Naciones Unidas A/HRC/48/31

Asamblea General

Distr. general
13 de septiembre de 2021
Español
Original: inglés

Consejo de Derechos Humanos
48º período de sesiones
13 de septiembre a 1 de octubre de 2021
Temas 2 y 3 de la agenda
Informe anual del Alto Comisionado de las Naciones Unidas para los Derechos Humanos e informes de la Oficina del Alto Comisionado y del Secretario General

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

El derecho a la privacidad en la era digital*

Informe de la Alta Comisionada de las Naciones Unidas para los Derechos Humanos

Resumen

En el presente informe, encargado por el Consejo de Derechos Humanos en su resolución 42/15, la Alta Comisionada analiza cómo el uso generalizado de la inteligencia artificial por parte de los Estados y las empresas, en particular en la elaboración de perfiles, la adopción automatizada de decisiones y las tecnologías de aprendizaje automático, afecta al disfrute del derecho a la privacidad y los derechos conexos. Tras presentar un panorama general del marco jurídico internacional, la Alta Comisionada destaca los aspectos de la inteligencia artificial que facilitan la injerencia en la vida privada y ofrece ejemplos de sus efectos en el derecho a la privacidad y los derechos conexos en cuatro sectores clave. A continuación, la Alta Comisionada examina varios enfoques para abordar los desafíos que se plantean y formula un conjunto de recomendaciones para los Estados y las empresas en relación con el diseño y la aplicación de salvaguardias con el objetivo de prevenir y reducir al mínimo los resultados negativos y facilitar el pleno disfrute de los beneficios que puede proporcionar la inteligencia artificial.

* Este informe se presentó con retraso para incluir en él la información más reciente.

I. Introducción

1. Este informe se presenta en cumplimiento de la resolución 42/15 del Consejo de Derechos Humanos, en la que el Consejo solicitó a la Alta Comisionada de las Naciones Unidas para los Derechos Humanos que organizase un seminario de expertos para examinar la forma en que la inteligencia artificial, incluidas la elaboración de perfiles, la adopción automatizada de decisiones y las tecnologías de aprendizaje automático, puede, sin las debidas salvaguardias, repercutir en el goce del derecho a la privacidad, y que preparase un informe temático sobre el tema y lo presentase al Consejo de Derechos Humanos en su 45º período de sesiones[1].

2. En los últimos años no se ha producido ningún avance tecnológico que haya cautivado más la imaginación del público que la inteligencia artificial (IA), en particular las tecnologías de aprendizaje automático[2]. De hecho, estas tecnologías pueden ser una herramienta extraordinariamente poderosa si se ponen al servicio del bien, al ayudar a las sociedades a superar algunos de los grandes problemas de nuestro tiempo. No obstante, también pueden tener efectos negativos, incluso catastróficos, si se implantan sin tener suficientemente en cuenta sus consecuencias para los derechos humanos.

3. Aunque el presente informe no se centra en la pandemia de la enfermedad por coronavirus (COVID-19), la actual crisis sanitaria mundial ilustra a la perfección y de manera muy evidente la velocidad, la escala y los efectos de la IA en diversas esferas de la vida en distintos lugares del mundo. Para seguir la propagación de la enfermedad, se han utilizado sistemas de rastreo de contactos que se sirven de diversos tipos de datos (geolocalización, tarjetas de crédito, sistemas de transporte, sanitarios y demográficos) e información procedente de las redes personales. Los sistemas de IA se han utilizado para señalar a personas potencialmente infectadas o infecciosas y obligarlas a aislarse o ponerse en cuarentena. También se han utilizado para atribuir calificaciones predictivas que han arrojado resultados discriminatorios para los estudiantes de las escuelas públicas y de los barrios más desfavorecidos. Estos hechos demuestran la amplia variedad de efectos que tienen estos sistemas en la vida cotidiana de las personas. En todos estos casos, el ejercicio del derecho a la privacidad se ve afectado, ya que la IA utiliza información personal y a menudo toma decisiones que tienen efectos tangibles en la vida de las personas. Sin embargo, la cuestión de la privacidad está íntimamente relacionada con diversos aspectos que afectan al disfrute de otros derechos, como el derecho a la salud, la educación, la libertad de circulación, la libertad de reunión pacífica, la libertad de asociación y la libertad de expresión.

4. En 2019, en el documento “La más alta aspiración: un llamamiento a la acción en favor de los derechos humanos”, el Secretario General de las Naciones Unidas reconoció que la era digital había abierto nuevas perspectivas en materia de bienestar humano, conocimiento e investigación. Subrayó que las tecnologías digitales proporcionaban nuevos medios para abogar por los derechos humanos, defenderlos y ejercerlos. Sin embargo, estas se utilizan con demasiada frecuencia para vulnerar estos derechos, especialmente los de las personas que ya son vulnerables o a las que se tiende a excluir, por ejemplo mediante medidas de vigilancia, represión, censura y acoso en línea, incluso contra defensores de los derechos humanos. La digitalización de los sistemas de bienestar, a pesar de su potencial para mejorar la eficiencia, comporta riesgos de exclusión para las personas más necesitadas. El Secretario General hizo hincapié en que los avances de las nuevas tecnologías no deben utilizarse para socavar los derechos humanos, agravar la desigualdad o exacerbar la discriminación existente. Subrayó que la gobernanza de la IA debe garantizar la equidad, la rendición de cuentas, la explicabilidad y la transparencia. En el ámbito de la seguridad, el Secretario General reiteró su llamamiento a la prohibición mundial de los sistemas de armas autónomos letales.

[1] La elaboración del informe fue aplazada. Véanse A/HRC/45/26 y A/HRC/47/61.

[2] No existe una definición generalmente aceptada del término “inteligencia artificial”. En el presente informe, se emplea para hacer referencia a una constelación de procesos y tecnologías que permiten que las computadoras complementen o reemplacen tareas específicas que de otro modo serían ejecutadas por seres humanos, como tomar decisiones y resolver problemas (A/73/348, párr. 3), lo que comprende, entre otras cosas, el aprendizaje automático y el aprendizaje profundo.

5. El presente informe se basa en los dos informes anteriores de la Oficina del Alto Comisionado sobre la cuestión del derecho a la privacidad en la era digital[3]. También tiene en cuenta la información obtenida en el seminario virtual de expertos que se organizó en cumplimiento de la resolución 42/15 del Consejo, celebrado los días 27 y 28 de mayo de 2020, así como las respuestas a la petición de la Alta Comisionada de que se presentaran contribuciones al presente informe[4].

II. Marco jurídico

6. El artículo 12 de la Declaración Universal de Derechos Humanos, el artículo 17 del Pacto Internacional de Derechos Civiles y Políticos y otros instrumentos internacionales y regionales de derechos humanos reconocen el derecho a la privacidad como un derecho humano fundamental[5]. Este derecho desempeña un papel fundamental en el equilibrio de poder entre el Estado y el individuo y es un derecho fundamental para una sociedad democrática[6]. Su importancia para el goce y el ejercicio de los derechos humanos dentro y fuera de Internet[7] en un mundo cada vez más centrado en los datos no deja de aumentar.

7. El derecho a la privacidad es una expresión de la dignidad humana y está vinculado a la protección de la autonomía y la identidad personales[8]. Entre los aspectos de la privacidad que revisten especial importancia en el contexto del uso de la IA está la privacidad de la información, a saber, la información que existe o puede obtenerse sobre una persona y su vida, así como las decisiones basadas en esa información[9], y la libertad de adoptar decisiones en lo que respecta a la propia identidad.

8. Ninguna injerencia en el derecho a la privacidad debe ser arbitraria o ilegal[10]. El término “ilegal” significa que los Estados solo pueden interferir en el derecho a la privacidad en aplicación de la ley y de conformidad con ella. La propia legislación debe ser conforme con las disposiciones, fines y objetivos del Pacto Internacional de Derechos Civiles y Políticos y especificar con detalle las circunstancias precisas en las que se permite dicha injerencia[11]. Con la introducción del concepto de arbitrariedad se pretende garantizar que incluso cualquier injerencia prevista en la ley esté en consonancia con las disposiciones, los propósitos y los objetivos del Pacto y sea, en todo caso, razonable en las circunstancias particulares del caso[12]. En consecuencia, toda injerencia en el derecho a la intimidad debe servir a un fin legítimo y ser necesaria y proporcionada para alcanzar ese fin legítimo[13]. Toda

[3] A/HRC/27/37 y A/HRC/39/29.

[4] Véase www.ohchr.org/EN/Issues/DigitalAge/Pages/cfi-digital-age.aspx para consultar la petición de contribuciones y las contribuciones recibidas.

[5] Véase el artículo 16 de la Convención sobre los Derechos del Niño, el artículo 14 de la Convención Internacional sobre la Protección de los Derechos de Todos los Trabajadores Migratorios y de Sus Familiares, el artículo 22 de la Convención sobre los Derechos de las Personas con Discapacidad, el artículo 10 de la Carta Africana sobre los Derechos y el Bienestar del Niño, el artículo 11 de la Convención Americana sobre Derechos Humanos y el artículo 8 del Convenio para la Protección de los Derechos Humanos y de las Libertades Fundamentales (Convenio Europeo de Derechos Humanos).

[6] A/HRC/39/29, párr. 11.

[7] Comité de los Derechos del Niño, observación general núm. 25 (2021), párrs. 67 y 68; y A/HRC/39/29, párr. 11.

[8] Comité de los Derechos del Niño, observación general núm. 25 (2021), párr. 67; y Tribunal Europeo de Derechos Humanos, *Goodwin v. the United Kingdom*, demanda núm. 28957/95, sentencia de 11 de julio de 2002, párr. 90

[9] A/HRC/39/29, párr. 5.

[10] Para obtener un análisis detallado de los términos “arbitrario” e “ilegal”, véase A/HRC/27/37, párrs. 21 a 27.

[11] Comité de Derechos Humanos, observación general núm. 16 (1988), párrs. 3 a 8.

[12] *Ibid*., párr. 4.

[13] *Toonen c. Australia* (CCPR/C/50/D/488/1992), párr. 8.3, *Van Hulst c. los Países Bajos* (CCPR/C/82/D/903/1999), párrs. 7.3 y 7.6, *Madhewoo c. Mauricio* (CCPR/C/131/D/3163/2018), párr. 7.5, y CCPR/C/USA/CO/4, párr. 22. Véase también Comité de los Derechos del Niño, observación general núm. 25 (2021), párr. 69.

restricción debe ser además la opción menos intrusiva disponible y no comprometer la esencia del derecho a la privacidad[14].

9. El derecho a la privacidad se aplica a todas las personas. Toda diferencia en la protección de este derecho basada en la raza, el color, el sexo, el idioma, la religión, la opinión política o de cualquier otra índole, el origen nacional o social, la posición económica, el nacimiento o cualquier otra condición es incompatible con el principio de no discriminación establecido en el artículo 2, párrafo 1, y el artículo 3 del Pacto Internacional de Derechos Civiles y Políticos. La discriminación basada en estos motivos vulnera el derecho a la igualdad ante la ley previsto en el artículo 26 del Pacto.

10. El artículo 2, párrafo 1, del Pacto Internacional de Derechos Civiles y Políticos dispone que los Estados se comprometen a "respetar y garantizar" a todos los individuos que se encuentren en su territorio y estén sujetos a su jurisdicción los derechos reconocidos en el Pacto, sin discriminación alguna. En otras palabras, los Estados no solo deben abstenerse de violar los derechos reconocidos en el Pacto[15], sino que también tienen la obligación de adoptar medidas positivas para proteger el goce de esos derechos. Esto implica el deber de adoptar medidas legislativas y de otra índole para proteger a las personas de las injerencias en su privacidad, provengan de las autoridades estatales o de personas físicas o jurídicas[16]. Este deber también se refleja en el primer pilar de los Principios Rectores sobre las Empresas y los Derechos Humanos, que señala el deber de los Estados de brindar protección frente a las consecuencias negativas para los derechos humanos que resultan de las actividades de las empresas.

11. Las empresas tienen la responsabilidad de respetar todos los derechos humanos internacionalmente reconocidos. Esto implica que deben abstenerse de infringir los derechos humanos de terceros y hacer frente a las consecuencias negativas sobre los derechos humanos en las que tengan alguna participación. El segundo pilar de los Principios Rectores sobre las Empresas y los Derechos Humanos proporciona un plan oficial aplicable a todas las empresas para dar cumplimiento a esa responsabilidad[17]. La responsabilidad de respetar los derechos humanos comprende todas las actividades y relaciones comerciales de una empresa.

III. Repercusiones de la inteligencia artificial en el derecho a la privacidad y otros derechos humanos

A. Características pertinentes de los sistemas de inteligencia artificial

12. El funcionamiento de los sistemas de IA puede facilitar y agravar de diversas maneras las intrusiones en la privacidad y las injerencias en otros derechos. Esto se debe, entre otras cosas, a aplicaciones totalmente nuevas o características de los sistemas de IA que amplían, intensifican o incentivan la interferencia con el derecho a la privacidad, sobre todo mediante el aumento de la recopilación y el uso de datos personales.

13. Los sistemas de IA suelen basarse en grandes conjuntos de datos, que a menudo incluyen datos personales. Esto incentiva la recopilación, el almacenamiento y el tratamiento de datos a gran escala. Muchas empresas optimizan sus servicios para recopilar la mayor cantidad de datos posible[18]. Por ejemplo, las empresas en línea, como las plataformas de medios sociales, dependen de la recopilación y monetización de cantidades masivas de datos

[14] Comité de Derechos Humanos, observación general núm. 31 (2004), párr. 6; A/HRC/27/37, párr. 22, y A/HRC/39/29, párr. 10.

[15] Comité de Derechos Humanos, observación general núm. 31 (2004), párr. 6.

[16] A/HRC/39/29, párr. 23. Véase también Comité de Derechos Humanos, observaciones generales núm. 16 (1988), párrs. 1 y 9, y núm. 31 (2004), párr. 8; y Comité de los Derechos del Niño, observación general núm. 25 (2021), párrs. 36 a 39.

[17] En su resolución 17/4, el Consejo de Derechos Humanos aprobó por unanimidad los Principios Rectores sobre las Empresas y los Derechos Humanos.

[18] Wolflie Christl, *Corporate surveillance in everyday life* (Viena, Cracked Lab – Institute for Critical Digital Culture, 2017).

sobre los usuarios de Internet[19]. La llamada Internet de las cosas es una fuente de datos de rápido crecimiento explotada tanto por las empresas como por los Estados. La recopilación de datos se realiza en espacios íntimos, privados y públicos[20]. Los corredores de datos adquieren, fusionan, analizan y comparten datos personales con innumerables destinatarios. Estas transacciones de datos escapan en gran medida al escrutinio público y están escasamente limitadas por los marcos jurídicos vigentes[21]. Los conjuntos de datos resultantes son vastos y la información recogida alcanza proporciones sin precedentes.

14. Además de exponer la vida privada de las personas a las empresas y los Estados, estos conjuntos de datos hacen a las personas vulnerables en diversos aspectos. Las filtraciones de datos han expuesto en reiteradas ocasiones información delicada de millones de personas[22]. Los grandes conjuntos de datos permiten innumerables formas de análisis e intercambio de datos con terceros, lo que a menudo conlleva otras intromisiones en la privacidad con consecuencias negativas para los derechos humanos. Los acuerdos que permiten a los organismos gubernamentales tener acceso directo a esos conjuntos de datos en poder de las empresas, por ejemplo, aumentan la probabilidad de que se produzcan injerencias arbitrarias o ilegales en el derecho a la privacidad de las personas afectadas[23]. Resulta especialmente preocupante que la fusión de datos procedentes de diversas fuentes pueda facilitar la desanonimización[24]. Al mismo tiempo, el diseño de los conjuntos de datos puede tener repercusiones para la identidad de las personas. Por ejemplo, un conjunto de datos que registra el género de forma binaria clasificaría de forma errónea a las personas que no se identifican ni como hombres ni como mujeres. El almacenamiento a largo plazo de datos personales también conlleva riesgos específicos, ya que estos pueden quedar expuestos a futuras formas de explotación que no se preveían en el momento de su recogida[25]. Con el tiempo, los datos pueden volverse inexactos, no pertinentes o arrastrar errores de identificación históricos, dando lugar a resultados sesgados o erróneos en el futuro tratamiento de los datos[26].

15. Cabe señalar que los sistemas de IA no se basan exclusivamente en el tratamiento de datos de carácter personal. No obstante, aunque los datos no sean personales, su uso puede tener consecuencias negativas para los derechos humanos[27], incluido el derecho a la privacidad, como se muestra a continuación.

16. Las herramientas de IA se utilizan profusamente para comprender los patrones de comportamiento humano. Con el acceso a los conjuntos de datos adecuados, es posible sacar conclusiones sobre el número de personas de un barrio concreto que suelen frecuentar un determinado lugar de culto, qué programas de televisión prefieren e incluso a qué hora suelen levantarse y acostarse. Estas herramientas permiten hacer deducciones de gran alcance sobre las personas, incluso sobre su estado mental y físico, y permitir la identificación de grupos de personas[28], por ejemplo en función de sus inclinaciones políticas o personales particulares. La IA también se utiliza para evaluar la probabilidad de comportamientos o acontecimientos futuros. Las deducciones y predicciones realizadas mediante IA, a pesar de su carácter

[19] Contribución de Ranking Digital Rights.

[20] Contribuciones del Centro para la Gobernanza de la Comunicación de la Universidad Nacional de Derecho de Delhi, Derechos Digitales, Digital Rights Watch, Global Partners Digital, International Center for Not-for-Profit Law y Universidade Federal de Uberlândia.

[21] Aaron Rieke y otros, *Data brokers in an open society* (Londres, Open Society Foundation, 2016).

[22] Véase, por ejemplo, www.ftc.gov/news-events/press-releases/2019/07/equifax-pay-575-million-part-settlement-ftc-cfpb-states-related.

[23] Contribución de Global Network Initiative.

[24] Contribuciones del Centro para la Gobernanza de la Comunicación de la Universidad Nacional de Derecho de Delhi, Derechos Digitales y Privacy International.

[25] Contribución de OVD-Info.

[26] Comité para la Eliminación de la Discriminación Racial, recomendación general núm. 36 (2020), párr. 33.

[27] Consejo de Europa, “Guidelines on addressing the human rights impacts of algorithmic systems”, (apéndice de la Recomendación CM/Rec(2020)1 del Comité de Ministros a los Estados miembros sobre el impacto de los sistemas algorítmicos en los derechos humanos), secc. A, párr. 6.

[28] Contribuciones de Derechos Digitales y Privacy International.

probabilístico, pueden servir de base para la adopción de decisiones que afectan a los derechos de las personas, a veces de forma totalmente automatizada.

17. Un gran número de deducciones y predicciones afectan profundamente al goce del derecho a la privacidad, en particular a la autonomía de las personas y su derecho a establecer los detalles de su identidad. También plantean muchas cuestiones en relación con otros derechos, como el derecho a la libertad de pensamiento y de opinión, el derecho a la libertad de expresión y el derecho a un juicio imparcial y otros derechos conexos.

18. Las decisiones basadas en la IA no están exentas de errores. De hecho, la escalabilidad de las soluciones en esta esfera puede aumentar drásticamente los efectos negativos de tasas de error aparentemente bajas[29]. Los resultados erróneos de los sistemas de IA proceden de diversas fuentes. Para empezar, los resultados de los algoritmos de IA tienen elementos probabilísticos, lo que significa que sus resultados albergan un margen de incertidumbre[30]. Además, a menudo la pertinencia y la exactitud de los datos resultan cuestionables. Por otra parte, las expectativas poco realistas pueden llevar a que se implanten herramientas de IA que no están preparadas para producir los objetivos deseados. Por ejemplo, en la esfera médica, un análisis de cientos de herramientas para el diagnóstico y la prevención de la COVID-19, cuyo desarrollo había suscitado grandes expectativas, reveló que ninguna de ellas resultaba apta para el uso clínico[31].

19. Los resultados de los sistemas de IA que se basan en datos erróneos pueden contribuir de muchas formas a la vulneración de derechos humanos por ejemplo, señalando erróneamente a una persona como posible terrorista o indicando que ha cometido un fraude en el cobro de prestaciones sociales. Resultan especialmente preocupantes los conjuntos de datos sesgados que conducen a decisiones discriminatorias basadas en sistemas de IA[32].

20. Los procesos de toma de decisiones de muchos sistemas de IA son opacos. La complejidad del entorno de datos, los algoritmos y los modelos que subyacen al desarrollo y funcionamiento de estos sistemas, así como el secretismo deliberado de los agentes gubernamentales y privados, son factores que obstaculizan la comprensión del público de sus consecuencias para los derechos humanos y la sociedad. Los sistemas de aprendizaje automático añaden un importante elemento de opacidad, ya que pueden llegar a identificar patrones y desarrollar prescripciones que son difíciles o imposibles de explicar[33]. Esto es lo que se suele denominar el problema de la "caja negra"[34]. La opacidad dificulta el examen significativo de los sistemas de IA y puede ser un obstáculo para la rendición de cuentas efectiva en los casos en que estos sistemas lleguen a causar daños[35]. No obstante, cabe señalar que estos sistemas no tienen por qué ser totalmente inescrutables[36].

[29] Contribución de Alemania.

[30] Agencia de los Derechos Fundamentales de la Unión Europea, "#BigData: Discrimination in data-supported decision-making" (Viena, 2018), pág. 4.

[31] Véase www.technologyreview.com/2021/07/30/1030329/machine-learning-ai-failed-covid-hospital-diagnosis-pandemic/.

[32] Comité para la Eliminación de la Discriminación Racial, recomendación general núm. 36 (2020), párrs. 31 a 36; y Panel de Alto Nivel sobre la Cooperación Digital, "The age of digital interdependence" (junio de 2019), págs. 17 y 18.

[33] Contribución de Alemania.

[34] Véase www.scientificamerican.com/article/demystifying-the-black-box-that-is-ai/.

[35] Véase www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6; y www.cambridge.org/core/journals/international-review-of-the-red-cross/article/abs/ai-for-humanitarian-action-human-rights-and-ethics/C91D044210CADF7A0E023862CF4EE758.

[36] Véase, por ejemplo, Inioluwa Deborah Raji y otros, "Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing", 3 de enero de 2020.

B. Preocupación por los sistemas de inteligencia artificial en sectores clave

21. En la presente sección se aportan ejemplos de cómo se materializan estas cuestiones en la práctica, examinando cuatro esferas clave en las que la aplicación de herramientas de IA ha suscitado preocupación.

La inteligencia artificial en la aplicación de la ley, la seguridad nacional, la justicia penal y la gestión de fronteras

22. Los Estados cada vez incorporan más sistemas de IA en sus mecanismos de aplicación de la ley, seguridad nacional, justicia penal y gestión de fronteras[37]. Aunque no hay duda de que muchas de esas aplicaciones pueden suscitar preocupación, en la presente sección solamente se abordarán algunos ejemplos representativos de diversas cuestiones que se están planteando en relación con los derechos humanos.

23. Los sistemas de IA se utilizan a menudo como herramientas de predicción. Estas aplican algoritmos para analizar grandes cantidades de datos, incluidos datos históricos, a fin de evaluar los riesgos y predecir las tendencias futuras. En función de cuál sea la finalidad, los datos de formación y los datos analizados pueden incluir, por ejemplo, antecedentes penales, actas de detención, estadísticas sobre delincuencia, informes de intervenciones policiales en barrios específicos, publicaciones en medios sociales, datos de comunicaciones y registros de viajes[38]. Las tecnologías pueden utilizarse para crear perfiles de personas, identificar lugares susceptibles de albergar una mayor actividad delictiva o terrorista, e incluso señalar a individuos como sospechosos probables y futuros reincidentes[39].

24. Las consecuencias de estas actividades para la privacidad y los derechos humanos en general son enormes. En primer lugar, los conjuntos de datos utilizados incluyen información sobre un gran número de personas, lo que afecta a su derecho a la privacidad. En segundo lugar, estos pueden desencadenar intervenciones del Estado, como registros, interrogatorios, detenciones y enjuiciamientos, aunque las evaluaciones de IA por sí mismas no deberían considerarse motivos razonables de sospecha dado el carácter probabilístico de las predicciones. Entre los derechos afectados se encuentran el derecho a la privacidad, el derecho a un juicio imparcial, el derecho a no ser objeto de detención y privación de libertad arbitrarias y el derecho a la vida. En tercer lugar, la opacidad inherente a las decisiones basadas en la IA plantea cuestiones especialmente apremiantes en lo que respecta la responsabilidad del Estado cuando toma como base la IA para la adopción de medidas coercitivas, más aún en las esferas en las que suele existir una falta general de transparencia, como las actividades de las fuerzas de lucha contra el terrorismo[40]. En cuarto lugar, las herramientas de predicción conllevan un riesgo inherente de perpetuar o incluso potenciar la discriminación, al reflejar prejuicios raciales y étnicos históricos integrados en los conjuntos de datos que se utilizan, como la tendencia a la aplicación desproporcionada de medidas policiales a determinadas minorías[41].

25. Los avances en el campo de la tecnología de reconocimiento biométrico han llevado a que cada vez sea más utilizada por las fuerzas del orden y los organismos de seguridad nacional. El reconocimiento biométrico se basa en la comparación de la representación digital de determinados rasgos de una persona, como el rostro, la huella dactilar, el iris, la voz o la

[37] Para un análisis exhaustivo de las consecuencias para los derechos humanos de la utilización de la IA y otras tecnologías digitales en la gestión de las fronteras, véase A/75/590.

[38] Contribucion de Privacy International. Véase también A/HRC/44/57, párr. 35.

[39] Véase www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf.

[40] A/74/335 y A/HRC/43/46, párrs. 37 y 38.

[41] Contribución de Tech Hive Advisory Limited. Véase también la recomendación general núm. 36 (2020) del Comité para la Eliminación de la Discriminación Racial, párr. 33 y el documento de sesión de la Alta Comisionada de las Naciones Unidas para los Derechos Humanos sobre la promoción y protección de los derechos humanos y las libertades fundamentales de los africanos y los afrodescendientes frente al uso excesivo de la fuerza y otras violaciones de los derechos humanos por agentes del orden (A/HRC/47/CRP.1), disponible en: www.ohchr.org/Documents/Issues/Racism/A_HRC_47_CRP_1.pdf, párrs. 15 y 19.

forma de andar, con otras características de este tipo recogidas en una base de datos[42]. A partir de la comparación se deduce una probabilidad mayor o menor de que la persona en cuestión sea aquella cuya identidad debe ser determinada. Estos procesos se realizan cada vez más en tiempo real y a distancia. En particular, las autoridades de todo el mundo hacen un uso cada vez mayor del reconocimiento facial remoto en tiempo real[43].

26. El reconocimiento biométrico a distancia en tiempo real plantea graves preocupaciones en lo que respecta al derecho internacional de los derechos humanos, tal y como la Alta Comisionada ha destacado con anterioridad[44]. Algunas de estas preocupaciones reflejan los problemas asociados a las herramientas de predicción, en particular la posibilidad de una identificación errónea de las personas y consecuencias desproporcionadas para los miembros de ciertos grupos[45]. Además, la tecnología de reconocimiento facial puede utilizarse para elaborar perfiles de personas en función de su etnia, raza, origen nacional, sexo y otras características[46].

27. El reconocimiento biométrico a distancia conlleva un grave riesgo de injerencia en el derecho a la privacidad. La información biométrica de una persona constituye uno de los atributos fundamentales de su personalidad, ya que revela características únicas que la distinguen de otras personas[47]. Además, el reconocimiento biométrico a distancia aumenta considerablemente la capacidad de las autoridades del Estado de identificar y rastrear sistemáticamente a las personas en los espacios públicos, lo que socava la capacidad de estas de hacer su vida sin ser observadas y tiene un efecto negativo directo en el ejercicio de los derechos a la libertad de expresión, de reunión pacífica y de asociación, así como a la libertad de circulación[48]. En este contexto, la Alta Comisionada acoge con satisfacción los recientes esfuerzos por limitar o prohibir el uso de tecnologías de reconocimiento biométrico en tiempo real[49].

28. También se han desarrollado herramientas de IA que supuestamente deducen el estado emocional y mental de las personas a partir de sus expresiones faciales y otros "datos biométricos predictivos" para determinar si representan una amenaza para la seguridad[50]. Los sistemas de reconocimiento facial de emociones parten de la premisa de que es posible inferir automática y sistemáticamente el estado emocional de los seres humanos a partir de sus expresiones faciales, lo que carece de una base científica sólida[51]. Los investigadores han visto que la relación entre las emociones y las expresiones faciales[52] son poco significativas y han destacado que las expresiones faciales varían según las culturas y los contextos[53], lo que hace que el reconocimiento de las emociones sea susceptible de sesgos y malas interpretaciones. Teniendo en cuenta estas preocupaciones, el uso de sistemas de reconocimiento de emociones por las autoridades públicas, por ejemplo para señalar a personas para la realización de controles policiales o detenciones policiales o para evaluar la

[42] A/HRC/31/64, párr. 14.
[43] Contribuciones de Derechos Digitales e International Center for Not-for-Profit Law.
[44] A/HRC/44/24.
[45] Contribución de Privacy International.
[46] A/HRC/44/57, párrs. 39 y 40.
[47] A/HRC/44/24, párr. 33. Véase también Tribunal Europeo de Derechos Humanos, *Reklos and Davourlis v. Greece*, demanda núm. 1234/05, sentencia pronunciada el 15 de abril de 2009, párr. 40.
[48] Véase Consejo Europeo de Protección de Datos y Supervisor Europeo de Protección de Datos, dictamen conjunto núm. 5/2021, párr. 30; y contribuciones del International Center for Not-for-Profit Law y Privacy International. Véase también A/HRC/44/24, párr. 34, y A/HRC/41/35.
[49] Contribución de la Unión Europea. Véase también https://nymag.com/intelligencer/2020/01/why-we-should-ban-facial-recognition-technology.htm; y Comisión Europea, Propuesta de Reglamento del Parlamento Europeo y del Consejo por el que se establecen normas armonizadas en materia de inteligencia artificial (Ley de Inteligencia Artificial) y se modifican determinados actos legislativos de la Unión, COM(2021) 206 final, 21 de abril de 2021, art. 5, párr. 1 d).
[50] Véase www.ohchr.org/Documents/Issues/Terrorism/biometricsreport.pdf.
[51] Véase https://www.nature.com/articles/d41586-020-00507-5.
[52] Véase https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780190613501.001.0001/acprof-9780190613501-chapter-7.
[53] Véase https://journals.sagepub.com/doi/10.1177/1529100619832930; y https://pubmed.ncbi.nlm.nih.gov/22509011/.

veracidad de las declaraciones durante los interrogatorios, podría socavar los derechos humanos, como los derechos a la privacidad, a la libertad y a un juicio imparcial.

Sistemas de inteligencia artificial y servicios públicos

29. Los sistemas de IA se utilizan cada vez más para contribuir a la prestación de servicios públicos, a menudo con el objetivo declarado de desarrollar sistemas más eficientes para una prestación de servicios oportuna y precisa. Esto también se observa cada vez más en contextos humanitarios en los que el suministro de bienes y servicios humanitarios puede estar vinculado a sistemas de IA. Aunque persiga objetivos legítimos, incluso loables, la implantación de herramientas de IA en la prestación de servicios públicos y humanitarios puede tener un efecto adverso en los derechos humanos si no se establecen las debidas garantías.

30. La IA se utiliza en diversos servicios públicos, desde para adoptar decisiones relativas al derecho a la asistencia social hasta para señalar a familias que requieren la visita de los servicios de atención a la infancia[54]. Estas decisiones se toman utilizando grandes conjuntos de datos, que no solo incluyen datos en poder del Estado, sino también información obtenida de entidades privadas, como empresas de medios sociales o corredores de datos, que a menudo se recopila fuera de los marcos jurídicos de protección[55]. Además, como normalmente los conocimientos informáticos y el poder sobre los sistemas de IA están en manos de empresas privadas, estos acuerdos suelen implicar que estas últimas obtengan acceso a conjuntos de datos que contienen información sobre grandes segmentos de la población. Esto plantea problemas en relación con la privacidad, así como preocupaciones sobre el modo en que el sesgo histórico inherente a los datos afectará a la toma de decisiones de las autoridades públicas.

31. Una de las principales preocupaciones en relación con el uso de la IA para los servicios públicos es que puede ser discriminatoria, particularmente en lo que respecta a los grupos marginados[56]. El Relator Especial sobre la extrema pobreza y los derechos humanos ha advertido de una “distopía de bienestar digital” en la que se utiliza la contrastación sin restricciones para exponer, encuestar y sancionar a los beneficiarios de la asistencia social y se les imponen condiciones que socavan su autonomía individual y su capacidad de elección[57]. Estas preocupaciones se han puesto de manifiesto recientemente en los Países Bajos, donde una sentencia judicial muy mediática ha prohibido un sistema digital de detección de fraudes a la seguridad social por considerar que vulnera el derecho a la intimidad. El sistema en cuestión otorgaba a las autoridades centrales y locales amplios poderes para compartir y analizar datos que antes se guardaban por separado, incluidos los relativos al empleo, la vivienda, la educación, las prestaciones y el seguro médico, así como otros tipos de datos identificables. Además, la herramienta se centraba en los barrios de bajos ingresos y habitados por minorías, lo que daba lugar a una discriminación *de facto* basada en el origen socioeconómico[58].

Uso de la inteligencia artificial en el contexto laboral

32. Una amplia variedad de empleadores, con empresas de todos los tipos y tamaños, demandan cada vez más el seguimiento y la gestión de los trabajadores mediante tecnologías basadas en datos, incluidos sistemas de IA. La llamada “analítica de recursos humanos”

[54] Véase A/74/493.

[55] Contribución de Privacy International.

[56] Contribución de Digital Rights Watch. Para un análisis exhaustivo de los efectos dispares de la automatización en los sistemas de bienestar social, véase Virginia Eubanks, *Automating Inequality* (Nueva York, St. Martin's Press, 2018).

[57] Véase A/74/493. Véase también la carta IRL 1/2020 del Relator Especial, en la que se expresaban preocupaciones similares en relación con una tarjeta de servicios digitales, y la respuesta correspondiente. En el presente informe se hacen varias referencias a las comunicaciones enviadas por los titulares de mandatos de procedimientos especiales del Consejo de Derechos Humanos. Todas esas comunicaciones y las respuestas correspondientes pueden consultarse en el siguiente enlace: https://spcommreports.ohchr.org/Tmsearch/TMDocuments.

[58] Véase www.ohchr.org/EN/NewsEvents/Pages/DisplayNews.aspx?NewsID=25522.

aspira a aportar información más eficaz y objetiva sobre los empleados. Esto puede incluir la toma de decisiones automatizada para la contratación, los planes de promoción o el despido.

33. Aunque la mayor parte de estas tecnologías se centran en el control del comportamiento y el rendimiento relacionados con el trabajo, la gama de aplicaciones de los sistemas de IA también se extiende a comportamientos y datos no relacionados con el trabajo[59]. La pandemia de COVID-19 ha acelerado esta tendencia de dos maneras. En primer lugar, algunas empresas que ofrecen a los trabajadores programas de prevención en materia de salud recogen cada vez más datos sanitarios. En segundo lugar, como el número de procesos realizados por vía digital ha aumentado con el teletrabajo, la supervisión de los sistemas de IA en el lugar de trabajo se ha introducido en los hogares de las personas. Estas dos tendencias aumentan el riesgo de fusionar los datos procedentes de la vigilancia en el lugar de trabajo con las entradas de datos no relacionados con la esfera laboral. Estas prácticas de supervisión basadas en la IA conllevan enormes riesgos para la privacidad a lo largo de todo el ciclo de vida de los datos. Además, los datos pueden utilizarse para fines distintos de los inicialmente comunicados a los empleados, lo que daría lugar a la llamada desviación de uso[60]. Al mismo tiempo, la base cuantitativa de ciencias sociales de muchos sistemas de IA utilizados para la gestión de personal no es sólida, y tiende a presentar sesgos. Por ejemplo, si una empresa utiliza un algoritmo de contratación entrenado a partir de conjuntos de datos históricos que favorecen a los hombres blancos de mediana edad, el algoritmo resultante desfavorecerá a las mujeres, a las personas de color y a las personas más jóvenes o mayores que habrían estado igualmente cualificadas para cubrir la vacante[61]. Al mismo tiempo, suelen faltar las estructuras de responsabilidad y la transparencia necesarias para proteger a los trabajadores, y estos se enfrentan cada vez más a escasas o nulas explicaciones sobre las prácticas de control basadas en la IA[62]. Aunque en algunas situaciones las empresas tienen un interés genuino en evitar las faltas de conducta en el lugar de trabajo, las medidas para defender ese interés no suelen justificar las prácticas ampliamente invasivas utilizadas para la cuantificación de los modos de interacción social y los objetivos de rendimiento conexos en el trabajo. En un entorno laboral, y teniendo en cuenta la relación de poder entre el empleador y el empleado, también se pueden prever posibles situaciones en las que los trabajadores se vean obligados a renunciar a su derecho a la privacidad a cambio de trabajo[63].

Inteligencia artificial para la gestión de la información en línea

34. Las plataformas de medios sociales utilizan sistemas de IA en apoyo a la adopción de decisiones relativas a la gestión de contenidos[64]. Las empresas recurren a estos sistemas para clasificar los contenidos y decidir cuáles ampliar y cuáles recortar, llegando a personalizar estas decisiones para usuarios particulares en función de sus perfiles. La automatización también se utiliza para aplicar restricciones de contenido, en particular para dar cumplimiento a diferentes prescripciones legales dentro de la misma jurisdicción o entre jurisdicciones[65]. La obligación impuesta a los intermediarios de aplicar filtros para evitar presuntos daños a través de Internet podría extender la confianza generalizada en la IA sin tener en cuenta los graves efectos de estos sistemas en los derechos a la privacidad y la libertad de expresión a nivel local y mundial.

35. Los vastos conjuntos de datos en los que se basan los sistemas de organización, amplificación y moderación se crean y alimentan continuamente a través de un amplio proceso de seguimiento en línea y de elaboración de perfiles de los usuarios de las plataformas y sus redes personales[66]. Este proceso constante de recogida de información y

[59] Véase https://journals.sagepub.com/doi/10.1177/20539517211013051.

[60] Christl, *Corporate surveillance in everyday life.*

[61] Contribución de Polonia. Véase también www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G.

[62] Véase https://journals.sagepub.com/doi/full/10.1177/2053951720938093.

[63] Véase www.californialawreview.org/print/3-limitless-worker-surveillance/.

[64] Véase www.theverge.com/2020/11/13/21562596/facebook-ai-moderation; y https://journals.sagepub.com/doi/full/10.1177/2053951720943234.

[65] Véase OTH 71/2018 y OTH 73/2020. Para un análisis exhaustivo del filtrado automático de contenido, véase también https://journals.sagepub.com/doi/full/10.1177/2053951720920686.

[66] A/73/348, párr. 17.

realización de inferencias a partir de ella, unido a la extrema concentración del mercado, ha llevado a que unas cuantas empresas posean y controlen a nivel mundial los perfiles de miles de millones de personas y la esfera pública en red en sentido amplio.

36. La organización de contenidos asistida por IA y realizada por empresas con un enorme poder de mercado suscita preocupación por sus efectos en la capacidad de las personas para formarse opiniones y desarrollarlas, como han señalado dos titulares sucesivos del mandato de Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión[67]. Además, los sistemas de recomendación de las plataformas tienden a centrarse en obtener el mayor grado de participación posible de los usuarios basándose en la información sobre las preferencias de las personas, así como en patrones demográficos y de comportamiento, lo que se ha demostrado que promueve contenidos sensacionalistas y puede reforzar las tendencias a la polarización[68]. Además, la información orientada de forma específica puede resultar inoportuna e incluso conducir a intrusiones peligrosas en la privacidad. Por ejemplo, se han dado casos en que los sistemas de recomendación han llevado a algunos supervivientes de violencia a descubrir a su agresor entre las sugerencias de amistad que les ofrecían las plataformas de medios sociales, y viceversa, lo que supondría un peligro para los supervivientes. Además, se ha demostrado que el sesgo de los grupos mayoritarios o dominantes, reflejado en los datos de los resultados de las búsquedas, afecta a la información sobre los grupos minoritarios o vulnerables o compartida por estos. Por ejemplo, la investigación ha demostrado un grado preocupante de sesgo de género[69] y racial en los resultados de búsqueda de Google[70].

IV. Abordar los problemas

37. La necesidad de adoptar un enfoque basado en los derechos humanos en lo que respecta a las nuevas tecnologías en general, y a la inteligencia artificial en particular, ha sido reconocida por un número creciente de expertos y partes interesadas, así como por la comunidad internacional[71]. Dicho enfoque proporciona herramientas para ayudar a las sociedades a identificar formas de prevenir y limitar los daños obteniendo los máximos beneficios posibles de los avances tecnológicos.

A. Principios fundamentales

38. Un enfoque de la IA basado en los derechos humanos exige la aplicación de varios principios básicos, como la igualdad y la no discriminación, la participación y la rendición de cuentas, principios que también constituyen el eje central de los Objetivos de Desarrollo Sostenible y los Principios Rectores sobre las Empresas y los Derechos Humanos. Además, a las tecnologías de IA deben aplicárseles sistemáticamente los requisitos de legalidad, legitimidad, necesidad y proporcionalidad[72]. Asimismo, la IA debería implantarse de modo que facilite el ejercicio efectivo de los derechos económicos, sociales y culturales garantizando que se cumplan sus elementos clave de disponibilidad, asequibilidad, accesibilidad y calidad[73]. Las personas que sufren vulneraciones de sus derechos humanos y

[67] *Ibid*., párr. 25, y A/HRC/47/25, párr. 36.
[68] Véase www.brookings.edu/techstream/how-youtube-helps-form-homogeneous-online-communities/.
[69] Contribuciones de Austria y Alemania.
[70] Safiya Umoja Noble, *Algorithms of Oppression* (Nueva York, New York University Press, 2018).
[71] Resolución 75/176 de la Asamblea General, párr. 6; resoluciones 47/16, párr. 8 d), y 47/23, decimosexto párrafo del preámbulo; A/73/348, párrs. 47 a 60, A/75/590, párr. 57, y A/HRC/43/29 del Consejo de Derechos Humanos; y contribuciones de Austria, el Comisionado de Protección de la Vida Privada de Canadá, Digital Rights Watch, Global Network Initiative y Privacy International.
[72] A/HRC/43/29, párr. 41.
[73] Véase el análisis detallado del papel de las nuevas tecnologías en el ejercicio efectivo de los derechos económicos, sociales y culturales en A/HRC/43/29.

abusos relacionados con el uso de la IA deben tener acceso a recursos judiciales y no judiciales efectivos[74].

39. Como se ha señalado anteriormente, las restricciones del derecho a la privacidad deben estar previstas en la ley y ser necesarias y proporcionadas para alcanzar un objetivo legítimo. En la práctica, eso significa que los Estados tienen que valorar detenidamente si una medida podrá alcanzar un objetivo establecido, hasta qué punto es importante ese objetivo y qué efectos tendrá esa medida. Los Estados también deben determinar si podrían lograr los mismos resultados con la misma eficacia aplicando enfoques menos invasivos y, en caso afirmativo, deberían adoptar esas medidas. La Alta Comisionada ya ha señalado la necesidad de esos límites y garantías en el contexto de la vigilancia por parte de los organismos de inteligencia y las fuerzas del orden[75]. Cabe señalar que las pruebas de necesidad y proporcionalidad también pueden llevar a la conclusión de que no deben adoptarse determinadas medidas. Por ejemplo, los requisitos de conservación general e indiscriminada de datos de comunicaciones impuestos a las empresas de telecomunicaciones y de otros sectores no pasarían la prueba de proporcionalidad[76]. Del mismo modo, sería desproporcionado imponer requisitos de identificación biométrica a los beneficiarios de la asistencia social sin ofrecerles una alternativa. Además, es fundamental que estas medidas no se evalúen de forma aislada, sino que se tengan debidamente en cuenta los efectos acumulativos de medidas distintas que interactúan. Por ejemplo, antes de decidir implantar nuevas herramientas de vigilancia basadas en la IA, un Estado debe hacer balance de las capacidades ya existentes y de sus efectos en el disfrute del derecho a la privacidad y otros derechos.

B. Legislación y reglamentación

40. La protección efectiva del derecho a la privacidad y los derechos conexos depende de los marcos jurídicos, reglamentarios e institucionales establecidos por los Estados[77].

41. Con la aparición de los sistemas de IA basados en datos, las medidas de protección jurídica efectivas en el marco de las leyes sobre privacidad de los datos han ganado importancia. Estas protecciones deben cumplir las normas mínimas señaladas en el anterior informe del Alto Comisionado sobre el derecho a la vida privada[78].

42. Los marcos relativos a la privacidad de los datos deben tener en cuenta las nuevas amenazas relacionadas con el uso de la IA[79]. Por ejemplo, las leyes podrían imponer limitaciones al tipo de datos que pueden inferirse y/o utilizarse y compartirse posteriormente de manera legal. Los legisladores también deberían valorar la posibilidad de reforzar los derechos de las personas, en particular reconociéndoles el derecho a una explicación significativa y a oponerse a las decisiones totalmente automatizadas que afectan a sus derechos[80]. A medida que evolucionen las tecnologías de IA, habrá que continuar reforzando las garantías previstas en los marcos de protección de la privacidad de los datos.

[74] Pacto Internacional de Derechos Civiles y Políticos, art. 2, párr. 3, y Principios Rectores sobre las Empresas y los Derechos Humanos, principio 15 c) y tercer pilar.

[75] A/HRC/30/39, párrs. 34 a 41

[76] *Ibid*., párr. 18; y Tribunal de Justicia de la Unión Europea, *Digital Rights Ireland and Others*, C-293/12 y C-594/12, párr. 69. Véase también Tribunal de Justicia de la Unión Europea, *Maximilan Schrems v. Data Protection Commissioner*, C-362/14, parr. 94, en el que se considera que "una normativa que permite a las autoridades públicas acceder de forma generalizada al contenido de las comunicaciones electrónicas lesiona el contenido esencial del derecho fundamental al respeto de la vida privada".

[77] A/HRC/39/29, párr. 26.

[78] *Ibid*., párrs. 28 a 33.

[79] El protocolo del Consejo de Europa que modifica el Convenio para la Protección de las Personas con respecto al Tratamiento Automatizado de Datos de Carácter Personal, adoptado en 2018, por ejemplo, es una respuesta a la aparición de nuevas prácticas de tratamiento de datos.

[80] Véase el Reglamento General de Protección de Datos de la Unión Europea, que recoge tales derechos, y la Ley de Derechos de Privacidad de California, que autoriza al organismo de reglamentación a adoptar normas a tal efecto.

43. Los organismos independientes de supervisión de la privacidad de los datos constituyen un elemento clave para contrarrestar la complejidad y opacidad crecientes del entorno mundial de datos, incluidas sus enormes asimetrías de información. Estos organismos deben disponer de facultades efectivas de aplicación de la ley y contar con los recursos adecuados. Las organizaciones de la sociedad civil deberían estar capacitadas para prestar apoyo a la aplicación de las leyes sobre privacidad de los datos, en particular mediante el establecimiento de mecanismos sólidos de denuncia.

44. Más allá de la legislación sobre la privacidad de los datos, es necesario revisar y, si es posible, adoptar una gama más amplia de leyes para abordar los desafíos que plantea la IA de modo que el respeto de los derechos quede garantizado[81].

45. Teniendo en cuenta la diversidad de aplicaciones, sistemas y usos de la IA, la regulación debe ser lo suficientemente específica como para abordar los problemas propios de cada sector y adaptar las respuestas a los riesgos que puedan plantearse[82]. Cuanto mayor sea el riesgo para los derechos humanos, más estrictos deben ser los requisitos legales para el uso de la tecnología de IA. Por lo tanto, los sectores en los que están más en juego los intereses de las personas, como la aplicación de la ley, la seguridad nacional[83], la justicia penal, la protección social, el empleo, la atención de la salud la educación y el sector financiero, deberían ser prioritarios. Un enfoque de la legislación y la reglamentación proporcionado a los riesgos exigirá la prohibición de determinadas tecnologías, aplicaciones o usos de la IA si tienen efectos potenciales o reales que no están justificados en virtud del derecho internacional de los derechos humanos, en particular si no superan las pruebas de necesidad y proporcionalidad. Además, los usos de la IA que son intrínsecamente incompatibles con la prohibición de la discriminación, no deberían autorizarse. Por ejemplo, con arreglo a estos principios, la puntuación social *(social scoring)* de las personas por parte de los gobiernos[84] o los sistemas de IA que clasifican a las personas en grupos por motivos discriminatorios ilícitos[85] deberían prohibirse. En el caso de los sistemas cuya utilización presenta riesgos para los derechos humanos cuando se implantan en determinados contextos, los Estados deberán regular su uso y venta para prevenir y mitigar los efectos adversos sobre los derechos humanos[86] que puedan tener tanto dentro como fuera del territorio del Estado. Debe imponerse la intervención humana obligatoria en la supervisión o adopción de decisiones cuando haya probabilidades de que se produzcan efectos adversos sobre los derechos humanos[87]. Como la evaluación y la gestión de los riesgos puede tomar un tiempo, los Estados también deben imponer moratorias en el uso de tecnología que podría ser de alto riesgo, como el reconocimiento facial remoto en tiempo real, hasta que se garantice que su uso no puede vulnerar los derechos humanos.

46. Los Estados también deben adoptar regímenes rigurosos de control de las exportaciones para el comercio transfronterizo de tecnologías de vigilancia, con el fin de impedir la venta de dichas tecnologías cuando exista el riesgo de que estas puedan utilizarse

[81] Véase Consejo de Europa, Recomendación CM/Rec (2020)1 del Comité de Ministros a los Estados miembros sobre el impacto de los sistemas algorítmicos en los derechos humanos.

[82] La propuesta de ley sobre IA de la Unión Europea adopta este enfoque basado en los riesgos. Las contribuciones de Freedom Online Coalition, Global Network Initiative y Global Partners Digital promueven la reglamentación basada en los riesgos.

[83] En A/HRC/27/37 y A/HRC/39/29, la Alta Comisionada aclaró los requisitos para las medidas de vigilancia adoptadas en el contexto de las investigaciones penales y con fines de protección de la seguridad nacional por los que se debería regir la legislación en esa esfera.

[84] Contribución de la Unión Europea; Catelijne Muller, "The impact of artificial intelligence on human rights, democracy and the rule of law", informe al Consejo de Europa, Comité ad hoc sobre Inteligencia Artificial (CAHAI(2020)06-fin), 24 de junio de 2020, párr. 75; y Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), "Proyecto de recomendación sobre la ética de la inteligencia artificial" (SHS/IGM-AIETHICS/2021/JUN/3 Rev.2), 25 de junio de 2021, párr. 26.

[85] Véase Consejo Europeo de Protección de Datos y Supervisor Europeo de Protección de Datos, dictamen conjunto núm. 5/2021, párr. 33.

[86] Contribución de Derechos Digitales.

[87] Véase www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6.

para violar los derechos humanos, en particular para perseguir a defensores de los derechos humanos o periodistas[88].

47. La diversidad de los riesgos derivados de los sistemas de IA hace que sea necesaria una supervisión adecuada, independiente e imparcial, de su desarrollo, implantación y uso. Esta puede estar a cargo de diversos órganos de supervisión administrativa, judicial, cuasijuidicial y/o parlamentaria[89]. Por ejemplo, además de las autoridades responsables de la privacidad de los datos, el sistema de supervisión también podría abarcar agencias de protección del consumidor, organismos de regulación sectorial, organismos de lucha contra la discriminación e instituciones nacionales de derechos humanos. Además, los organismos de regulación intersectorial encargados de supervisar el uso de la IA pueden ayudar a definir normas fundamentales y garantizar la coherencia de las políticas y de su aplicación.

C. Diligencia debida en materia de derechos humanos

48. Los Estados y las empresas deben velar por que se observe de manera general la diligencia debida en materia de derechos humanos cuando se adquieran, desarrollen, implanten y exploten sistemas de IA, así como antes de que se compartan o utilicen los macrodatos que se hayan recogido sobre personas particulares[90]. Además de financiar y dirigir estos procesos, los Estados también pueden exigir a las empresas que observen en general la diligencia debida en materia de derechos humanos, o incentivarlas de otro modo para que lo hagan.

49. Los procesos de diligencia debida en materia de derechos humanos tienen por objetivo identificar, evaluar, prevenir y mitigar los efectos adversos sobre los derechos humanos que una entidad pueda causar o a los que pueda contribuir o estar directamente vinculada[91]. Cuando los procesos de diligencia debida revelen que un uso de la IA es incompatible con los derechos humanos, debido a la falta de medios efectivos para mitigar los daños, este tipo de uso debería abandonarse. La evaluación de los efectos sobre los derechos humanos es un elemento esencial de los procesos de diligencia debida en materia de derechos humanos[92]. La diligencia debida debe observarse a lo largo de todo el ciclo de vida de los sistemas de IA[93]. Debe prestarse especial atención a los efectos desproporcionados de estos sobre las mujeres y las niñas, las personas lesbianas, gais, bisexuales, transgénero y *queer*, las personas con discapacidad, las personas pertenecientes a minorías, las personas mayores, las personas en situación de pobreza y otras personas en situación de vulnerabilidad.

50. Deberían llevarse a cabo consultas significativas con los titulares de derechos potencialmente afectados y la sociedad civil, y para la evaluación de los efectos, incluida la elaboración y evaluación de medidas de mitigación, se debería contar con la participación de expertos multidisciplinarios. Los Estados y las empresas deberían supervisar constantemente los efectos de los sistemas de IA que utilizan para verificar si afectan negativamente a los derechos humanos. Los resultados de las evaluaciones de impacto sobre los derechos

[88] A/HRC/41/35, párr. 49, y A/HRC/44/24, párr. 40. En estos informes, el Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión y la Alta Comisionada pedían también una moratoria de la concesión de licencias para la exportación de tecnologías de vigilancia.

[89] Véase https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64.

[90] En el marco del proyecto B-Tech de la Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos (ACNUDH) se están elaborando orientaciones sobre la aplicación de los Principios Rectores sobre las Empresas y los Derechos Humanos en el sector tecnológico, incluidas medidas de respuesta a los efectos del uso de las tecnologías de IA sobre los derechos humanos. Véase www.ohchr.org/EN/Issues/Business/Pages/B-TechProject.aspx.

[91] Para una visión general del ejercicio de la diligencia debida en materia de derechos humanos en el contexto de la IA, véase https://international-review.icrc.org/articles/ai-humanitarian-action-human-rights-ethics-913, págs. 174 a 178.

[92] Para un resumen conciso de los métodos de evaluación del impacto en los derechos humanos, véase https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64.

[93] A/HRC/43/29, párr. 62 g), y A/HRC/44/24, párrs. 38, 53 j) i) y 54 c).

humanos, las medidas adoptadas para hacer frente a los riesgos para los derechos humanos y las propias consultas públicas deberían divulgarse[94].

D. Vínculo entre el Estado y las empresas

51. Las situaciones en las que existe un vínculo estrecho entre un Estado y una empresa tecnológica requieren una atención específica[95]. El Estado es un importante agente económico que puede influir en el desarrollo y el uso de la IA, más allá de la labor estatal de adoptar medidas jurídicas y políticas. Cuando los Estados colaboren con desarrolladores de IA y proveedores de servicios del sector privado, deberán tomar medidas adicionales para garantizar que la IA no se utiliza con fines incompatibles con los derechos humanos. Estas medidas deberían aplicarse en la gestión de las empresas estatales, la financiación de la investigación y el desarrollo, el apoyo financiero y de otro tipo que los Estados prestan a las empresas de tecnología de IA, las medidas de privatización y las prácticas de contratación pública.

52. Cuando los Estados actúan como agentes económicos, siguen siendo los principales garantes de derechos en virtud del derecho internacional de los derechos humanos y deben cumplir sus obligaciones de forma proactiva. Al mismo tiempo, las empresas siguen teniendo la responsabilidad de respetar los derechos humanos cuando colaboran con los Estados y deben buscar la manera de cumplir sus obligaciones en esa esfera cuando se enfrentan a exigencias del Estado que entran en conflicto con las normas de derechos humanos[96]. Por ejemplo, cuando se enfrentan a demandas de acceso a datos personales que no cumplen las normas de derechos humanos, deben utilizar su influencia para resistirse o mitigar el daño que se podría causar[97].

53. Los Estados pueden reforzar la protección de los derechos humanos exigiendo sistemáticamente una conducta empresarial responsable. Por ejemplo, cuando los organismos de crédito a la exportación ofrezcan apoyo a las empresas de tecnología de IA, deben velar por que estas empresas tengan un sólido historial de conducta respetuosa con los derechos y puedan demostrarlo a través de procesos rigurosos de diligencia debida.

54. Cuando los Estados confíen en empresas de IA para el suministro de bienes o servicios públicos, deben asegurarse de poder supervisar el desarrollo y la implantación de los sistemas de IA. Esto puede hacerse exigiendo información sobre la precisión y los riesgos de una determinada aplicación de IA y evaluando dicha información. Cuando sea imposible mitigar eficazmente los riesgos, los Estados no deben utilizar la IA para suministrar bienes o servicios públicos.

E. Transparencia

55. Los desarrolladores, comercializadores, operadores y usuarios de sistemas de IA deberían aumentar drásticamente sus medidas para promover la transparencia en lo que respecta al uso de estos sistemas. Como primer paso, los Estados, las empresas y otros usuarios de IA deberían facilitar información sobre el tipo de sistemas que utilizan, los fines que persiguen con su uso y la identidad del desarrollador y el operador de los sistemas[98]. Las personas afectadas deben ser informadas sistemáticamente cuando se tomen o se hayan tomado decisiones de forma automática o con la ayuda de herramientas de automatización[99]. También se les debe informar cuando los datos personales que proporcionan pasen a formar

[94] A/73/348, párr. 68.

[95] Véase www.ohchr.org/Documents/Issues/Business/B-Tech/b-tech-foundational-paper-state-duty-to-protect.pdf.

[96] Principios Rectores sobre las Empresas y los Derechos Humanos, principio 23 b).

[97] A/HRC/32/38, párr. 58. Véase también www.ohchr.org/Documents/Issues/Terrorism/biometricsreport.pdf, págs. 39 y 40.

[98] A/HRC/43/29, párr. 52, y A/73/348, párr. 49.

[99] Consejo de Europa, “Guidelines on addressing the human rights impacts of algorithmic systems”, (apéndice de la Recomendación CM/Rec(2020)1 del Comité de Ministros a los Estados miembros sobre el impacto de los sistemas algorítmicos en los derechos humanos), secc. B, párr. 4.2.

parte de un conjunto de datos utilizados por un sistema de IA[100]. Además, en el caso de las aplicaciones de importancia fundamental para los derechos humanos, los Estados deben establecer registros que contengan información esencial sobre las herramientas de IA y su utilización[101]. Debe garantizarse el cumplimiento efectivo de las obligaciones de transparencia y los derechos de acceso, supresión y rectificación de datos previstos en los marcos de protección de la privacidad de los datos. Debe prestarse especial atención a facilitar que los individuos comprendan y controlen mejor los perfiles recopilados en relación con ellos[102].

56. Las medidas de promoción de la transparencia deberían ir más lejos e incluir esfuerzos constantes para superar el problema de la "caja negra" descrito anteriormente. El desarrollo y la implantación sistemática de metodologías destinadas a hacer que los sistemas de IA sean más explicables —lo que a menudo se denomina transparencia algorítmica— es de suma importancia para garantizar una protección adecuada de los derechos[103]. Esto resulta crucial cuando la IA se utiliza para determinar cuestiones decisivas en el marco de procesos judiciales o en relación con servicios sociales que son esenciales para el ejercicio efectivo de los derechos económicos, sociales y culturales. Los investigadores ya han desarrollado una serie de enfoques que favorecen ese objetivo[104], y es esencial aumentar las inversiones en esa esfera. Los Estados también deben adoptar medidas para garantizar que las disposiciones para la protección de la propiedad intelectual no impidan un examen significativo de los sistemas de IA que tienen efectos sobre los derechos humanos[105]. Las normas de contratación pública deben actualizarse para reflejar la necesidad de transparencia, lo que incluye la posibilidad de supervisar los sistemas de IA[106]. En particular, los Estados deberían evitar el uso de sistemas de IA que puedan tener efectos materiales adversos sobre los derechos humanos pero que no puedan ser objeto de una supervisión significativa[107].

V. Conclusiones y recomendaciones

A. Conclusiones

57. **En el presente informe se han puesto de relieve los efectos innegables y cada vez mayores de las tecnologías de IA en el ejercicio del derecho a la privacidad y otros derechos humanos, tanto para bien como para mal. En él se señala que se han producido eventos inquietantes, como la expansión de un ecosistema caracterizado por procesos en gran medida opacos de recopilación e intercambio de datos personales que subyacen a ciertas partes de los sistemas de IA utilizados de forma generalizada. Estos sistemas afectan a los enfoques gubernamentales de las actividades policiales y la administración de justicia, determinan la accesibilidad de los servicios públicos, deciden quién tiene la opción de ser contratado para un trabajo y afectan a la información que la gente ve y puede compartir en línea. Además, el riesgo de discriminación ligado a las decisiones basadas en la IA es demasiado real. El informe presenta diversas formas de abordar los problemas fundamentales asociados con la IA, subrayando que solo un enfoque integral**

[100] A/73/348, párr. 49.
[101] A/HRC/43/29, párr. 52. La propuesta de ley sobre IA de la Unión Europea contiene disposiciones relativas al registro de sistemas de IA de alto riesgo.
[102] Véase https://link.springer.com/article/10.1007/s12394-008-0003-1, pág. 67.
[103] Para una visión general de los elementos de la transparencia algorítmica, véase www.cambridge.org/core/journals/international-and-comparative-law-quarterly/article/international-human-rights-law-as-a-framework-for-algorithmic-accountability/1D6D0A456B36BA7512A6AFF17F16E9B6, págs. 320 a 323.
[104] Véase https://arxiv.org/abs/2001.00973 and https://arxiv.org/pdf/1711.01134.pdf.
[105] Consejo de Europa, "Guidelines on addressing the human rights impacts of algorithmic systems" (apéndice de la Recomendación CM/Rec(2020)1 del Comité de Ministros a los Estados miembros sobre el impacto de los sistemas algorítmicos en los derechos humanos), secc. B, párr. 4.1.
[106] Véanse las contribuciones de Alemania, Derechos Digitales, Freedom Online Coalition y Global Partners Digital.
[107] A/73/348, párr. 55, y A/HRC/43/29, párr. 54.

basado en los derechos humanos puede garantizar soluciones sostenibles beneficiosas para todos.

58. **No obstante, dada la diversidad de cuestiones nuevas que surgen en el contexto de la IA, el presente informe no es más que una instantánea de una situación en constante evolución. Entre las esferas que merecen un mayor análisis se encuentran la sanidad, la educación, la vivienda y los servicios financieros. Los Estados, las organizaciones internacionales y las empresas tecnológicas hacen cada vez más recurso de las tecnologías biométricas, por lo que se necesita urgentemente más orientación sobre derechos humanos en lo que respecta a esta esfera. Además, uno de los puntos centrales de las labores futuras en relación con los derechos humanos debería ser la búsqueda de fórmulas para llenar el inmenso vacío de responsabilidad que existe en el entorno mundial de datos. Por último, urge identificar y poner en práctica soluciones para hacer frente a la discriminación facilitada por la IA.**

B. Recomendaciones

59. **La Alta Comisionada recomienda a los Estados que:**

a) Reconozcan plenamente la necesidad de proteger y reforzar todos los derechos humanos en el desarrollo, el uso y la gobernanza de la IA como objetivo fundamental, y garanticen en la misma medida el respeto y la observancia de todos los derechos humanos, tanto en línea como en entornos no electrónicos;

b) Velen por que el uso de la IA respete todos los derechos humanos y por que cualquier injerencia en el derecho a la privacidad y otros derechos humanos mediante el uso de la IA esté prevista en la ley, persiga un objetivo legítimo, cumpla con los principios de necesidad y proporcionalidad y no comprometa la esencia de los derechos en cuestión;

c) Prohíban expresamente las aplicaciones de IA que no puedan utilizarse con arreglo a la legislación internacional sobre derechos humanos e impongan moratorias a la venta y el uso de sistemas de IA que entrañen un alto riesgo para el disfrute de los derechos humanos, a menos que se establezcan garantías adecuadas para proteger tales derechos y hasta el momento en que eso suceda;

d) Impongan una moratoria del uso de la tecnología de reconocimiento biométrico a distancia en los espacios públicos, al menos hasta que las autoridades responsables puedan demostrar el cumplimiento de las normas de privacidad y protección de datos, así como la ausencia de problemas significativos de precisión y de efectos discriminatorios, y hasta que se apliquen las recomendaciones formuladas en el párrafo 53 j) (i a v) del documento A/HRC/44/24;

e) Adopten y apliquen de forma efectiva, a través de autoridades independientes e imparciales, legislación sobre privacidad de los datos para los sectores público y privado como requisito indispensable para la protección del derecho a la privacidad en el contexto de la IA;

f) Adopten marcos legislativos y reglamentarios que prevengan y mitiguen adecuadamente los efectos adversos sobre los derechos humanos de diversa naturaleza vinculados al uso de la IA por parte de los sectores público y privado;

g) Garanticen que las víctimas de violaciones de derechos humanos y abusos relacionados con el uso de sistemas de IA tengan acceso a recursos efectivos;

h) Exijan la explicabilidad adecuada de todas las decisiones basadas en la IA que puedan afectar significativamente a los derechos humanos, particularmente en el sector público;

i) Aumenten las medidas de lucha contra la discriminación vinculada al uso de sistemas de IA por parte de los Estados y las empresas, entre otras cosas realizando, exigiendo y promoviendo evaluaciones sistemáticas y supervisando los resultados de los sistemas de IA y los efectos de su implantación;

j) Velen por que las asociaciones público-privadas sean transparentes en el suministro y el uso de tecnologías de IA y estén sujetas a una supervisión independiente en lo que respecta a los derechos humanos, y que no se traduzcan en una renuncia de los gobiernos a sus responsabilidades en materia de derechos humanos.

60. **La Alta Comisionada recomienda a los Estados y las empresas que:**

a) Observen sistemáticamente la diligencia debida en materia de derechos humanos a lo largo del ciclo de vida de los sistemas de IA que diseñan, desarrollan, implantan, venden, obtienen o explotan. Un elemento fundamental de la diligencia debida en materia de derechos humanos debería ser la realización de evaluaciones periódicas y exhaustivas de los efectos de estos sistemas sobre tales derechos;

b) Aumenten considerablemente la transparencia en su uso de la IA, entre otras cosas, informando adecuadamente al público y a las personas afectadas y permitiendo una supervisión independiente y externa de los sistemas automatizados. Cuanto más probables y graves sean los efectos potenciales o reales sobre los derechos humanos relacionados con el uso de la IA, más necesaria será la transparencia;

c) Garanticen la participación de todas las partes interesadas en las decisiones sobre el desarrollo, la implantación y el uso de la IA, en particular de las personas y grupos afectados;

d) Promuevan la explicabilidad de las decisiones basadas en la IA, entre otras cosas mediante la financiación y la realización de investigaciones con ese objetivo.

61. **La Alta Comisionada recomienda a las empresas que:**

a) Hagan todo lo posible para cumplir con su responsabilidad de respetar todos los derechos humanos, en particular dando plena aplicación a los Principios Rectores sobre las Empresas y los Derechos Humanos;

b) Redoblen sus esfuerzos por luchar contra la discriminación relacionada con el desarrollo, la venta o la explotación de sistemas de IA, lo que incluye la realización de evaluaciones sistemáticas y el seguimiento de los resultados de estos sistemas y los efectos de su implantación;

c) Adopten medidas decisivas para garantizar la diversidad del personal encargado del desarrollo de la IA;

d) Den reparación o contribuyan a dar reparación mediante procedimientos legítimos cuando hayan provocado o contribuido a provocar efectos negativos en los derechos humanos, entre otras cosas mediante mecanismos efectivos de presentación de reclamaciones a nivel operativo.