



# Asamblea General

Distr. general  
17 de julio de 2024  
Español  
Original: inglés

---

## Septuagésimo noveno período de sesiones

Tema 71 b) del programa provisional\*

**Promoción y protección de los derechos humanos:  
cuestiones de derechos humanos, incluidos otros  
medios de mejorar el goce efectivo de los derechos  
humanos y las libertades fundamentales**

## Los derechos humanos y la solidaridad internacional

### Nota del Secretario General

El Secretario General tiene el honor de transmitir a la Asamblea General el informe de la Experta Independiente sobre los derechos humanos y la solidaridad internacional, Cecilia M. Bailliet, presentado en cumplimiento de la resolución 53/5 del Consejo de Derechos Humanos.

---

\* [A/79/150](#).



## **Informe de la Experta Independiente sobre los derechos humanos y la solidaridad internacional, Cecilia M. Bailliet**

### **Inteligencia artificial y solidaridad internacional – hacia la solidaridad internacional en el ámbito de la inteligencia artificial centrada en el ser humano desde el diseño**

#### *Resumen*

Este es el primer informe que la Experta Independiente sobre los derechos humanos y la solidaridad internacional, Cecilia M. Bailliet, prepara para la Asamblea General. Se presenta en cumplimiento de la resolución [53/5](#) del Consejo de Derechos Humanos. La Experta Independiente aborda el desafío actual del aumento de la desigualdad global y explica cómo una estrategia de diseño de la solidaridad internacional en el ámbito de la inteligencia artificial puede dilucidar los deberes de los Estados, las empresas y la sociedad civil con miras a fortalecer la igualdad de acceso a la tecnología y la inclusión no discriminatoria de grupos e individuos en situación de vulnerabilidad. En el informe, la Experta Independiente propone que la gobernanza de la inteligencia artificial debería apoyar los mecanismos de procedimiento de diligencia debida de los Estados y las empresas para implicar a las partes interesadas directas e indirectas en el procesamiento de datos y la toma de decisiones en el ciclo de vida de la inteligencia artificial.

## I. Introducción

1. La solidaridad internacional es un principio de derecho internacional y un valor universal que ha captado la atención de los encargados de formular políticas para solucionar problemas globales, como los relacionados con las pandemias, el cambio climático, la desigualdad y la energía limpia<sup>1</sup>. En todo el mundo está surgiendo una nueva conciencia orientada hacia un cambio transformador que se aleje de la desigualdad neoliberal y se oriente hacia una administración, solidaridad y acción colectiva impulsada por una revolución basada en valores<sup>2</sup>. Es imperativo que la inteligencia artificial (IA) se utilice para unir y no para dividir a la humanidad. De acuerdo con los derechos humanos, la IA debe promover acciones colectivas de solidaridad internacional más allá de las fronteras<sup>3</sup>. En 2022, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) publicó una Recomendación sobre la Ética de la Inteligencia Artificial en la que se reconocía la necesidad de reforzar la solidaridad para facilitar un acceso equitativo a las tecnologías de IA y abordar los retos que plantean en la diversidad y la interconectividad de las culturas y los sistemas éticos. Además, la solidaridad puede contribuir a mitigar el posible uso indebido de la IA, ayudar a aprovechar todo su potencial y ayudar a garantizar que las estrategias nacionales de IA se rigen por principios éticos.

2. Como se señala en un estudio, la IA tiene una presencia cada vez mayor en sociedades en las que se están creando “ciudades inteligentes” para facilitar el seguimiento y la supervisión de los ciudadanos, potencialmente en todos los aspectos de la vida, en entornos en los que aumenta el uso del registro biométrico y como herramienta en el marco de las acciones de respuesta humanitaria y desarrollo internacional ante los desplazamientos y otras crisis<sup>4</sup>. Se prevé que el uso de la IA en instituciones gubernamentales e instituciones subcontratadas (incluidas empresas y organizaciones de la sociedad civil) para procesar y tomar decisiones en causas y reclamaciones, utilizando datos personales, aumente significativamente en todo el mundo. La vigilancia mediante IA afecta especialmente a las personas pobres, ya que las instituciones estatales la utilizan para detectar el uso indebido de prestaciones sociales, como herramienta de “sobrevigilancia” de barrios marginados y en el contexto del seguimiento de la migración irregular<sup>5</sup>. Nunca se insistirá lo suficiente en la importancia de reconocer la vulnerabilidad intersectorial a la discriminación en el ámbito de la IA por motivos de raza, etnia, religión, género, ubicación, nacionalidad y situación socioeconómica, entre otros<sup>6</sup>.

3. Además, en el *Artificial Intelligence Index Report 2024* se indicó que “los desarrolladores de IA no son transparentes, sobre todo en lo que respecta a la divulgación de metodologías y datos de entrenamiento. Esta falta de apertura dificulta los esfuerzos por comprender mejor la solidez y seguridad de los sistemas de IA”<sup>7</sup>. Además, las empresas de tecnología de la información y las comunicaciones (TIC) colaboran en mayor medida con las organizaciones de la sociedad civil del Norte Global, lo que puede debilitar la solidaridad con las organizaciones de la sociedad

<sup>1</sup> Véase Global Nation, *Global Solidarity Report 2023* (septiembre de 2023).

<sup>2</sup> Kurt April, “AI-Induced Solidarity Economy: The Need for Stewardship Orientation” en *Effective Executive*, vol. 26, núm. 3 (2023).

<sup>3</sup> Contribución del Foro para la Gobernanza de Internet. Véase también <https://intgovforum.org/en/content/pnai-report>.

<sup>4</sup> Linnet Taylor, “What is data justice? The case for connecting digital rights and freedoms globally”, *Big Data and Society*, vol. 4, núm. 2 (julio-diciembre de 2017).

<sup>5</sup> *Ibid.*

<sup>6</sup> *Ibid.*

<sup>7</sup> Institute for Human-Centered AI, Universidad de Stanford, *Artificial Intelligence Index Report 2024* (Universidad de Stanford, California, 2024).

civil del Sur Global<sup>8</sup>. Los esfuerzos por corregir ese desequilibrio se ven complicados por el limitado acceso de los grupos del Sur Global a las redes que abordan la rendición de cuentas de la IA por violaciones de derechos humanos.

4. Para elaborar el informe, la Experta Independiente recabó contribuciones de Estados Miembros, organizaciones de la sociedad civil, empresas y círculos académicos. En julio de 2024 se llevó a cabo una consulta con las partes interesadas de la sociedad civil. Se hizo una revisión de la literatura académica y los informes de los órganos creados en virtud de tratados de las Naciones Unidas y los órganos basados en la Carta de las Naciones Unidas, y se analizaron las normas jurídicas internacionales, regionales y nacionales.

## II. Criterios de solidaridad de una inteligencia artificial responsable: transparencia, equidad, no discriminación e inclusión

5. Dado que el procesamiento de datos y la toma de decisiones se delegan cada vez más en la IA, existe una necesidad fundamental de transparencia para garantizar la equidad, la no discriminación y la inclusión. En el *Artificial Intelligence Index Report 2024* se define la equidad como “la creación de algoritmos equitativos, evitando los sesgos o la discriminación y teniendo en cuenta las diversas necesidades y circunstancias de todas las partes interesadas, y alineándose así con las normas sociales más amplias de equidad”<sup>9</sup>. En el informe se indica además que “los resultados muestran que, aunque la mayoría de las empresas han aplicado plenamente al menos una medida de equidad, sigue faltando una integración global. El promedio global de medidas de equidad adoptadas se sitúa en 1,97 de un máximo de 5 medidas” para las que se recopilaban datos en la encuesta Global State of Responsible AI. La World Benchmarking Alliance señaló que en 2023, solo una cuarta parte de las 200 empresas tecnológicas más influyentes del mundo cumplían una norma mínima en la divulgación de la adopción de principios éticos de IA sobre inclusión digital<sup>10</sup>.

6. En el Llamamiento de Roma por la Ética de la IA se subraya la inclusión como principio clave, y en el evento AI Ethics for Peace celebrado en Hiroshima en julio de 2024, el presidente del Foro de Abu Dabi por la Paz, el jeque Abdallah Bin Bayyah, subrayó que “la cooperación, la solidaridad y el trabajo conjunto son necesarios para hacer frente a los avances de la inteligencia artificial, en la que se mezclan intereses, perjuicios y beneficios, con miras a garantizar que sus sistemas y productos no solo sean técnicamente avanzados, sino también moralmente sólidos”<sup>11</sup>. El borrador del Pacto Digital Global de 26 de junio de 2024 establece un objetivo inclusivo específico: “Nuestra cooperación cerrará las brechas digitales dentro de los Estados y entre ellos y avanzará hacia un entorno digital que promueva y permita la diversidad”<sup>12</sup>.

7. De acuerdo con los Principios Rectores sobre las Empresas y los Derechos Humanos, cabe esperar que las empresas tecnológicas lleven a cabo evaluaciones del impacto sobre los derechos humanos y detecten los riesgos como parte de sus

<sup>8</sup> Véase <https://www.business-humanrights.org/en/from-us/briefings/dismantling-the-facade-a-global-south-perspective-on-the-state-of-engagement-with-tech-companies/dismantling-the-facade-a-global-south-perspective-on-the-state-of-engagement-with-tech-companies/>.

<sup>9</sup> *Artificial Intelligence Index Report 2024*.

<sup>10</sup> Véase <https://www.worldbenchmarkingalliance.org/impact/investor-statement-for-ethical-ai-2024/>.

<sup>11</sup> Véase <https://www.romecall.org/> y <https://www.romecall.org/ai-ethics-for-peace-hiroshima-july-9th-2024/>.

<sup>12</sup> Véase el borrador de 26 de junio, disponible en <https://www.un.org/techenvoy/global-digital-compact>.

procesos de control de calidad, y que colaboren con las partes interesadas externas en las evaluaciones de derechos humanos<sup>13</sup>. Se puede sugerir que un enfoque de solidaridad internacional en el ámbito de la IA se centraría en la necesidad de establecer de manera concreta enfoques de procedimiento para abordar la discriminación y reforzar la inclusión. Se recomienda incorporar la solidaridad intergeneracional a la normativa sobre IA<sup>14</sup>.

8. La solidaridad social puede invocarse para compartir los beneficios y los costos del despliegue de la IA, promover la diversidad en las trayectorias de IA y fomentar la transparencia y el cumplimiento para corregir las asimetrías de información en la IA<sup>15</sup>. La Experta Independiente destaca la importancia de incorporar la solidaridad como perspectiva al evaluar el diseño y la aplicación de la IA.

## A. Inclusión y no discriminación algorítmica como objetivos solidarios de la inteligencia artificial

9. Es fundamental regular la IA para corregir los sesgos y la discriminación y garantizar la seguridad. La IA puede ser emancipadora al facilitar la educación en lugares remotos, puede proporcionar accesibilidad lingüística mediante funciones de traducción y puede utilizarse para luchar contra los estereotipos y el discurso de odio. Por ello, sigue siendo necesario apoyar la investigación sobre el impacto social de la IA. Las instituciones que están diseñando la IA en el marco de los procesos de toma de decisiones (incluidos los ámbitos judicial y administrativo) deben ser inclusivas y tener en cuenta la opinión de los grupos en situación de vulnerabilidad y los actores cívicos orientados a la democracia, en todas las fases, desde la planificación hasta la aplicación, con el fin de evitar las violaciones de los derechos humanos y mitigar el daño. Además, es importante crear un mecanismo de supervisión independiente que se ocupe de la protección de datos, capaz de regular y establecer directrices sobre la recolección y el tratamiento de datos personales durante las distintas fases de planificación y despliegue de la IA. La Experta Independiente apoya el modelo de disparidad demográfica condicional que sirve como conjunto estándar de pruebas estadísticas para los casos de discriminación automatizada<sup>16</sup>.

10. Una de las contribuciones de este informe presenta la visión de una IA inclusiva que cumpla el principio de solidaridad internacional mediante la creación de plataformas que amplifiquen las perspectivas de poblaciones tradicionalmente infrarrepresentadas en los diálogos mundiales, como la traducción mediante IA, las plataformas de medios sociales utilizadas por los activistas para establecer contactos, y mediante el empoderamiento de la sociedad civil para exigir responsabilidades a los Gobiernos, como en los casos de corrupción<sup>17</sup>.

<sup>13</sup> Kate Jones, *AI Governance and Human Rights: Resetting the relationship*, documento de investigación, Programa de Derecho Internacional (Londres, Royal Institute of International Affairs, 10 de enero de 2023).

<sup>14</sup> Sébastien Fassiaux, "Preserving consumer autonomy through European Union regulation of artificial intelligence: a long-term approach", *European Journal of Risk Regulation*, vol. 14, número especial 4 (diciembre de 2023). Véase también Jon Truby y otros, "A sandbox approach to regulating high-risk artificial intelligence applications", *European Journal of Risk Regulation*, vol. 13, núm. 2 (junio de 2022), en el que se reconoce la solidaridad como principio dentro de la regulación de la IA en la Unión Europea.

<sup>15</sup> Juan C. Mateos-García, "The Complex Economics of Artificial Intelligence" (2 de diciembre de 2018) (documento de trabajo inédito).

<sup>16</sup> Sandra Wachter, Brent Mittelstadt y Chris Russell, "Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI", *Computer Law and Security Review*, vol. 41 (2021).

<sup>17</sup> Contribución de Jake Okechukwu Effoduh, Profesor Adjunto, Lincoln Alexander School of Law, Universidad Metropolitana de Toronto, Canadá.

## B. Normas nacionales sobre inclusión y no discriminación

11. A nivel nacional, existen numerosas normas (la mayoría de ellas aspiracionales, no vinculantes o en forma de borrador) que se ajustan a los principios de igualdad, inclusión y no discriminación. Sin embargo, es frecuente que falte una competencia institucional independiente y mecanismos procesales específicos que garanticen el cumplimiento y que proporcionen recursos en caso de violación.

12. Australia cuenta con principios éticos voluntarios y no vinculantes en materia de IA que reconocen que los sistemas de IA deben beneficiar a las personas, la sociedad y el medio ambiente y respetar los derechos humanos, la diversidad y la autonomía de las personas<sup>18</sup>. Del mismo modo, China cuenta con medidas provisionales para la gestión de servicios de inteligencia artificial generativa que exigen la prevención de la discriminación basada en el origen étnico, la religión, la nacionalidad, el origen geográfico, el género, la edad, la ocupación y la salud, así como la adopción de medidas para aumentar la diversidad y el contenido no discriminatorio de los datos de entrenamiento<sup>19</sup>. La India cuenta con una estrategia nacional para la inteligencia artificial que aborda la igualdad y la no discriminación<sup>20</sup>. En ella se dispone que los sistemas de IA deben tratar por igual a las personas que se encuentren en las mismas circunstancias pertinentes para la decisión, que los sistemas de IA no deben negar oportunidades a una persona calificada por razones de identidad y que no deben ahondar en las perjudiciales divisiones históricas y sociales basadas en la religión, la raza, la casta, el género, la ascendencia, el lugar de nacimiento o de residencia en cuestiones como la educación, el empleo o el acceso a espacios públicos. Además, la estrategia establece como objetivo esforzarse por evitar la exclusión injusta de servicios o prestaciones. El Brasil ha propuesto un reglamento sobre IA en el que figura una norma de no discriminación que incluye el derecho a impugnar decisiones y solicitar la intervención humana y la corrección de sesgos discriminatorios directos, indirectos, ilegales o abusivos<sup>21</sup>. España creó una autoridad independiente para la igualdad de trato y la no discriminación que supervisa y promueve el uso de una IA ética, fiable y respetuosa con los derechos fundamentales<sup>22</sup>.

13. Los Estados Unidos de América tienen un proyecto de ley sobre la IA titulado “Blueprint for an AI bill of rights: making automated systems work for the American people” que apoya un decreto en el que se pide una evaluación independiente y la presentación de informes para garantizar la no discriminación<sup>23</sup>. Kenia cuenta con un proyecto de código de prácticas de IA que exige la creación de un mecanismo para aplicar normas de no discriminación, entre otras cosas mediante la documentación de evaluaciones de imparcialidad, el registro de las medidas adoptadas para hacer frente a los sesgos y la documentación de las políticas contra los resultados discriminatorios<sup>24</sup>.

<sup>18</sup> Véase <https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles>.

<sup>19</sup> Véase <https://www.chinalawtranslate.com/en/generative-ai-interim/>.

<sup>20</sup> Véase <https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf>; <https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf>; <https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf>.

<sup>21</sup> Véase el proyecto de ley núm. 2338 de 2023. Disponible en <https://www25.senado.leg.br/web/atividade/materias/-/materia/157233>.

<sup>22</sup> Contribución de España.

<sup>23</sup> Véase <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>; véase también <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>.

<sup>24</sup> Véase [https://www.dataguidance.com/sites/default/files/kebs-tc\\_094\\_n66\\_public\\_review\\_kenya\\_standard\\_dks\\_3007\\_ai\\_code\\_of\\_practice.pdf](https://www.dataguidance.com/sites/default/files/kebs-tc_094_n66_public_review_kenya_standard_dks_3007_ai_code_of_practice.pdf).

### C. Riesgos de la aplicación sesgada de la inteligencia artificial

14. Se ha reconocido que “el acceso asequible e informado a Internet se ha convertido en una necesidad fundamental para la plena realización de los derechos humanos y libertades fundamentales, la democracia, el desarrollo y la justicia social”<sup>25</sup>. Esto ha dado lugar a una serie de iniciativas privadas en materia de derechos digitales que tratan de abordar la infraestructura, la aplicación y el uso de Internet desde la perspectiva de una mayor accesibilidad. El sesgo de la IA somete a individuos o grupos a estereotipos o prejuicios, comparándolos con otros mediante el uso de algoritmos de IA. El sesgo de la IA aparece en el diseño de un sistema de IA, la recopilación e interpretación de datos y a través de la interacción con las partes interesadas directas e indirectas. La Relatora Especial sobre las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia confirmó el desafío permanente que suponen “los problemas relacionados con los datos, las cuestiones de diseño de algoritmos, el uso intencionadamente discriminatorio de la inteligencia artificial y las cuestiones relativas a la rendición de cuentas” (véase A/HRC/56/68). En un informe sobre la IA elaborado en 2019 por la UNESCO se señala la solidaridad como un elemento de la creación de una “sociedad del conocimiento” y se explica la urgencia de abordar la discriminación de la IA que surge de la programación discriminatoria intencional o no intencional, el sesgo en los datos de entrenamiento para los algoritmos de aprendizaje automático u otros factores<sup>26</sup>:

Muchos tipos de discriminación pueden ser indirectos. Por ejemplo, un algoritmo que utiliza los patrones de uso del teléfono móvil para determinar la solvencia crediticia de una persona es discriminatorio si asigna un alto riesgo crediticio a las mujeres de comunidades que i) tienen un bajo uso del teléfono móvil o ii) no poseen teléfonos móviles. La condición que se aplica puede parecer igual y justa, pero pone en desventaja a un grupo en particular. Los algoritmos pueden causar y exacerbar estas múltiples formas de discriminación. Los prejuicios sociales y políticos existentes se están sistematizando en los algoritmos de aprendizaje automático de muchas maneras. Además, vale la pena investigar las posibles nuevas formas de discriminación que puede provocar la IA, como las exclusiones decididas sobre la base de correlaciones estadísticas que no corresponden necesariamente a características socialmente destacadas, pero que sin embargo están muy vinculadas a la identidad personal de cada uno.

15. En el *Artificial Intelligence Index Report 2024* se explica de qué manera la tokenización del lenguaje (en que el lenguaje se descompone en componentes para su análisis) afecta negativamente a los idiomas no occidentales ya que la eficacia de la IA es mayor en el caso de los idiomas occidentales<sup>27</sup>. También hay organizaciones de la sociedad civil que abordan directamente el sesgo algorítmico, la exclusión y la discriminación de minorías raciales, étnicas o religiosas, así como por motivos de género u otra identidad, y que trabajan para mejorar la accesibilidad<sup>28</sup>. Amnistía Internacional y la National Association for the Advancement of Colored People han

<sup>25</sup> George A. Walker, “Technology Law, Rights and Ethics - One Choice, One Future”, *The International Lawyer*, vol. 56, núm. 1 (2023).

<sup>26</sup> Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), *El aporte de la inteligencia artificial y las TIC avanzadas a las sociedades del conocimiento: una perspectiva de derechos, apertura, acceso y múltiples actores* (París, 2019), pág. 71.

<sup>27</sup> *Artificial Intelligence Index Report 2024*.

<sup>28</sup> Véase <https://digitalrightsfoundation.pk/wp-content/uploads/2021/03/Policy-brief-2.pdf>; <https://www.hrw.org/report/2023/12/21/metals-broken-promises/systemic-censorship-palestine-content-instagram-and>; <https://www.apc.org/es/member/7amleh-arab-center-social-media-advancement>; <https://smex.org/>; <https://www.derechosdigitales.org/>; y <https://www.accessnow.org/>.

expresado su preocupación por las herramientas policiales predictivas<sup>29</sup>, el uso de sistemas automatizados para determinar el acceso a la atención médica y los servicios sociales, la vigilancia de los movimientos de refugiados y migrantes, y el impacto sesgado de las herramientas de IA que se utilizan para el reconocimiento facial y la detección de fraudes en las comunidades racializadas<sup>30</sup>. Las personas sometidas a estos procesos de IA pueden desconocer y no estar al tanto del uso de la IA y no suelen tener acceso a ninguna reparación por el trato discriminatorio que reciben de las instituciones estatales. Los grupos de la sociedad civil insisten en la necesidad de incluir a las partes interesadas en el diseño de normas en materia de IA. Se ha sugerido que la Unión Europea mantiene el uso de la IA, a pesar de las críticas, en sistemas de vigilancia discriminatorios, como los sistemas de evaluación de riesgos y análisis predictivos para facilitar las devoluciones sumarias<sup>31</sup>, mientras que en el contexto de la migración, la IA se está desarrollando con diferentes fines: a) verificación de la identidad mediante el desarrollo de tecnología de reconocimiento facial; b) evaluación de riesgos; c) evaluaciones relacionadas con la detención; y d) vigilancia. En estos contextos se ha destacado el riesgo que supone minimizar el control humano<sup>32</sup>.

16. Dado que las instituciones estatales y las empresas que utilizan la IA para el procesamiento de datos y la toma de decisiones no explican de manera transparente su desarrollo, entrenamiento, uso o aplicación, las personas o grupos que denuncian violaciones de derechos humanos a menudo no pueden presentar las pruebas que relacionan la IA con los actos de violación. Las violaciones pueden caracterizarse en primer lugar por el rápido bucle de retroalimentación entre diferentes sistemas estatales conectados a través de la IA, que da lugar a un trato o una toma de decisiones perjudiciales por parte de la institución estatal. Los delincuentes también utilizan la IA para apropiarse fraudulentamente de las remesas que los emigrantes envían a sus familias, o del dinero que los familiares envían a los migrantes, tanto en el país de origen como en el de residencia.

17. La Asamblea General ha reconocido la necesidad de luchar contra la discriminación algorítmica<sup>33</sup>, detectar vulnerabilidades, mejorar la accesibilidad y ofrecer reparación por las violaciones de los derechos humanos<sup>34</sup>. Los órganos creados en virtud de tratados de derechos humanos y los procedimientos especiales de las Naciones Unidas han prestado especial atención al uso de la IA en las instituciones estatales encargadas de abordar las cuestiones relacionadas con la migración, la delincuencia, la atención sanitaria y la atención a las personas mayores. El Comité para la Eliminación de la Discriminación Racial formuló observaciones finales en las que expresó su preocupación por el impacto discriminatorio del uso de la IA en el contexto del asilo<sup>35</sup>. El Comité para la Eliminación de la Discriminación contra la Mujer ha formulado observaciones finales en las que pide que se establezcan salvaguardias adecuadas para prevenir los estereotipos de género asociados a los

<sup>29</sup> En relación con la vigilancia policial predictiva que conduce a una vigilancia y un control policial desproporcionados de las comunidades negras, véase <https://naacp.org/resources/artificial-intelligence-predictive-policing-issue-brief>.

<sup>30</sup> Véase <https://www.amnesty.org/es/latest/campaigns/2024/01/the-urgent-but-difficult-task-of-regulating-artificial-intelligence/>.

<sup>31</sup> Contribución de Indira Boutier, Universidad Glasgow Caledonian.

<sup>32</sup> *Ibid.* Véase también Petra Molnar y Lex Gill, *Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System*, (Programa Internacional de Derechos Humanos (Facultad de Derecho de la Universidad de Toronto) y Citizen Lab (Munk School of Global Affairs and Public Policy, Universidad de Toronto), 2018), págs. 31 a 34.

<sup>33</sup> Véase la resolución 78/265 de la Asamblea General, párr. 6 h).

<sup>34</sup> Resolución 78/213 de la Asamblea General.

<sup>35</sup> CERD/C/DEU/CO/23-26, párrs. 45 y 46.

sistemas biométricos, de vigilancia y de elaboración de perfiles algorítmicos utilizados por las autoridades policiales en la prevención e investigación de delitos, y adoptar medidas para eliminar el sesgo algorítmico relacionado con la inteligencia artificial y los servicios algorítmicos<sup>36</sup>. La Experta Independiente sobre el disfrute de todos los derechos humanos por las personas de edad expresó su agradecimiento por la oportunidad brindada durante su visita de conocer cómo la inteligencia artificial puede ser útil en relación con el envejecimiento, la atención y los servicios de salud, pero también recomendó que se revisara el uso de la recopilación de datos, a fin de garantizar que el uso de esos datos respete las obligaciones en materia de no discriminación en relación con las personas de edad de diversos orígenes, lo que implica una perspectiva de solidaridad de la IA<sup>37</sup>.

18. Los grupos de la sociedad civil subrayaron el impacto negativo que la IA tuvo en la capacidad de los activistas LGBTQI para participar en acciones de solidaridad.

### III. Iniciativas regionales para abordar los procesos institucionales de inteligencia artificial

19. Los principios de la Organización de Cooperación y Desarrollo Económicos (OCDE) sobre inteligencia artificial se actualizaron en 2024 y subrayan la importancia de fomentar la inclusión de las poblaciones infrarrepresentadas, reducir las desigualdades económicas, sociales, de género y de otro tipo, y promover la diversidad, la equidad, la justicia social, la transparencia, la capacidad de acción humana y la supervisión<sup>38</sup>. Un problema que plantea el creciente interés de las instituciones estatales por utilizar la IA para la tramitación de casos y la toma de decisiones es que muchos de estos sistemas están desarrollados por empresas y, por lo tanto, existe el riesgo de que carezcan de transparencia. Las instituciones estatales pueden no ser conscientes de los riesgos que plantea el uso de la IA, como las “alucinaciones”, es decir, aquellos casos en que la IA crea contenido que no se basa en el contenido, y de la necesidad esencial de un control humano para supervisar el uso de la información y detectar el sesgo algorítmico. El catálogo de la OCDE de herramientas y métricas para una IA fiable ofrece ejemplos de diversos enfoques técnicos, educativos y de procedimiento para perseguir la solidaridad con la IA<sup>39</sup>.

20. La Ley de Inteligencia Artificial de la Unión Europea aborda la solidaridad internacional mediante la detección de prácticas prohibidas de IA en su artículo 5 1) b), c) y g). La Comisión Europea ha propuesto una Directiva sobre responsabilidad para su aplicación por los tribunales nacionales, y que por lo tanto establece normas probatorias<sup>40</sup>. La Comisión Europea aprobó en 2018 la Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno, en la que pidió que se realizaran auditorías externas del procesamiento de datos para detectar casos

<sup>36</sup> CEDAW/C/ITA/CO/8, párr. 26. Véanse también observaciones finales similares en el documento CEDAW/C/DEU/CO/9, párrs. 27 y 28. El Comité expresó su preocupación por la IA y la violencia de género en las observaciones finales que figuran en el documento CEDAW/C/ESP/CO/9, párrs. 21, 23, 31 y 33.

<sup>37</sup> A/HRC/45/14/Add.1, párr. 93.

<sup>38</sup> Véase <https://oecd.ai/en/ai-principles>; véase también <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>.

<sup>39</sup> Véase <https://oecd.ai/en/>.

<sup>40</sup> Véase <https://artificialintelligenceact.eu/es/el-acto/>. Véase también la propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial (Directiva sobre responsabilidad en materia de IA). Disponible en <https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:52022PC0496>.

de discriminación<sup>41</sup>. La Comisión reconoció que el uso de la IA en la gestión de la migración, el asilo y el control de fronteras y en la administración de justicia y los procesos democráticos conlleva un alto riesgo, por lo que en 2024 aprobó la Ley de Inteligencia Artificial. La Ley prevé la creación de una Oficina de Inteligencia Artificial (Oficina de IA) europea dentro de la Comisión para supervisar la aplicación y el cumplimiento efectivos de la Ley por parte de los proveedores de modelos de IA de propósito general. El Consejo de Inteligencia Artificial europeo estará compuesto por representantes de los Estados miembros y de las empresas. La Oficina de IA podrá invitar a los proveedores de modelos de IA de propósito general y a las autoridades nacionales competentes pertinentes a participar en la creación de códigos de buenas prácticas, mientras que la sociedad civil, la industria, el mundo académico, los proveedores intermedios y los expertos independientes podrán apoyar el proceso<sup>42</sup>. Esa norma resulta insuficiente por su vaguedad en cuanto a la capacidad de la sociedad civil para contribuir de forma directa a las evaluaciones del impacto.

21. El Tribunal de Justicia europeo detectó riesgos de discriminación en relación con su futuro uso de la IA en un documento estratégico: “Uno de los principales riesgos asociados a la adopción de la tecnología de IA es la posibilidad de que se introduzcan sesgos involuntarios durante el entrenamiento de los modelos de IA, que den lugar a discriminación involuntaria”<sup>43</sup>. La Subcomisión de Derechos Humanos del Parlamento Europeo realizó un análisis de la IA en el que subrayó el riesgo de que se produzca un afianzamiento de la discriminación estructural: “Los sistemas de IA pueden perpetuar y amplificar los sesgos, lo que conduce a la discriminación en diversos sectores, como el empleo, la aplicación de la ley y la calificación crediticia. Hay pruebas sustanciales que demuestran que la IA puede afianzar las disparidades socioeconómicas al reflejar los prejuicios presentes en los datos de entrenamiento o el diseño”<sup>44</sup>.

22. El Comité de Ministros del Consejo de Europa aprobó el 17 de mayo un Convenio Marco sobre Inteligencia Artificial y Derechos Humanos, Democracia y Estado de Derecho, que se abrirá a la firma en septiembre de 2024. El artículo 10 del Convenio Marco, sobre igualdad y no discriminación, presenta la idea de un enfoque basado en el ciclo de vida para detectar y corregir sistemáticamente los sesgos; los sistemas de IA deben revisarse para detectar posibles discriminaciones a lo largo del período de uso con arreglo a lo dispuesto en la legislación aplicable<sup>45</sup>.

1. Cada Parte adoptará o mantendrá medidas con vistas a garantizar que las actividades durante el ciclo de vida de los sistemas de inteligencia artificial respeten la igualdad, incluida la igualdad de género, y la prohibición de la discriminación, con arreglo a lo dispuesto en el Derecho internacional y nacional aplicable.

2. Cada Parte se compromete a adoptar o mantener medidas destinadas a superar las desigualdades para lograr resultados justos y equitativos, en consonancia con sus obligaciones nacionales e internacionales aplicables en materia de derechos humanos, en relación con las actividades dentro del ciclo de vida de los sistemas de inteligencia artificial.

<sup>41</sup> Véase <https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment>.

<sup>42</sup> Véase <https://artificialintelligenceact.eu/es/high-level-summary/>.

<sup>43</sup> Véase [https://curia.europa.eu/jcms/upload/docs/application/pdf/2023-11/cjeu\\_ai\\_strategy.pdf](https://curia.europa.eu/jcms/upload/docs/application/pdf/2023-11/cjeu_ai_strategy.pdf).

<sup>44</sup> Véase [https://www.europarl.europa.eu/RegData/etudes/IDAN/2024/754450/EXPO\\_IDA\(2024\)754450\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/IDAN/2024/754450/EXPO_IDA(2024)754450_EN.pdf); véase también Ahmet Bilal Aytakin, “Algorithmic bias in the context of European Union anti-discrimination directives”, ponencia presentada en el European Workshop on Algorithmic Fairness, 7 a 9 de junio de 2023, Winterthur (Suiza); y [https://fra.europa.eu/sites/default/files/fra\\_uploads/fra-2022-bias-in-algorithms\\_en.pdf](https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf).

<sup>45</sup> Véase <https://rm.coe.int/1680afae3c>.

23. También ha habido otras iniciativas regionales para promover la inclusión digital, como la resolución 580 (LXXVIII) 2024 sobre los cortes de Internet y las elecciones en África, aprobada por la Comisión Africana de Derechos Humanos y de los Pueblos<sup>46</sup>. La Corte Interamericana de Derechos Humanos celebró un diálogo con jueces en noviembre de 2023, durante el cual debatieron sobre el *Kit de herramientas global sobre IA y el estado de derecho para el poder judicial* de la UNESCO. En el Kit de herramientas se describen los problemas de la exclusión:

El problema es que [...] los datos pueden estar impregnados de sesgos [...]. Por ejemplo, los ensayos clínicos a menudo excluyen a las mujeres y las personas de color, lo que lleva a una representación inadecuada de los datos. Esto podría tener graves consecuencias si se utilizan algoritmos entrenados con dichos datos para analizar imágenes de la piel o priorizar la atención a los pacientes. Como resultado, es crucial garantizar que los algoritmos de IA estén capacitados utilizando datos representativos para evitar tales sesgos y garantizar resultados equitativos para todos<sup>47</sup>.

24. En el Kit de herramientas también se aborda la cuestión de la representatividad de la IA:

Las brechas digitales en muchos países del hemisferio sur han llevado a la “invisibilidad de los datos”, que probablemente afectará a grupos históricamente marginados como mujeres, castas, comunidades tribales, minorías religiosas y lingüísticas y mano de obra migrante. La utilidad y validez de los algoritmos de IA desarrollados a partir de datos fácilmente disponibles pueden verse limitadas por sesgos perpetuados por la invisibilidad de los datos. Esto subraya los requisitos de transparencia y responsabilidad algorítmica.

En el Kit de herramientas se pone de relieve el problema de la discriminación por proxy, como el uso de información sobre códigos postales, nivel educativo o ingresos por parte de los bancos al evaluar las solicitudes de préstamos que pueden asociarse a minorías étnicas o raciales, lo cual es un método con un riesgo inherente de perpetuar políticas y prácticas discriminatorias interseccionales. Se sugiere que la IA puede ayudar a detectar patrones de sesgo contra personas y grupos marginados por parte de actores públicos y privados, incluido el poder judicial.

25. En el Diálogo Interamericano con los jueces también se debatió la herramienta de la UNESCO para la evaluación del impacto ético de la IA, en la que se pide que se tenga en cuenta “la diversidad en la composición del equipo del proyecto de IA, especialmente en cuanto a género, edad, raza, color, ascendencia, idioma, religión, origen nacional, origen étnico, origen social, condición económica o social, discapacidad y orientación sexual, pero sin limitarse a esos aspectos, y en particular el modo en que refleja la complejidad y diversidad de la población de usuarios prevista, y la manera en que podría introducir sesgos”<sup>48</sup>.

## A. Solidaridad para la alfabetización digital contra la desinformación

26. Aumentan las solicitudes de reglamentación en las “campañas de desinformación electoral impulsadas por extremistas”<sup>49</sup>. La mejora de la calidad de los contenidos generados por IA hace que sea difícil detectar la desinformación y los

<sup>46</sup> Véase <https://achpr.au.int/en/adopted-resolutions/580-internet-shutdowns-elections-africa-achpres580-lxxvii>.

<sup>47</sup> Véase <https://nataliazuazo.com/2023/11/20/ai-and-the-rule-of-law-at-the-inter-american-court-of-human-rights/>. Véase también [https://unesdoc.unesco.org/ark:/48223/pf0000387331\\_spa](https://unesdoc.unesco.org/ark:/48223/pf0000387331_spa).

<sup>48</sup> Véase <https://unesdoc.unesco.org/ark:/48223/pf0000386276>.

<sup>49</sup> Contribución de Jake Okechukwu Effoduh.

contenidos falsos, y cada vez se producen más cortes de Internet por parte de los Estados, lo que tiene un impacto desproporcionado en el acceso a la información. Se necesita una estrategia de solidaridad internacional que incluya una cooperación legislativa, técnica y educativa para aumentar la alfabetización digital<sup>50</sup>. Se ha hecho referencia a la solidaridad receptiva como medio para gestionar emergencias relacionadas con el uso malintencionado de la IA, como “cuando durante una jornada electoral en un país con antecedentes de genocidio circulan miles de videos con contenidos ultrafalsos en los que se muestra violencia étnica”<sup>51</sup>.

27. La Unión Europea aprobó un Reglamento de Servicios Digitales en 2022, que entró en vigor en 2024, en el que se pedía a los motores de búsqueda y las plataformas en línea de gran tamaño que utilizaran códigos de conducta y que aplicaran medidas de respuesta a la desinformación (incluidas las noticias falsas y la publicidad política falsa) cuya aplicación sería absolutamente necesaria y específica<sup>52</sup>. La Ley apoya el Código de Buenas Prácticas en materia de Desinformación reforzado de 2022 (aplicable a las entidades autorreguladas), que incluye orientaciones para desmonetizar la difusión de desinformación, garantizar la transparencia en la publicidad política, mejorar la cooperación con los verificadores de datos y facilitar el acceso de los investigadores a los datos<sup>53</sup>. No obstante, cabe señalar que “aunque puede que estos avances tengan un efecto unificador, el enfoque continental de la desinformación está actualmente fragmentado”<sup>54</sup>. Además, preocupa que pueda haber conflictos entre los códigos de conducta y lo dispuesto en las decisiones del Tribunal Europeo de Derechos Humanos<sup>55</sup>.

28. La Ley de Inteligencia Artificial de la Unión Europea también contiene normas para frenar la influencia y el impacto de los contenidos ultrafalsos (*deepfakes*). Según el artículo 50 4), el implantador, es decir, la persona que utiliza un sistema de IA, debe revelar por lo general que el contenido ha sido generado por IA. Sin embargo, la propia ley contiene excepciones a esta obligación de revelar esa información, lo que significa que su eficacia debe ser objeto de examen.

29. La Agencia de Desarrollo de la Unión Africana pide protección jurídica y marcos reguladores para luchar contra la discriminación algorítmica<sup>56</sup>. El Consejo Ejecutivo de la Unión Africana apoya la creación de una estrategia continental basada en el marco conceptual sobre la IA adoptado por el Consejo Ejecutivo<sup>57</sup>. En la región, existe una tendencia hacia la regulación de los “contenidos tecnológicos”, que implica

<sup>50</sup> *Ibid.*

<sup>51</sup> Miguel Luengo-Oroz, “Solidarity should be a core ethical principle of AI”, *Nature Machine Intelligence*, vol. 1 (noviembre de 2019). Véase también Patrik Hummel y Matthias Braun, “Just data? Solidarity and justice in data-driven medicine”, *Life Sciences, Society and Policy*, vol. 16, núm. 8 (2020).

<sup>52</sup> Reglamento (UE) 2022/2065 del Parlamento Europeo y del Consejo de 19 de octubre de 2022 relativo a un mercado único de servicios digitales y por el que se modifica la Directiva 2000/31/CE (Reglamento de Servicios Digitales).

<sup>53</sup> Véase <https://digital-strategy.ec.europa.eu/es/library/2022-strengthened-code-practice-disinformation>.

<sup>54</sup> Contribución de Dorothy Andersen, Investigadora en materia de Derechos Humanos del Programa de Derecho Internacional y Comparado de la Facultad de Derecho de la Universidad George Washington, Washington, D.C.

<sup>55</sup> Paolo Cavaliere, “The Truth in Fake News: How Disinformation Laws Are Reframing the Concepts of Truth and Accuracy on Digital Platforms”, *European Convention on Human Rights Law Review*, vol. 3, núm. 4 (noviembre de 2022). Véase también Ethan Shattock, “Fake News in Strasbourg: Electoral Disinformation and Freedom of Expression in the European Court of Human Rights”, *European Journal of Law and Technology*, vol. 13, núm. 1 (2022). Véase también <https://edmo.eu/wp-content/uploads/2022/01/Case-law-for-policy-making-Report-2022.pdf>.

<sup>56</sup> Véase <https://www.nepad.org/publication/ai-and-future-of-work-africa-white-paper>.

<sup>57</sup> Véase <https://au.int/en/newsevents/20240419/multistakeholder-consultative-sessions-development-continental-strategy>.

censurar o bloquear el acceso a los contenidos para evitar las noticias falsas o el discurso de odio, y hacia la regulación de los “contenidos jurídicos”, que implica debatir, redactar y aprobar proyectos de ley para regular las noticias falsas y el discurso de odio. Existe el riesgo de que ambas medidas se apliquen de forma demasiado amplia o se dirijan contra la oposición política y debiliten de esa forma la independencia de las elecciones.

30. La Arabia Saudita informó a la Experta Independiente de la creación de la Autoridad Saudita de Datos e Inteligencia Artificial como un avance significativo para el gran modelo lingüístico árabe (conocido como ALLaM). La iniciativa pretende aprovechar las tecnologías de IA y las aplicaciones digitales para favorecer la diversidad cultural y beneficiar a toda la humanidad. Uno de los proyectos, SauTech, se centra en la localización de la tecnología de IA y la preservación de la cultura local, y es capaz de reconocer y transformar la voz en texto en un amplio abanico de dialectos árabes para encontrar soluciones adaptadas a los contextos locales<sup>58</sup>. La Autoridad Saudita de Datos e Inteligencia Artificial sensibiliza sobre la IA mediante programas de capacitación. Del mismo modo, Colombia informó sobre la capacitación en IA a diferentes niveles<sup>59</sup>.

## B. Solidaridad de género en la inteligencia artificial

31. También se presta cada vez más atención al potencial de un enfoque de solidaridad internacional para luchar contra la discriminación de género en relación con la IA<sup>60</sup>. Según el *Artificial Intelligence Index Report 2024*, el número de varones con estudios universitarios en el campo de las tecnologías de la información en los países europeos es superior al de mujeres aunque se sigue avanzando lentamente en la reducción de la brecha de género en ese ámbito<sup>61</sup>. El Comité para la Eliminación de la Discriminación contra la Mujer formuló observaciones finales respecto de varios informes de Estados partes en las que expresaba su preocupación por la infrarrepresentación de las niñas y las mujeres en el campo de la inteligencia artificial, y la ausencia de medidas concretas para evitar la integración de estereotipos de género y la violencia y discriminación algorítmica por razón de género en la codificación de la IA, y pidió que se establecieran salvaguardias efectivas<sup>62</sup>.

32. En su contribución, la Arabia Saudita informó de que las mujeres constituían más del 50 % de la plantilla técnica de la Autoridad Saudita de Datos e Inteligencia Artificial.<sup>63</sup> El Centro Internacional de Investigación y Ética de la Inteligencia Artificial ubicado en el país puso en marcha la iniciativa “Elevate Program” en colaboración con Google Cloud, con el objetivo de formar a 25.000 mujeres en el ámbito de los datos y la IA. El Instituto de Democracia y Derechos Humanos del Perú informó de que la Presidenta de ese país había apoyado un programa de capacitación de proyectos de negocio desarrollados con IA. Los ciudadanos inscritos desde mayo

<sup>58</sup> Contribución de la Arabia Saudita.

<sup>59</sup> Contribución de Colombia.

<sup>60</sup> Keratso Georgiadou, “Solidarity, Gender, Dialogue in the AI era”, en *Kritische Pädagogik und Bildungsforschung: Anschlüsse an Paulo Freire*, Wassilios Baros, Rita Braches-Chyrek, Solvejg Jobst y Joachim Schroeder, eds. (Wiesbaden, Alemania, Springer Fachmedien, 2024), págs. 477–489.

<sup>61</sup> *Artificial Intelligence Index Report 2024*.

<sup>62</sup> Véase [CEDAW/C/ITA/CO/8](#), [CEDAW/C/DEU/CO/9](#), párrs. 27 y 28; [CEDAW/C/ESP/CO/9](#), párrs. 21 y 23, 31 y 33; [CEDAW/C/TJK/CO/7](#), párrs. 47, 48 y 55. Véanse también las observaciones finales que figuran en [CEDAW/C/TUR/CO/8](#), párr. 46; [CEDAW/C/CRI/CO/8](#), párrs. 27, 28 y 37.

<sup>63</sup> Contribución de la Arabia Saudita.

de 2024 tuvieron la oportunidad de solicitar becas de estudio y el 50 % de esas becas se concedieron a mujeres en situación de vulnerabilidad<sup>64</sup>.

33. El Centro de Información sobre Empresas y Derechos Humanos describe los retos que sigue planteando el hecho de que la IA actúe en contra de la solidaridad con las mujeres<sup>65</sup>:

Las mujeres y las personas de género no conforme siguen sintiendo un efecto desproporcionado de las tecnologías que no están debidamente diseñadas, desarrolladas y desplegadas, como lo demuestra el hecho de que las últimas herramientas tecnológicas, incluida la IA generativa, se lanzan al mercado con defectos y sesgos que, entre otras cosas, amplifican el sexismo y los estereotipos de género, facilitan el control social de las mujeres, restan poder a las pacientes, discriminan a las solicitantes de empleo, facilitan la persecución de las defensoras de los derechos humanos y perjudican a las mujeres en su acceso a las prestaciones sociales.

34. Las mujeres están cada vez más expuestas a través de la IA al discurso de odio, el ciberacoso, los contenidos sexuales no consentidos, la pornovenganza y el hostigamiento, que afectan negativamente a su capacidad para lograr sus aspiraciones personales y profesionales. Las organizaciones de la sociedad civil se refirieron a una iniciativa innovadora para poner fin a la violencia de género que se celebró en la Universidad Internacional de Kampala en 2023 y que incluyó capacitación sobre seguridad digital<sup>66</sup>. También señalaron que el uso de los medios sociales era esencial para que la solidaridad internacional pudiera exigir el enjuiciamiento del feminicidio y la reforma legislativa en Kenya, la declaración del feminicidio y la violencia contra las mujeres como una emergencia en ese país, y la creación de una comisión encargada específicamente de tratar estos delitos y romper el ciclo de impunidad<sup>67</sup>. Asimismo, pidieron que las empresas tecnológicas, las organizaciones no gubernamentales (ONG) y los gobiernos reforzaran su colaboración en el ámbito internacional con el fin de desarrollar herramientas eficaces y éticas de moderación de contenidos basadas en IA para detectar y señalar la violencia contra las mujeres en línea, que puedan adaptarse a diferentes idiomas y contextos culturales.

### C. Solidaridad en el diseño de la inteligencia artificial

35. En la segunda revisión del Pacto Digital Global se indica que este se aplicará con un espíritu de solidaridad mundial y se recalca específicamente la inclusión al centrarse en la necesidad de que las empresas de tecnología digital y los desarrolladores colaboren con los usuarios de todos los orígenes y capacidades para incorporar sus perspectivas y necesidades en el ciclo de vida de las tecnologías digitales, lograr una distribución justa de los beneficios digitales y apoyar la accesibilidad digital y la diversidad lingüística y cultural en el espacio digital<sup>68</sup>. Además, la comunidad internacional pide que al desarrollar y aplicar estrategias nacionales y locales de conectividad digital se tengan en cuenta las necesidades de las personas en situación de vulnerabilidad y las que viven en zonas desatendidas, rurales y remotas, y subraya la necesidad de orientar y adaptar la creación de capacidad a las mujeres y las niñas, los niños y los jóvenes, así como a las personas de edad, las personas con discapacidad y las personas en situación de vulnerabilidad,

<sup>64</sup> Contribución del Instituto de Democracia y Derechos Humanos, Perú.

<sup>65</sup> Véase [https://media.business-humanrights.org/media/documents/BHRRRC\\_Submission\\_\\_Gender\\_and\\_Role\\_of\\_Business\\_2023\\_j2HkLgP.pdf](https://media.business-humanrights.org/media/documents/BHRRRC_Submission__Gender_and_Role_of_Business_2023_j2HkLgP.pdf).

<sup>66</sup> Contribución del consorcio Our Voices, Our Futures.

<sup>67</sup> *Ibid.*

<sup>68</sup> Disponible en <https://www.un.org/techenvoy/global-digital-compact>.

y de garantizar su participación significativa en el diseño y la aplicación de los programas. La comunidad internacional también busca desarrollar y llevar a cabo encuestas nacionales de inclusión digital con un desglose sistemático de datos por edad, discapacidad y género, a fin de detectar las lagunas de aprendizaje y establecer las prioridades en contextos específicos.

36. El Pacto Digital Global apoya la aprobación de legislación nacional en materia de IA y de mecanismos eficaces de supervisión y reparación. También pide a las empresas tecnológicas y a los desarrolladores de IA que reconozcan la obligación de respetar los derechos humanos, apliquen la diligencia debida y realicen evaluaciones de impacto. Hace un llamamiento a las empresas tecnológicas y los desarrolladores de IA para que desarrollen conjuntamente marcos de rendición de cuentas del sector (en consulta con los gobiernos y otras partes interesadas) que aumenten la transparencia en torno a sus sistemas y procesos, definan responsabilidades y se comprometan a respetar las normas e informes públicos auditables. Pide a las plataformas de medios sociales que establezcan mecanismos de denuncia seguros y accesibles para que los usuarios y sus defensores puedan denunciar las violaciones de las políticas (incluidos mecanismos especiales de denuncia adaptados a los niños y las personas con discapacidad). Aboga por que las empresas de tecnología de IA y las plataformas de medios sociales mejoren la transparencia y la rendición de cuentas de sus sistemas, incluidas las condiciones de servicio, la moderación de contenidos y los algoritmos y el tratamiento de los datos personales de los usuarios en los idiomas locales, para empoderar a los usuarios a tomar decisiones informadas y dar o retirar su consentimiento informado.

37. La UNESCO hace hincapié en la inclusión a la hora de diseñar mecanismos receptivos de los interesados en IA como elemento de transparencia<sup>69</sup>:

- ¿El gobierno promueve la participación de otros interesados en la gobernanza nacional de la IA?
- ¿Hay asociaciones activas de profesionales de la IA, consumidores y otras comunidades de interesados?
- ¿El gobierno involucra activamente a otros grupos de interesados en el desarrollo de políticas para la gobernanza global de la IA?

También se pide que se apliquen indicadores que garanticen la inclusión y midan la participación de los diversos grupos interesados (mujeres y personas de género diverso, jóvenes y grupos marginados) en la elaboración de políticas de IA y en los mecanismos, procesos y foros internacionales y regionales de IA.

38. Se ha sugerido que: “las tecnologías de IA dentro del sistema sociotécnico pueden servir entonces para facilitar o dar forma a las interacciones sociales y, por lo tanto, pueden socavar o promover la solidaridad”<sup>70</sup>. Es esencial incluir a los interesados en la toma de decisiones sobre el diseño de la IA, teniendo en cuenta que su relevancia puede cambiar con el tiempo y que, por lo tanto, pueden surgir problemas de solidaridad imprevistos. Los derechos humanos son la medida de la solidaridad en relación con la IA<sup>71</sup>. Cabe considerar que el uso de la IA puede llevar a situaciones de violación del derecho a la privacidad debido al reconocimiento facial, o de violación de la no discriminación de las mujeres o las minorías en las contrataciones, o denegación de la autorrealización de objetivos vitales (o un proyecto de vida) como la denegación de la

<sup>69</sup> UNESCO, *El aporte de la inteligencia artificial y las TIC avanzadas a las sociedades del conocimiento*.

<sup>70</sup> Catharina Rudschies, “Exploring the Concept of Solidarity in the Context of AI: An Ethics in Design Approach Digital Society”, vol. 2, núm. 1 (2023).

<sup>71</sup> *Ibid.*, pág. 12.

solicitud de un préstamo para vivienda o educación. Las evaluaciones del impacto en los derechos humanos deben considerarse elementos centrales del diseño de la solidaridad en materia de IA. Además, en el proceso de diseño se deben detectar y tener en cuenta los intereses y necesidades de los usuarios. En el diseño participativo se invitaría a los usuarios a debatir las opciones de diseño y a participar en el proceso de adopción de decisiones, que debería incluir una toma de decisiones colectiva que permita la solidaridad dentro del grupo. Sin embargo, si el foco se pone únicamente en los usuarios no se tiene en cuenta a los interesados indirectos, es decir aquellas personas que se ven afectadas de manera involuntaria. Una perspectiva solidaria de la IA exigiría ir más allá de la atención a los usuarios para centrarse en todos los interesados afectados dentro del sistema sociotécnico. Implicaría no solo la evaluación de los daños y beneficios en general, sino también una evaluación de cómo estos se distribuyen entre la sociedad y los grupos de interesados. La cuestión fundamental es cómo hacer que las empresas tecnológicas acepten la obligación de la solidaridad y el reparto de riesgos, además del reparto de beneficios. Un enfoque solidario de la IA llevaría a cabo una evaluación del riesgo para los derechos humanos, detectaría el derecho de las personas a autorrealizarse y regularía la distribución de la responsabilidad por los daños, además de la distribución de los beneficios del procesamiento de datos y la toma de decisiones<sup>72</sup>. Por lo tanto, también será necesaria la solidaridad política para apoyar la toma de decisiones colectiva, así como un marco de gobernanza solidario para la IA, para hacer frente a las consecuencias de la IA.

39. Para aplicar un enfoque de solidaridad internacional en el ámbito de la IA puede que sea necesario establecer reglamentos en distintas esferas, como la sanidad o la energía. El diálogo en el diseño y desarrollo de IA debe promover la colaboración entre usuarios, desarrolladores de tecnología de IA y otras partes interesadas para encontrar soluciones de IA. Es necesario crear vías de colaboración que favorezcan el diálogo con los desarrolladores de casos de uso de IA, desarrolladores de soluciones de IA u otros desarrolladores de IA/aprendizaje automático, para realizar constantes actualizaciones que tengan en cuenta los riesgos y beneficios cambiantes de los usos de soluciones de IA. También es necesario establecer planes de comunicación e interacción adaptados a cada contexto y basados en riesgos para que los clientes puedan comprender fácilmente las explicaciones sobre cómo se ha desarrollado la solución de IA, su rendimiento y mantenimiento, y el modo en que se ajusta a los requisitos reglamentarios y las mejores prácticas más recientes.

40. El reconocimiento y el respeto deben considerarse elementos de la solidaridad de datos que dan “un lugar conceptual a las experiencias concretas de marginación y a las experiencias de injusticia en contextos de uso intensivo de datos”<sup>73</sup>. La solidaridad como principio de la IA debe tener un efecto de equidad sostenida en el reparto de los beneficios y las cargas y evitar la desigualdad<sup>74</sup>. Existe una propuesta académica de un marco general de solidaridad en el ámbito de la IA basado en incentivos para evitar la desigualdad, como el pago de regalías cada vez que se utilice un sistema de IA entrenado con datos de alguna persona, recompensas a los médicos que utilicen un modelo de IA para realizar diagnósticos o remuneración a las personas que produzcan textos para un generador de texto automático de IA cada vez que el robot escriba un artículo público<sup>75</sup>. Otros abogan por un enfoque solidario que mida el impacto externo en la sociedad y promueven el uso de salvaguardias para el control individual de los datos que se procesan<sup>76</sup>. Otra perspectiva es la de un marco de

---

<sup>72</sup> *Ibid.*, pág. 13.

<sup>73</sup> Hummel y Braun, “Just data? Solidarity and justice in data-driven medicine”.

<sup>74</sup> Luengo-Oroz, “Solidarity should be a core ethical principle of AI”.

<sup>75</sup> *Ibid.*

<sup>76</sup> Hummel y Braun, “Just data? Solidarity and justice in data-driven medicine”.

solidaridad de datos que haga visibles los procesos de datos para el bien público<sup>77</sup>. Se sugiere que esto puede hacer que las partes interesadas de datos públicos y corporativos compartan tanto los riesgos como los beneficios del acceso, la producción y la puesta en común de datos. La solidaridad de datos se presenta como la promoción de la visibilidad de los procesos de los conjuntos de datos para detectar y corregir la discriminación basada en la marginación<sup>78</sup>. Se hace un llamamiento a la acción colectiva, utilizando la solidaridad como principio de gobernanza de datos para la creación de conjuntos de datos públicos con el fin de empezar a generar confianza y favorecer la rendición de cuentas. Existe la opinión de que la comunidad internacional se enfrenta a un punto de inflexión en el que la solidaridad de datos debe promover la creación de infraestructuras y debe configurarse de acuerdo con valores democráticos, como la inclusión<sup>79</sup>. Además, esto puede requerir la creación de una entidad independiente de gobernanza global compuesta por representantes del sector, los Estados, la sociedad civil, las organizaciones internacionales y el mundo académico para promover normas sobre IA basadas en los derechos humanos<sup>80</sup>. Cabe considerar la “Constitución de Claude” de Anthropic, que incluye principios que fomentan la inclusión de perspectivas no occidentales, además de los derechos humanos<sup>81</sup>.

41. Algunas empresas tecnológicas hicieron aportaciones sobre la implementación de la solidaridad internacional en el ámbito de la IA desde el diseño:

- a) Las sesiones de capacitación se diseñan de manera accesible, culturalmente sensible e inclusiva;
- b) Los algoritmos se entrenan en diversos conjuntos de datos para representar a las personas a las que sirven;
- c) Auditorías periódicas, pruebas en diferentes grupos demográficos y aplicación de métodos de aprendizaje automático que tengan en cuenta la equidad;
- d) Directrices éticas para la IA que den prioridad a la equidad, la transparencia y la rendición de cuentas en el desarrollo y despliegue de algoritmos;
- e) Supervisión y ajuste continuos de los sistemas de IA para evitar la discriminación;
- f) Colaboración con grupos de defensa de derechos, ONG y representantes comunitarios de poblaciones en situación de vulnerabilidad para conocer sus preocupaciones y recabar su opinión sobre el impacto de las tecnologías;
- g) Capacitación continua a los equipos sobre la equidad algorítmica y las implicaciones éticas de la IA;
- h) Políticas de moderación de contenidos que prohíben la difusión de información falsa, el discurso de odio y contenidos nocivos;
- i) Aplicación de las leyes mediante herramientas automatizadas y moderadores humanos entrenados para detectar y eliminar la desinformación;

<sup>77</sup> Mercedes Bunz y Photini Vrikki, “From Big to Democratic Data: Why the Rise of AI Needs Data Solidarity”, en *Democratic Frontiers: Algorithms and Society*, Michael Filimowicz, ed. (Londres, Taylor & Francis, 2022).

<sup>78</sup> *Ibid.*

<sup>79</sup> *Ibid.*

<sup>80</sup> Ana Beduschi, “Human rights and the governance of artificial intelligence”, documento de investigación (Academia de Derecho Internacional Humanitario y Derechos Humanos de Ginebra, 2020).

<sup>81</sup> Véase <https://www.anthropic.com/news/claude-constitution>.

- j) Colaboración con organizaciones reconocidas de verificación de datos;
- k) Priorización de la transparencia en sus algoritmos para evitar la amplificación involuntaria de la desinformación;
- l) Diseño de algoritmos para promover fuentes creíbles y reducir la visibilidad de contenidos engañosos;
- m) Campañas informativas y notificaciones en plataformas con el fin de educar a los usuarios para que detecten y denuncien la desinformación;
- n) Colaboración con organismos gubernamentales y no gubernamentales e instituciones académicas para compartir ideas y mejores prácticas para luchar contra la desinformación de manera eficaz.

#### IV. Orientación nacional hacia prácticas positivas

42. España informó de que cuenta con una estrategia nacional de inteligencia artificial y un organismo para su supervisión<sup>82</sup>. Indicó que establecerá un marco ético y normativo que refuerce la protección de los derechos individuales y colectivos, a efectos de garantizar la inclusión y el bienestar social. España tiene previsto elaborar una carta de derechos digitales y poner en marcha un modelo de gobernanza nacional de la ética en la IA a través de un consejo asesor de IA, en colaboración con un consejo consultivo de transformación digital. En el artículo 23 de la ley núm. 15/2022, de 12 de julio de 2022, se dispone que deben tomarse medidas para reducir los sesgos y, al mismo tiempo, contribuir a una mayor transparencia y rendición de cuentas cuando se utilice la IA en la toma de decisiones de las administraciones públicas. Estas medidas incluyen el análisis del diseño y de los datos de entrenamiento y su evaluación para detectar cualquier impacto discriminatorio<sup>83</sup>. Además, el país tiene previsto crear foros de diálogo y promover la sensibilización y la participación nacional e internacional de la IA dirigidos a fomentar el diálogo entre gobierno, ciencia, sector privado y sociedad civil.

43. La Arabia Saudita informó que la Autoridad Saudita de Datos e Inteligencia Artificial cuenta con un sistema para organizar, desarrollar y manejar datos de IA y proporcionar servicios gubernamentales, que busca apoyar de manera confiable la transformación digital y la preservación de datos a través de la inclusión de las partes interesadas en el desarrollo y la implementación de la IA<sup>84</sup>. Promueve siete principios éticos de la IA, entre ellos la equidad; la privacidad y protección; la humanidad; los beneficios sociales y medioambientales; la fiabilidad y seguridad; la transparencia y explicabilidad; y la rendición de cuentas y la responsabilidad<sup>85</sup>. La Autoridad anunció un programa de incentivos para ayudar a las empresas a cumplir voluntariamente la ética de la IA. El proceso comienza por detectar y evaluar todos los riesgos potenciales y la gravedad de su impacto.

44. En su contribución, Malasia brindó información sobre su mecanismo de supervisión propuesto, una unidad de coordinación e implementación de la IA, que actuará como órgano gubernamental en todos los asuntos relacionados con la IA<sup>86</sup>. El

<sup>82</sup> Véase <https://www.lamoncloa.gob.es/presidente/actividades/Documents/2020/ENIA2B.pdf>.

<sup>83</sup> Contribución de España.

<sup>84</sup> Contribución de la Arabia Saudita.

<sup>85</sup> En la contribución de Colombia también se señala el respeto de principios como la transparencia, la privacidad, el control humano y la no discriminación, lo que garantiza que las tecnologías sean inclusivas y justas.

<sup>86</sup> Contribución de Malasia.

país tiene previsto crear un comité de prospectiva que se encargará del escaneo de horizontes, la prospectiva y la promoción de políticas.

45. Alemania utiliza la IA como parte de su programa Match'In para mejorar la integración de los refugiados en los municipios, consultándoles para conocer sus preferencias en cuanto a vivienda, empleo/experiencia profesional, aficiones, ocio, atención sanitaria, situación familiar, comunidad religiosa y otras cuestiones<sup>87</sup>. Ese enfoque promueve el reconocimiento del derecho de los refugiados a la autorrealización, en consonancia con un enfoque de solidaridad internacional. Se ha observado que el programa Match'In es el único que incluye las perspectivas de los refugiados en el programa de diseño<sup>88</sup>.

46. La República Dominicana indicó que cuenta con una estrategia nacional de inteligencia artificial que prioriza el uso de la IA en la administración pública en sectores clave como la justicia, la salud, la educación, el medio ambiente y la seguridad, y como modelos analíticos predictivos para diseñar servicios para la ciudadanía. Su objetivo es disponer de un código que incluya acciones preventivas, garantías procesales y mecanismos de rendición de cuentas para garantizar una aplicación responsable de la IA. Además, se establecerán mecanismos de supervisión para verificar el cumplimiento ético. El hub de talento humano e innovación YoSoyFuturoRD dará prioridad a los sectores vulnerables. El país propuso la creación de un ecosistema regional de IA fuerte y colaborativo que contribuya significativamente al avance tecnológico, el desarrollo económico y social y la cooperación dentro de la región. Afirma que establecerá mecanismos de supervisión, compensación y recurso contra los daños causados por los sistemas de IA, con el fin de salvaguardar los derechos de sus ciudadanos.

47. La municipalidad de Rawson en la Argentina utiliza la IA para promover programas de asistencia social e incentivos, como la Red de Economía Social y Solidaria, con miras a generar circuitos alimentarios sostenibles en el marco de la soberanía alimentaria. También cuenta con centros de atención barrial en la Municipalidad de Rawson (Municerca), que tramitan solicitudes de mejora en obras de espacios públicos y calles, el inicio y la consulta de expedientes, y la recepción de reclamos municipales.

## A. Solidaridad sanitaria e inteligencia artificial

48. La solidaridad sanitaria debe ser la base de la reglamentación sobre el uso de la IA para garantizar el acceso y la equidad, pero resulta difícil alcanzar ese objetivo porque en muchos países la sanidad es privada<sup>89</sup>. Las desigualdades en salud preexistentes por motivos de raza, clase, edad y otros factores podrían reproducirse en los sistemas médicos de IA que no abordan la injusticia estructural. La IA utilizada para prestar asistencia sanitaria debe diseñarse teniendo en cuenta la atención sanitaria de las comunidades y, más concretamente, debe desplegarse en zonas de bajos recursos y facilitar de ese modo la cobertura sanitaria en entornos de bajos recursos. Algunos han propuesto una visión de la solidaridad sanitaria en la IA que situaría a las personas en posición de compartir o retirar datos mediante el control de los flujos de datos, la creación de infraestructuras para la controlabilidad y la orientación de la gobernanza hacia los resultados, con el fin de prevenir y mitigar las

<sup>87</sup> Véase <https://matchin-projekt.de/en/>.

<sup>88</sup> Véase [https://www.rsc.ox.ac.uk/files/files-1/automating-immigration-and-asylum\\_afar\\_9-1-23.pdf](https://www.rsc.ox.ac.uk/files/files-1/automating-immigration-and-asylum_afar_9-1-23.pdf).

<sup>89</sup> Nicolas Terry, "Of Regulating Healthcare AI and Robots", *Yale Journal of Law and Technology*, número especial 133 (2019).

injusticias<sup>90</sup>. El Brasil cuenta con el programa Bolsa Família, que utiliza herramientas digitales para distribuir ayudas a las familias de ingresos bajos. El programa ha reducido la pobreza y fomentado el desarrollo de capital humano al vincular las transferencias en efectivo a condiciones específicas, como la asistencia escolar y los controles médicos. Malasia utiliza la IA para ayudar al diagnóstico médico, desarrollar tratamientos personalizados, proporcionar información a través de chats y ayudar con análisis predictivos en el establecimiento de medidas proactivas<sup>91</sup>.

49. La Arabia Saudita señaló su inversión en la solidaridad sanitaria en el ámbito de la IA. El país cuenta con un centro de excelencia para la inteligencia artificial en el sector de la salud. Los expertos e ingenieros saudíes desarrollaron la herramienta Eyenai, que ha revolucionado el diagnóstico médico en la región. La retinopatía diabética es una enfermedad causada por la diabetes de tipo 1 y 2 y una de las principales causas de ceguera en la Arabia Saudita. El diagnóstico precoz es fundamental para poder evitar complicaciones posteriores, y la herramienta Eyenai contribuye a ello mediante la detección y el diagnóstico precisos. Una aplicación desarrollada por la Autoridad Saudita de Datos e Inteligencia Artificial (Tawakkalna) recibió el premio de Administración Pública de las Naciones Unidas en 2022 por haber ofrecido una respuesta innovadora a la pandemia de la enfermedad por coronavirus (COVID-19).

## B. Solidaridad entre los trabajadores e inteligencia artificial

50. La fragmentación de la mano de obra resultante del uso de la IA dificulta la capacidad de los trabajadores de organizarse de forma solidaria: “La facilidad con la que una empresa de plataforma puede contratar a nuevos trabajadores, a veces incluso de todo el mundo, genera competencia y aislamiento, y dificulta considerablemente la formación de una identidad colectiva. En resumen, el trabajo basado en plataformas fragmenta la identidad colectiva de los trabajadores y sabotea la acción colectiva, especialmente la acción basada en la solidaridad y la confianza”<sup>92</sup>.

51. No obstante, se ha sugerido que la intermediación solidaria puede ayudar a los trabajadores en tiempos de crisis<sup>93</sup>. Una cuestión central es si la IA puede concebirse como un instrumento para fomentar la solidaridad entre los trabajadores<sup>94</sup>. El Banco de Previsión Social del Uruguay estableció procedimientos centralizados y automatizados para los empleadores que promovieron la autogestión mediante servicios en línea, una aplicación móvil y asistencia multicanal. Según se informó, un chatbot respondió al 97 % de las consultas con un índice de satisfacción del 100 %. Entre los resultados cabe destacar una reducción significativa del 24,4 % en la tasa de evasión de las cotizaciones a la seguridad social: el 57 % de los empresarios se registraron en línea y el 42 % efectuaron pagos en línea<sup>95</sup>. La Superintendencia de Riesgos del Trabajo de la Argentina ha lanzado con éxito un chatbot de inteligencia

<sup>90</sup> Hummel y Braun, “Just data? Solidarity and justice in data-driven medicine”.

<sup>91</sup> Contribución de Malasia.

<sup>92</sup> Tammy Katsabian, “The rule of technology - how technology is used to disturb basic labor law protections”, *Lewis and Clark Law Review*, vol. 25, núm. 3 (2021).

<sup>93</sup> Saiph Savage y Mohammad H. Jarrahi, “Solidarity and AI for transitioning to crowd work during COVID-19”, ponencia elaborada para el simposio virtual *The New Future of Work*, agosto de 2020. Véase también Kurt April, “AI-Induced Solidarity Economy: The Need for Stewardship Orientation”.

<sup>94</sup> Frances Flanagan y Michael Walker, “How can unions use artificial intelligence to build power? The use of AI chatbots for labour organising in the US and Australia”, *New Technology, Work and Employment*, vol. 36, núm. 2 (2021).

<sup>95</sup> Véase <https://www.issa.int/sites/default/files/documents/2024-06/2-AI%20in%20SecSoc%202024.pdf>.

artificial llamado Julieta<sup>96</sup>. Por un lado, hay nuevas formas de solidaridad horizontal que son claramente más abiertas y diversas. También existe la tendencia del “discurso distribuido” que eliminó las barreras burocráticas entre los activistas y los responsables sindicales, aumentó la transparencia y los miembros ordinarios reforzaron su poder para reformular y desafiar a la oligarquía sindical<sup>97</sup>. Además, está surgiendo un “pluralismo acelerado”, que se ha definido como “la fragmentación existente en la política de grupo basada en los intereses en favor de una política de grupo más fluida y basada en los problemas, y nuevos espacios seguros para el debate democrático entre las mujeres y otros trabajadores que anteriormente habían sido marginados por las estructuras sindicales tradicionales”<sup>98</sup>. El aumento de la capacidad de las trabajadoras agrícolas para acceder a la IA con capacidad predictiva con objeto de prever las cosechas y las pautas meteorológicas puede considerarse un ejemplo de solidaridad en el ámbito de la IA<sup>99</sup>. Otro ejemplo es el movimiento contra el apartheid dirigido por los trabajadores del sector tecnológico<sup>100</sup>. Por otro lado, cabe destacar la gran labor de la California Nurses Association para exigir salvaguardias en materia de IA que protejan a los pacientes contra un tratamiento deficiente<sup>101</sup>. No obstante, se ha observado que “los empleadores se han contramovilizado en el espacio en línea y se ha observado que la incierta situación jurídica del discurso en los medios sociales ejerce un efecto disuasorio adicional en el discurso en línea de los empleados”<sup>102</sup>.

52. Por lo tanto, el contexto del tipo de actividad sindical, así como las diferencias políticas y organizativas de los movimientos sindicales, afectarán al uso de la IA en relación con la solidaridad internacional en cuestiones relacionadas con la oposición al poder, y refleja la identidad del grupo, la cohesión interna, la cultura, la estrategia, la gobernanza y la comunidad<sup>103</sup>.

## V. Procedimientos de reclamación en materia de diligencia debida

53. La creciente preocupación por la transparencia está dando lugar a reclamaciones para que se revisen las extralimitaciones de las empresas en cuanto al acceso y uso de los datos personales. Por ejemplo, Meta anunció que tenía previsto establecer una configuración predeterminada que permitiría utilizar los contenidos de los usuarios para entrenar modelos de inteligencia artificial. El Consejo de Consumidores de Noruega consideró que el proceso para optar por no participar (*opt-out*) era complejo y presentó una denuncia contra la empresa por infringir el Reglamento General de

<sup>96</sup> *Ibid.*

<sup>97</sup> Véase Andy Hodder, David Houghton, “Union use of social media: a study of the University and College Union on Twitter”, *New Technology, Work and Employment*, vol. 30, núm. 3 (noviembre de 2015).

<sup>98</sup> Véase Anne-Marie Greene y Gill Kirton, “Possibilities for remote participation in trade unions: mobilising women activists”, *Industrial Relations Journal*, vol. 34, núm. 4 (octubre de 2003).

<sup>99</sup> Véase CEDAW/C/OMN/CO/4. En relación con la IA y la predicción meteorológica, véase: <https://deepmind.google/discover/blog/graphcast-ai-model-for-faster-and-more-accurate-global-weather-forecasting/>.

<sup>100</sup> Véase <https://www.notechforapartheid.com/>.

<sup>101</sup> Véase <https://www.nationalnursesunited.org/press/cna-demand-patient-safeguards-against-artificial-intelligence>.

<sup>102</sup> Flanagan y Walker, “How can unions use artificial intelligence to build power? The use of AI chatbots for labour organising in the US and Australia”. Véase también Louise Thornthwaite, “Chilling times: Social media policies, labour law and employment relations”, *Asia Pacific Journal of Human Resources*, vol. 54, núm. 3 (agosto de 2015).

<sup>103</sup> *Ibid.*

Protección de Datos de la Unión Europea<sup>104</sup>. En el principio 15 de los Principios Rectores sobre las Empresas y los Derechos Humanos se establece la obligación de diligencia debida de las empresas, que incluye un llamamiento a la adopción de enfoques preventivos así como un proceso de reparación; en el principio 25 se establece el marco para el acceso a la justicia; y en el principio 29 se insta a las empresas a establecer mecanismos de reparación.

54. Los Principios Globales de las Naciones Unidas para la Integridad de la Información piden la creación de mecanismos de supervisión y el encargo de auditorías externas independientes periódicas en materia de derechos humanos que abarquen, entre otras cosas: las condiciones de servicio, las normas de la comunidad, las políticas de publicidad, la moderación de contenidos, los procesos de reclamación, el acceso de los investigadores a los datos, la evaluación del impacto en la vulnerabilidad y la marginación, la igualdad de género y los derechos de la infancia. Las auditorías deben ser públicas, accesibles y comprensibles para todos los usuarios<sup>105</sup>.

55. En su Recomendación sobre la Ética de la Inteligencia Artificial, la UNESCO indicó que se espera que los Estados establezcan mecanismos de supervisión que incluyan a todas las partes interesadas, y que alienten a esas partes a desarrollar los derechos humanos. El artículo 14 de la Directiva (UE) 2024/1760 del Parlamento Europeo y del Consejo de 13 de junio de 2024 sobre diligencia debida de las empresas en materia de sostenibilidad dispone la obligación de los Estados de velar por que las empresas establezcan un mecanismo de reclamación justo, públicamente disponible, accesible, previsible y transparente para las personas, las organizaciones de la sociedad civil, los defensores de los derechos humanos, los sindicatos y otros, con el fin de abordar las violaciones de los derechos humanos y el impacto ambiental de las empresas, sus filiales u otras entidades de la cadena de actividades<sup>106</sup>. Estas medidas deben respetar las normas de confidencialidad e impedir las represalias contra la persona o entidad que haya presentado la reclamación. La directiva establece un marco que complementa el mecanismo de comunicación previsto en el proyecto de declaración revisado sobre el derecho a la solidaridad internacional (art. 6, párr. 3)). Además, el proyecto de declaración revisado (art. 8, párr. 3)) establece un marco flexible para el posible diseño e intercambio de políticas y prácticas de solidaridad internacional en el ámbito de la IA durante el examen periódico universal. De conformidad con el párrafo 3 del artículo 8 del proyecto de declaración revisado, los Estados tienen el deber de adoptar medidas, dentro de sus respectivas capacidades, para facilitar la protección de los espacios reales y virtuales de comunicación, incluido el acceso a Internet y a las infraestructuras, con el fin de permitir que las personas y los pueblos compartan ideas solidarias. Estas normas proporcionan una base para apoyar mecanismos creativos de supervisión de la solidaridad internacional en el ámbito de la IA que podrían promover una mayor inclusión participativa de la sociedad civil.

## VI. Conclusión

56. La Experta Independiente coincide con la opinión de que es necesario crear un modelo de gobernanza mundial de múltiples partes interesadas para prevenir y reparar la discriminación y otras violaciones de los derechos humanos en los sistemas

<sup>104</sup> Véase <https://www.forbrukerradet.no/side/legal-complaint-against-metas-use-of-personal-content-for-ai-training/>.

<sup>105</sup> Véase [https://www.un.org/sites/un2.un.org/files/principios\\_globales\\_onu\\_integridad\\_informacion.pdf](https://www.un.org/sites/un2.un.org/files/principios_globales_onu_integridad_informacion.pdf).

<sup>106</sup> Véase [https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L\\_202401760](https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L_202401760).

de IA<sup>107</sup>. En cuanto a la gobernanza, el Órgano Asesor de Alto Nivel sobre Inteligencia Artificial recomienda un enfoque interoperable que abarque a los actores públicos y privados y se extienda a todas las jurisdicciones, incluidos los niveles internacional, regional y nacional, como la Ley de Inteligencia Artificial de la Unión Europea<sup>108</sup>. No obstante, un indicador clave de la evolución del sistema es el uso de códigos y acuerdos voluntarios, como el acuerdo contra los contenidos ultrafalsos (*deepfakes*) para luchar contra la injerencia electoral, firmado por empresas tecnológicas, desarrolladores de IA y empresas de seguridad<sup>109</sup>. En la actualidad, la comunidad internacional no se reparte de manera equitativa los beneficios y riesgos de la solidaridad en el ámbito de la IA y parece existir una falta fundamental de confianza entre las organizaciones de la sociedad civil del Sur Global y las empresas tecnológicas y los desarrolladores de IA del Norte Global. Además, la mayoría de los principios éticos de la IA se redactan en el Norte Global y puede que no presten atención a las preocupaciones específicas del Sur Global, incluso cuando este último proporciona los minerales y la energía que son esenciales para el desarrollo de la infraestructura de datos de IA<sup>110</sup>. Existe una concentración de poder por parte de las empresas tecnológicas y los desarrolladores de IA que aumenta el riesgo de que se refuercen las medidas antisolidarias relacionadas con la IA, lo que agravaría la brecha digital entre los países y dentro de ellos, así como entre los distintos sectores de la sociedad. El Foro para la Gobernanza de Internet subraya los problemas relacionados con la escasa participación de los representantes del Sur Global en las consultas mundiales como consecuencia de la falta de financiación, la conectividad precaria y el hecho de que el inglés sea el principal idioma para el diálogo<sup>111</sup>. La solidaridad internacional en el ámbito de la IA, desde el diseño, puede constituir un valor clave para el futuro desarrollo de la IA y para la aplicación de mecanismos de supervisión que defiendan la justicia procesal y la participación inclusiva de todos. Con el fin de apoyar un cambio político hacia la aplicación de una solidaridad internacional sostenible en el ámbito de la IA, la Experta Independiente formula recomendaciones que deben adoptarse de manera urgente en los próximos cinco años.

## VII. Recomendaciones

**57. La Experta Independiente recomienda a los Estados, las empresas y los actores de la sociedad civil:**

**a) Garantizar que las instituciones gubernamentales, las empresas tecnológicas y los grupos de la sociedad civil que aplican la IA para el procesamiento de datos y la toma de decisiones incluyen la participación activa de todas las personas y grupos, independientemente de su raza, color, ascendencia, género, edad, idioma, religión, opiniones políticas, origen nacional o étnico, antecedentes sociales o económicos, discapacidad o cualquier otro factor, de conformidad con la Recomendación sobre la Ética de la Inteligencia Artificial de la UNESCO;**

**b) Aprobar reglamentos nacionales para promover la solidaridad internacional en el ámbito de la IA desde el diseño, aplicando un enfoque no**

<sup>107</sup> Véase [https://www.unwomen.org/en/news-stories/explainer/2024/05/artificial-intelligence-and-gender-equality?gad\\_source=1&gclid=CjwKCAjw4f6zBhBVEiwATEHFVvzWNyAYvJV56epBISUMbQTVZ4hm\\_tCsn\\_VJGGvIzoRMMXfutYebdhoCzqEQAvD\\_BwE](https://www.unwomen.org/en/news-stories/explainer/2024/05/artificial-intelligence-and-gender-equality?gad_source=1&gclid=CjwKCAjw4f6zBhBVEiwATEHFVvzWNyAYvJV56epBISUMbQTVZ4hm_tCsn_VJGGvIzoRMMXfutYebdhoCzqEQAvD_BwE).

<sup>108</sup> Véase [https://www.un.org/sites/un2.un.org/files/un\\_ai\\_advisory\\_body\\_governing\\_ai\\_for\\_humanity\\_interim\\_report.pdf](https://www.un.org/sites/un2.un.org/files/un_ai_advisory_body_governing_ai_for_humanity_interim_report.pdf). Véase también <https://oecd.ai/en/accountability>.

<sup>109</sup> Véase <https://www.techradar.com/pro/top-tech-companies-ai-developers-and-security-firms-sign-anti-deepfake-agreement-to-combat-election-interference>.

<sup>110</sup> Contribución del Foro para la Gobernanza de Internet.

<sup>111</sup> *Ibid.*

discriminatorio que garantice la inclusión de la sociedad civil y otras partes interesadas pertinentes en la planificación, la selección, el diseño y la aplicación de la IA en las instituciones gubernamentales y las empresas tecnológicas que participan en el procesamiento de datos y la toma de decisiones;

c) Los gobiernos y las empresas deben realizar continuas evaluaciones del impacto en los derechos humanos con contribuciones directas de la sociedad civil, antes de recopilar, procesar o desplegar datos, y ofrecer total transparencia, en particular en el entrenamiento de la IA;

d) Garantizar que las instituciones gubernamentales y las empresas tecnológicas que implementan la IA para el procesamiento de datos y la toma de decisiones realizan auditorías de transparencia y evaluaciones del impacto en los derechos humanos independientes, externas y sistemáticas (para identificar sesgos en el entrenamiento, los algoritmos y la toma de decisiones) con la solidaridad internacional como valor y objetivo durante todo el ciclo de vida de la IA<sup>112</sup>;

e) Garantizar la inclusión de mecanismos de denuncia y apelación independientes, transparentes, accesibles y eficaces en la legislación nacional y regional, que permitan la rendición de cuentas en casos de discriminación o exclusión mediante la IA o por la IA utilizada en instituciones gubernamentales, empresas tecnológicas o grupos de la sociedad civil que participan en el procesamiento de datos o la toma de decisiones;

f) Reconocer que en los casos de discriminación u otras violaciones de los derechos humanos debido al uso de la IA para el procesamiento de datos o la toma de decisiones, la carga de la prueba en relación con el uso de la tecnología de IA debe recaer en la institución gubernamental pertinente o en la empresa tecnológica;

g) Aprobar normas claras de IA en las evaluaciones del impacto en relación con el procesamiento de datos y la toma de decisiones institucionales y mecanismos para la protección de los derechos cuando el desarrollo implique el uso de datos sensibles (por ejemplo, datos de salud) o cantidades masivas de datos (por ejemplo, bases de datos nacionales);

h) Garantizar que las instituciones estatales y las empresas tecnológicas o los grupos de la sociedad civil informan, de manera fácilmente comprensible, a todas las personas y grupos en situación de vulnerabilidad de que se utilizará la IA en el procesamiento de datos o la toma de decisiones, y de que se solicitará previamente el consentimiento sin coacción directa o indirecta;

i) Garantizar mecanismos de acceso a la justicia que sean independientes para abordar la discriminación relacionada con la IA u otras violaciones de derechos humanos en ese ámbito;

j) Los gobiernos y las empresas tecnológicas deben proporcionar información jurídica de manera explicable a los grupos en situación de vulnerabilidad en relación con sus derechos y mecanismos de reparación cuando sean objeto de violaciones relacionadas con el uso de la IA;

k) Invertir en educación para la alfabetización en el uso de datos destinada al público en general para hacer frente a los contenidos ultrafalsos (*deepfakes*), la desinformación y el discurso de odio mediante IA, a fin de crear sociedades más resilientes;

<sup>112</sup> Véanse <https://digital-strategy.ec.europa.eu/es/library/ethics-guidelines-trustworthy-ai> y <https://ieeexplore.ieee.org/document/8058187>.

l) Reconocer el derecho de todas las personas a la autorrealización como criterio clave de un enfoque desde el diseño de la solidaridad internacional en el ámbito de la IA, con procesos que aborden, entre otras cosas, el acceso a la educación, la vivienda, el empleo y la atención sanitaria;

m) Respetar la autonomía de los titulares de datos en la recopilación y el procesamiento de datos;

n) Se espera que los gobiernos y las empresas tecnológicas informen a los titulares de datos de sus derechos en materia de protección de datos y hagan todo lo necesario para garantizar que todas las personas tengan acceso a derechos en materia de protección de datos, incluidos los derechos a la eliminación de datos, el acceso y la no participación (*opt-out*);

o) Los gobiernos y las empresas tecnológicas que procesen datos con otros usos distintos al propósito original (uso secundario de datos) deben solicitar el consentimiento libre, previo e informado de las personas, independientemente de cómo se hayan obtenido los datos; las autoridades de seguridad no deben perseguir a los activistas de la solidaridad internacional con herramientas de IA destinadas a la lucha contra las actividades extremistas o terroristas, en cumplimiento de las normas jurídicas relativas a la necesidad, la proporcionalidad, la legalidad y la no discriminación;

p) Establecer un marco jurídico y un procedimiento sistemático para determinar el alcance del interés público y categorizar la urgencia y la necesidad del procesamiento de datos; las personas deben conservar el derecho a no participar (*opt-out*) para proteger su privacidad;

q) Los Estados y las empresas deben prohibir la creación de ultrafalsificaciones de imágenes íntimas sin consentimiento, sintéticas o manipuladas, de acuerdo con una solidaridad internacional de género en el ámbito de la IA.

58. La Experta Independiente recomienda a la sociedad civil:

a) Seguir abogando por la inclusión en la creación de herramientas para la moderación de contenidos basadas en IA éticas con miras a prevenir, detectar y eliminar la violencia y la discriminación en línea contra las mujeres y las personas objeto de discriminación intersectorial;

b) Seguir proporcionando ejemplos de discriminación algorítmica a las Naciones Unidas, los gobiernos nacionales y las empresas tecnológicas o desarrolladores de IA con el fin de ofrecer recomendaciones para mejorar las prácticas;

c) Los sindicatos de trabajadores deben promover la inclusión y la consulta de los trabajadores en el codiseño de la IA, el entrenamiento y la evaluación de riesgos, además del acceso a la alfabetización digital.

59. La Experta Independiente recomienda a las empresas:

a) Establecer mecanismos concretos para amplificar las voces de los grupos tradicionalmente marginados con el fin de promover un entorno en línea seguro e inclusivo que reconozca como principio el respeto mutuo por la diversidad, incluidas consultas en distintos idiomas y la creación de becas para facilitar la participación de las partes interesadas del Sur Global;

b) Las empresas tecnológicas y los desarrolladores de IA deben adoptar medidas de solidaridad internacional en el ámbito de la IA para respetar los límites planetarios desde el diseño; realizar evaluaciones de riesgos

medioambientales independientes, transparentes y continuas; promover algoritmos de eficiencia energética; apoyar el desarrollo digital sostenible utilizando energías renovables; financiar la economía solidaria en el ámbito de la IA para la creación de datos que apoyen el conocimiento medioambiental y del cambio climático de zonas comunes accesibles al público (como los fondos marinos y el océano); gestionar los desechos de equipo eléctricos y electrónicos; y aplicar principios de economía circular que incluyan a las comunidades locales;

c) Los financiadores y desarrolladores de IA deben apoyar la solidaridad internacional en el ámbito de la IA mediante la realización de evaluaciones de los derechos humanos preventivas y sistemáticas para detectar el riesgo de sesgos algorítmicos que violen los principios de no discriminación y la igualdad, en todo el ciclo de vida de los datos, y exigir la inclusión de las partes interesadas directas e indirectas de la sociedad civil (incluidas las comunidades estructuralmente silenciadas) en un mecanismo de supervisión independiente para descubrir las agendas contrarias a la igualdad de género y la democracia y hacerles frente;

d) Las empresas tecnológicas deben crear mecanismos de solidaridad preventivos y reactivos para hacer frente a las violaciones de los derechos humanos relacionadas con el uso de la IA, como las campañas de información errónea y desinformación que desembocan en violencia social o en el acoso, la vigilancia, la discriminación o la censura desproporcionada de comunidades estructuralmente silenciadas<sup>113</sup>;

e) Las empresas tecnológicas deben proporcionar equipos de respuesta rápida dotados de recursos suficientes y con poder de decisión para responder a tiempo a las reclamaciones y ofrecer soluciones.

60. La Experta Independiente recomienda a las Naciones Unidas:

a) Crear una plataforma de solidaridad digital a través de la cual las organizaciones de la sociedad civil (del Norte Global y el Sur Global) puedan intercambiar ideas de solidaridad internacional y, además, debatir políticas y prácticas de inclusión y no discriminación, retos e innovaciones en el ámbito de la IA con representantes de empresas tecnológicas, desarrolladores de IA y funcionarios del Estado, en línea con las recomendaciones esbozadas en el informe de la Experta Independiente sobre sociedad civil y solidaridad internacional<sup>114</sup>;

b) Crear un fondo para apoyar el establecimiento de modelos lingüísticos de gran tamaño en todos los idiomas, con miras a proteger la diversidad cultural.

---

<sup>113</sup> La recomendación está en consonancia con los Principios Globales de las Naciones Unidas para la Integridad de la Información. “Mejorar la respuesta ante las crisis. Colaborar con partes interesadas que operen áreas de alto riesgo, establecer procesos tempranos de advertencia y escalonamiento con tasas de respuesta aceleradas y oportunas en contextos de crisis y conflicto. Establecer mecanismos para habilitar el acceso destacado y puntual a información fiable y precisa que sirva al interés público.”

<sup>114</sup> [A/HRC/56/57](#).