa de los derechos humanos en pro de la buena gobernanza

A. Sesgo en los datos y desequilibrio mundial en el desarrollo de la inteligencia artificial: problemas para los derechos humanos

50. Un problema fundamental es el sesgo en los datos y los algoritmos, agravado por el hecho de que el desarrollo de la IA se concentra en un grupo reducido de actores y contextos. Los sistemas de IA suelen reflejar los valores, supuestos y limitaciones inherentes a su diseño y a los datos con los que se entrenan. La industria y la investigación en este ámbito siguen dominados por un pequeño número de Estados tecnológicamente avanzados y un puñado de grandes empresas. Por ello, estos sistemas encarnan un conjunto limitado de puntos de vista y es difícil imaginar que puedan responder y representar a la diversidad de contextos culturales, sociales y regionales existentes.

51. Este fenómeno se ha descrito como una forma de colonialismo digital, en la que el dominio de unos pocos grandes proveedores de tecnología y sus marcos normativos perpetuarían los patrones históricos de exclusión⁶⁶. Los países donde la IA es un sector emergente suelen adoptar sistemas y marcos regulatorios elaborados fuera de sus fronteras y diseñados en su mayor parte pensando en las economías tecnológicamente avanzadas. En ocasiones, estas soluciones se implantan con muy escasas aportaciones de los responsables locales de la formulación de políticas o de las comunidades afectadas. Los desequilibrios en la capacidad de influencia durante los procesos de implantación pueden agravar las desigualdades existentes. Un enfoque uniforme de la IA en el que la tecnología no se adapte a los contextos culturales y sociales particulares puede, sin pretenderlo, empujar a los sistemas de conocimientos locales, en particular los de los Pueblos Indígenas, a la irrelevancia. En esos casos, la aplicación de marcos homogéneos importados de fuentes externas puede marginar los valores y conocimientos locales⁶⁷. Las repercusiones de los sesgos inherentes a los datos y algoritmos en los derechos humanos son directas y considerables. Los algoritmos sesgados o discriminatorios pueden generar resultados desiguales e injustos que supongan una vulneración efectiva de los principios de igualdad y no discriminación. Si no se tienen en cuenta estos aspectos, se afianzará lo que se ha denominado “injusticia automatizada”, en la que decisiones algorítmicas opacas niegan a las personas un trato equitativo o el acceso a servicios u oportunidades sobre la base de datos erróneos o sesgados. Las personas pertenecientes a grupos marginados y vulnerables tienen más probabilidades de verse afectadas, lo que ahondará aún más las desigualdades existentes.

⁶⁶ Jake Okechukwu Effoduh y Miracle Okumu Mudeyi, “TWAILing AI governance: will third world countries be left behind?”, TWAIL Review (7 de agosto de 2025).

⁶⁷ Contribución de People’s Solidarity for Participatory Democracy.

52. Hacer frente a estos problemas exige integrar los principios de derechos humanos, la equidad y la diversidad en todas las fases del desarrollo de la IA⁶⁸. Los ajustes técnicos no bastan por sí solos, ya que es necesario introducir cambios más profundos en lo que respecta a la diversidad de las personas que desarrollan los sistemas de IA y los objetivos para los que estos se diseñan. Garantizar una participación diversa y un diseño inclusivo en todo el ciclo de vida de la IA, desde la recopilación de datos hasta el diseño e implantación de los algoritmos, es esencial para mitigar los sesgos y promover la rendición de cuentas. Al incorporar una mayor variedad de perspectivas, los sistemas de IA y los marcos de gobernanza en los que funcionan pueden reflejar mejor las diversas necesidades y valores de la sociedad.

53. Están surgiendo iniciativas alentadoras a nivel comunitario para democratizar la IA. Científicos de datos de varias regiones, por ejemplo de África y Asia, están colaborando para crear conjuntos de datos y herramientas adaptados a las lenguas con pocos recursos y las necesidades locales. El objetivo es que la IA sea más inclusiva y culturalmente adecuada y huya de los enfoques uniformes que no tienen en cuenta la diversidad de los contextos. Las iniciativas internacionales de elaboración de normas están empezando a reflejar estas preocupaciones⁶⁹. En 2021, la UNESCO instó a los Estados a que velaran por que el desarrollo de la IA respetara los derechos humanos, promoviera la diversidad y protegiera la equidad⁷⁰. Aplicar estos principios exige cuestionar los supuestos de diseño predeterminados y recabar la participación de expertos locales en el proceso de desarrollo. En lugar de imponer inadvertidamente normas externas, los sistemas de IA deberían tener en cuenta una pluralidad de valores jurídicos y culturales. Además, el propio marco internacional de derechos humanos debe aplicarse a la IA tomando en consideración las diferencias culturales. Algunos expertos advierten de que aplicar el derecho internacional de los derechos humanos sin la debida consideración a los contextos locales puede responder a una perspectiva implícitamente individualista que no refleja bien las perspectivas colectivas o comunitarias de los derechos que predominan en determinados entornos culturales. Por consiguiente, las normas internacionales relativas a la protección de datos o la transparencia de los algoritmos deberían interpretarse y aplicarse de modo que se respeten los derechos humanos fundamentales teniendo en cuenta al mismo tiempo las diferencias culturales y jurídicas pertinentes en los mecanismos de rendición de cuentas a nivel local.

B. Promoción de la participación efectiva y la inclusión

54. Una de las principales dificultades a la hora de aprovechar la IA para la buena gobernanza es la falta de participación efectiva de las distintas partes interesadas en los procesos decisivos clave. Por regla general, quienes elaboran las estrategias de alto nivel en materia de IA y las directrices éticas son expertos de un reducido grupo de Estados en los que se concentra el desarrollo de estas tecnologías, sin aportaciones significativas de otras regiones. Los más afectados por los servicios públicos o los sistemas de vigilancia basados en la IA, como la población en general, las personas pertenecientes a comunidades marginadas y los Estados más pequeños o menos desarrollados tecnológicamente, suelen quedar excluidos de los procesos de formulación de políticas, lo que refleja el amplio déficit democrático de que adolece la gobernanza de las tecnologías de la información. Los principios relativos a la IA que reciban el respaldo internacional serán difíciles de llevar a la práctica si no tienen en cuenta las prioridades de la mayor parte de la población mundial.

55. El grado de preparación y la capacidad de gobernanza de los países en relación con la IA es muy dispar. Algunos Gobiernos, especialmente en los países desarrollados, han establecido marcos nacionales y regionales sobre la IA y participan activamente en los foros internacionales, mientras que muchos otros apenas empiezan a elaborar políticas o directrices al respecto. La cooperación regional puede ser una vía para dar mayor visibilidad a las voces menos representadas. En algunas regiones, los Estados han comenzado a colaborar a través de las organizaciones regionales con el fin de fijar posiciones comunes sobre la ética de la

⁶⁸ Véase [A/HRC/59/32](#).

⁶⁹ Véase [A/HRC/53/42](#).

⁷⁰ “Recomendación sobre la Ética de la Inteligencia Artificial”.

IA y su aplicación a la gobernanza que reflejen los valores locales compartidos, además de los principios universales de derechos humanos, lo que ilustra que los países en desarrollo pueden trabajar con espíritu colectivo para ejercer su influencia en la formulación de las normas mundiales relativas a la IA. Sin embargo, es necesario que dichas iniciativas reciban un amplio apoyo de la comunidad internacional.

56. La participación efectiva también es esencial a nivel nacional. Cuando los Gobiernos integran la IA en la administración pública, ya sea para automatizar la asignación de prestaciones sociales, desplegar sistemas de vigilancia en las llamadas ciudades inteligentes o utilizar herramientas asistidas por IA en el ámbito judicial, las personas y las comunidades afectadas, además de gozar del derecho a la información y a la transparencia, deberían tener una participación efectiva en el diseño y aplicación de esos sistemas. La participación de las personas en las decisiones que afectan a sus derechos constituye un principio general del derecho que debe aplicarse también a la adopción en la gobernanza de tecnologías basadas en la IA. Pese a ello, en muchos países la participación de la sociedad civil en las políticas tecnológicas sigue siendo escasa o superficial.

57. Para subsanar esta deficiencia en materia de gobernanza, los Estados deberían poner en marcha mecanismos formales de consulta y supervisión siempre que se vaya a adoptar un sistema de IA con un importante impacto social. Deberían establecerse procedimientos de consulta pública antes de toda implantación de un sistema basado en la IA que pueda afectar a los derechos humanos o al interés público. Debería garantizarse que la comunidad y la sociedad civil estén representadas en los órganos de supervisión y crear consejos de evaluación ética que examinen las iniciativas gubernamentales en este ámbito. Dichos órganos y consejos, si cuentan con representantes de las comunidades afectadas y de organizaciones de derechos humanos y con expertos independientes, pueden evaluar con mayor eficacia las repercusiones sociales de la implantación de la IA y asesorar sobre las medidas de protección necesarias. Estos grupos aportan conocimientos muy valiosos, que los enfoques tecnocráticos suelen pasar por alto, sobre las repercusiones en los derechos humanos y sobre los contextos culturales. Para que la IA contribuya a los objetivos de la buena gobernanza, su aplicación debe gobernarse con la participación efectiva de aquellos a quienes debe servir, lo que reforzará la confianza de los ciudadanos y garantizará que su uso redunde en el interés público.

C. Definición de las responsabilidades de las distintas partes interesadas para una implantación de la inteligencia artificial basada en los derechos

58. Lograr un ecosistema de IA que respete los derechos humanos y prevenga los abusos exige una acción concertada de todas las partes interesadas que se base en los principios de equidad, soberanía⁷¹ y justicia social. Los Estados tienen la obligación primordial de asegurarse de que el desarrollo y la implantación de la IA se ajusten a las normas internacionales de derechos humanos. Dicha obligación comienza por el establecimiento de marcos jurídicos y regulatorios sólidos, que exijan transparencia, rendición de cuentas y supervisión en la toma de decisiones algorítmica en el sector público⁷². Los Gobiernos también deberían invertir en educación pública y alfabetización digital para que las personas entiendan el uso que se hace de la IA en la gestión pública y conozcan sus derechos en este ámbito. Los países en desarrollo necesitan un espacio de políticas y un apoyo suficientes para regular la IA en función de sus prioridades sociales particulares. Dado que muchos países en desarrollo carecen de los recursos necesarios para elaborar de forma autónoma marcos de gobernanza eficaces para los sistemas algorítmicos, la cooperación internacional debería centrarse en el fomento de capacidades sustantivas mediante transferencia de tecnología,

⁷¹ En este contexto, la soberanía se refiere a la capacidad de los Estados, en particular de los países en desarrollo, para ejercer un control efectivo y autónomo sobre las decisiones que atañen a las políticas relativas a la IA, la gobernanza de los datos y la infraestructura digital, sin someterse a factores o agentes externos. Sobre los problemas derivados de la concentración del desarrollo de la IA y los riesgos de la dependencia digital, véanse los párrs. 50 a 54.

⁷² Véase [A/HRC/48/31](#).

formación en materia de regulación e intercambio de conocimientos Sur-Sur, con el fin de capacitar a todos los países para que aprovechen la IA en favor de la buena gobernanza. Por consiguiente, las normas internacionales deberían fijar los principios mostrando al mismo tiempo flexibilidad para que los Estados puedan aplicarlas en función de su contexto institucional y cultural⁷³. Los Gobiernos deben fundamentar la gobernanza de la IA en los valores universales de los derechos humanos, al tiempo que orientan activamente el desarrollo tecnológico para que responda a las necesidades nacionales y locales.

59. Los Estados deben cumplir sus obligaciones de derechos humanos también cuando delegan la prestación de servicios públicos a actores privados, por ejemplo en los ámbitos de la educación, la sanidad y el suministro de agua. Cuando se utilizan sistemas de IA en la prestación de servicios públicos, se aplican las obligaciones de derechos humanos de conformidad con los principios establecidos, incluida la atribución de un comportamiento al Estado cuando proceda y, en cualquier circunstancia, el deber de actuar con la debida diligencia en relación con los actores privados. Los Estados deben asegurarse de que recurrir a esos sistemas no atente contra la disponibilidad, accesibilidad, aceptabilidad y calidad de los servicios ni dé lugar a discriminación ni a ningún otro efecto negativo para los derechos humanos⁷⁴. Dicha obligación forma parte de las obligaciones positivas de los Estados en materia de derechos humanos, entre las que figuran establecer y aplicar normas regulatorias que se apliquen por igual a la prestación de servicios privados apoyados por la IA, llevar a cabo un seguimiento y una evaluación continuos, velar por que se investiguen con eficacia las violaciones de los derechos humanos y ofrecer recursos efectivos a las personas cuyos derechos se vean afectados. Las obligaciones del Estado en este contexto tienen plena aplicación con independencia de los medios tecnológicos utilizados para prestar los servicios⁷⁵.

60. Gran parte de la tecnología basada en la IA que los Gobiernos adquieren o implantan ha sido desarrollada por actores del sector privado. De conformidad con los Principios Rectores sobre las Empresas y los Derechos Humanos, todas las empresas, incluidas las que diseñan y suministran sistemas de IA, tienen la obligación de respetar los derechos humanos. En la práctica, las empresas deberían ejercer la diligencia debida en materia de derechos humanos en todos sus procesos de desarrollo relacionados con la IA⁷⁶, lo que implica consultar con las partes interesadas, evaluar los posibles riesgos, como los sesgos algorítmicos, la discriminación y otras formas de uso indebido, y adaptar las decisiones de diseño para mitigar los daños. Las empresas también deben combatir los sesgos sistémicos y el predominio de los puntos de vista de un número reducido de regiones en sus herramientas de IA, lo que pueden lograr diversificando su fuerza de trabajo y los conjuntos de datos que utilizan para entrenar los modelos a fin de reflejar mejor la diversidad existente en el mundo. Dichas medidas, si van acompañadas de la inclusión efectiva de investigadores locales y del debido respeto por los contextos locales, pueden contribuir a que los productos basados en la IA se adapten mejor a una mayor variedad de entornos sociales y culturales.

61. La sociedad civil, en la que se incluyen las organizaciones no gubernamentales, las instituciones académicas, los medios de comunicación, los grupos comunitarios y los jóvenes, desempeña un papel esencial para lograr que la IA pueda utilizarse en la gobernanza respetando los derechos humanos. Por consiguiente, debería capacitársela para participar en los procesos de toma de decisiones y facilitarle vías oficiales para ello. Los conocimientos de la sociedad civil y el mundo académico pueden contribuir al establecimiento de las garantías necesarias, como exigir auditorías independientes de los algoritmos o crear mecanismos para que la comunidad exprese sus opiniones sobre los sistemas de IA utilizados en el sector público. Las organizaciones de la sociedad civil también desempeñan una función fundamental de concienciación y alfabetización en relación con la IA: al dar a conocer sus

⁷³ UNESCO, “Recomendación sobre la Ética de la Inteligencia Artificial”.

⁷⁴ Véase [A/HRC/59/53](#).

⁷⁵ Comité de Derechos Económicos, Sociales y Culturales, observación general núm. 24 (2017), párr. 11; y Comisión Africana de Derechos Humanos y de los Pueblos, observación general núm. 7, párr. 43.

⁷⁶ Véase [A/HRC/59/32](#).

repercusiones en los derechos humanos, contribuyen a que las comunidades puedan exigir la rendición de cuentas.

62. Las organizaciones internacionales, como las Naciones Unidas, tienen la responsabilidad de establecer marcos inclusivos que orienten en materia de uso de la IA para una buena gobernanza, en consonancia con los derechos humanos. Las medidas que se adopten a nivel mundial deben integrar de forma explícita el punto de vista de los países en desarrollo. Voces procedentes de los países en desarrollo y las comunidades marginadas advierten de que las iniciativas internacionales deben evitar reproducir las jerarquías de poder existentes bajo el pretexto de la gobernanza global. Por consiguiente, las normas internacionales, en forma de directrices no vinculantes o de posibles instrumentos jurídicamente vinculantes, deberían elaborarse mediante procesos inclusivos y transparentes, en los que la opinión de todos los Estados cuente por igual. Tal vez sea necesario plantearse nuevos mecanismos de gobernanza global para que la IA refuerce su conformidad con los derechos humanos.

IV. Recomendaciones

A. Estados

63. Los Estados deberían adoptar marcos legislativos y regulatorios integrales sobre el uso de la IA en la administración pública que obliguen a realizar evaluaciones del impacto en los derechos humanos antes de su implantación y periódicamente tras esta, exijan la transparencia y explicabilidad de los algoritmos, establezcan órganos de supervisión independientes que vigilen los sesgos, la precisión y el cumplimiento del derecho internacional de los derechos humanos en los sistemas de IA y garanticen la investigación eficaz y el acceso a recursos cuando la implantación de la IA dé lugar a violaciones de los derechos humanos.

64. Los Estados deberían garantizar una participación pública efectiva en las decisiones relativas a la introducción de sistemas de IA en el sector público mediante el establecimiento de procedimientos de consulta transparentes que cuenten con la participación activa de la sociedad civil, los expertos técnicos y las comunidades afectadas, a fin de que las tecnologías de gobernanza reflejen las diversas necesidades y valores de la sociedad.

65. Los Estados deberían regular estrictamente el uso de la vigilancia basada en la IA para prevenir vulneraciones de los derechos a la privacidad, a la libertad de expresión y a la libertad de asociación, y velar por que su aplicación se ajuste a los principios de legalidad, necesidad y proporcionalidad y esté sujeta a una supervisión judicial efectiva e independiente.

66. Los Estados deberían salvaguardar los procesos democráticos y la integridad de los ecosistemas informativos, y para ello promover la alfabetización mediática y digital, exigir transparencia en la publicidad política y el etiquetado de los contenidos generados por IA, estudiar o facilitar mecanismos adecuados para atestiguar la participación humana y apoyar el pluralismo, velando al mismo tiempo por que las medidas adoptadas para combatir la desinformación mantengan un equilibrio adecuado con la libertad de expresión y no den lugar a una censura indebida.

67. Los Estados deberían velar por que la integración de la IA en los procedimientos judiciales y los entornos penitenciarios no elimine la supervisión humana, las debidas garantías ni la dignidad inherente al ser humano, prohibir que las decisiones que afectan a la libertad se basen exclusivamente en sistemas automatizados y preservar un contacto humano apreciable en los centros de reclusión.

68. Los Estados deberían hacer frente con proactividad a las repercusiones socioeconómicas de la automatización basada en la IA mediante la adaptación de los sistemas educativos, la puesta a punto de programas de reciclaje profesional y el refuerzo de los mecanismos de protección social a fin de defender el derecho al trabajo y lograr transiciones justas.

69. Los Estados deberían evaluar y mitigar los efectos ambientales de la implantación de la IA, como la considerable cantidad de energía y agua que consumen los centros de datos o la generación de desechos electrónicos, velando por que la adopción de una gobernanza facilitada por esas tecnologías no atente contra el derecho a un medio ambiente limpio, saludable y sostenible ni imponga cargas ambientales a los países en desarrollo.

B. Comunidad internacional

70. La comunidad internacional debería trabajar en favor de una gobernanza global inclusiva y equitativa de la IA, y para ello garantizar que los países en desarrollo y los representantes del Sur Global tengan una participación efectiva en los procesos internacionales de elaboración de normas y sus opiniones cuenten igual que las de los demás, y velar por que no se impongan marcos regulatorios uniformes que no reflejen los contextos locales o perpetúen patrones de dependencia digital.

71. La comunidad internacional debería facilitar fomento de la capacidad, transferencia de tecnología y apoyo financiero para que todos los Estados, en particular los países de ingresos bajos, puedan elaborar y aplicar marcos de gobernanza de la IA eficaces y adaptados a cada contexto, y debería imponer que la prestación de asistencia internacional para el desarrollo a proyectos digitales estuviera condicionada al establecimiento de salvaguardias sólidas en materia de derechos humanos.

72. La comunidad internacional debería considerar la posibilidad de crear mecanismos específicos en el marco de las Naciones Unidas, como una relatoría especial o un grupo de trabajo sobre las tecnologías digitales nuevas y emergentes y sus repercusiones en los derechos humanos, con el fin de mantenerse al día de las novedades y prestar asesoramiento sobre posibles medidas de protección. También debería considerar la posibilidad de elaborar un marco específico bajo los auspicios de las Naciones Unidas que imponga a los Estados obligaciones concretas en relación con las repercusiones de la IA en los derechos humanos.

C. Empresas

73. Las empresas que diseñan, desarrollan o implantan sistemas de IA deberían respetar los derechos humanos, entre otras cosas ejerciendo la diligencia debida en esta esfera a lo largo de todo el ciclo de vida de la IA, evaluando y mitigando los riesgos de que se produzcan sesgos algorítmicos, discriminación y vulneraciones de la privacidad, consultando con las comunidades afectadas y estableciendo mecanismos de reclamación eficaces, o participando en ellos, para que las personas puedan denunciar los perjuicios sufridos y obtener reparación.

74. Las empresas deberían promover una IA inclusiva y equitativa, y para ello diversificar su fuerza de trabajo y los conjuntos de datos que utilizan para entrenar los modelos a fin de que reflejen la diversidad existente en el mundo, velar por que las herramientas de IA se adapten a los distintos contextos sociales, culturales y regionales, y evitar las prácticas que afiancen los sesgos históricos o agraven las desigualdades digitales.

D. Todas las partes interesadas

75. Todas las partes interesadas, incluidos los Estados, las organizaciones internacionales, las empresas, la sociedad civil, el mundo académico y las comunidades afectadas, deberían colaborar a través de procesos multipartitos transparentes e inclusivos a fin de formular medidas en materia de gobernanza que defiendan los derechos humanos. Dicha colaboración debería ir acompañada de un seguimiento independiente de los efectos sociales de la IA, de vías de recurso accesibles y de instrumentos comunes, como marcos de evaluación del impacto de los algoritmos o normas de fiscalización, a fin de que la gobernanza de la IA permanezca centrada en

los derechos humanos, responda a las necesidades de todos y respete la dignidad humana.

76. Todas las partes interesadas deberían apoyar activamente a las organizaciones de la sociedad civil, especialmente en los países en desarrollo, mediante iniciativas de financiación y fomento de la capacidad, con el objetivo de mejorar la comprensión del público, realizar un seguimiento de los efectos de la IA y lograr que las personas más afectadas por las tecnologías emergentes puedan influir en el modo en que se implantan.
