

## RECOMENDACIÓN SOBRE LA ÉTICA DE LA INTELIGENCIA ARTIFICIAL

### PREÁMBULO

La Conferencia General de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), reunida en París del 9 al 24 de noviembre de 2021, en su 41ª reunión,

**Reconociendo** las repercusiones positivas y negativas profundas y dinámicas de la inteligencia artificial (IA) en las sociedades, el medio ambiente, los ecosistemas y las vidas humanas, en particular en la mente humana, debido en parte a las nuevas formas en que su utilización influye en el pensamiento, las interacciones y la adopción de decisiones de los seres humanos y afecta a la educación, las ciencias sociales y humanas, las ciencias exactas y naturales, la cultura y la comunicación y la información,

**Recordando** que, con arreglo a su Constitución, la UNESCO se propone contribuir a la paz y a la seguridad estrechando, mediante la educación, la ciencia, la cultura y la comunicación y la información, la colaboración entre las naciones, a fin de asegurar el respeto universal a la justicia, a la ley, a los derechos humanos y a las libertades fundamentales que se reconocen a todos los pueblos del mundo,

**Convencida** de que la Recomendación que aquí se presenta, en su calidad de instrumento normativo elaborado mediante un enfoque mundial, basado en el derecho internacional y centrado en la dignidad humana y los derechos humanos, así como en la igualdad de género, la justicia social y económica y el desarrollo, el bienestar físico y mental, la diversidad, la interconexión, la inclusión y la protección del medio ambiente y de los ecosistemas, puede dar una orientación responsable a las tecnologías de la IA,

**Guiada** por los propósitos y principios de la Carta de las Naciones Unidas,

**Considerando** que las tecnologías de la IA pueden ser de gran utilidad para la humanidad y que todos los países pueden beneficiarse de ellas, pero que también suscitan preocupaciones éticas fundamentales, por ejemplo, en relación con los sesgos que pueden incorporar y exacerbar, lo que puede llegar a provocar discriminación, desigualdad, brechas digitales y exclusión y suponer una amenaza para la diversidad cultural, social y biológica, así como generar divisiones sociales o económicas; la necesidad de transparencia e inteligibilidad del funcionamiento de los algoritmos y los datos con los que han sido entrenados; y su posible impacto en, entre otros, la dignidad humana, los derechos humanos y las libertades fundamentales, la igualdad de género, la democracia, los procesos sociales, económicos, políticos y culturales, las prácticas científicas y de ingeniería, el bienestar animal y el medio ambiente y los ecosistemas,

**Reconociendo también** que las tecnologías de la IA pueden agravar las divisiones y desigualdades existentes en el mundo, dentro de los países y entre ellos, y que es preciso defender la justicia, la confianza y la equidad para que ningún país y ninguna persona se queden atrás, ya sea mediante el acceso equitativo a las tecnologías de la IA y el disfrute de los beneficios que aportan o mediante la protección contra sus consecuencias negativas, reconociendo al mismo tiempo las diferentes circunstancias de los distintos países y respetando el deseo de algunas personas de no participar en todos los avances tecnológicos,

**Consciente** de que todos los países se enfrentan a una aceleración del uso de las tecnologías de la información y la comunicación y las tecnologías de la IA, así como a una necesidad cada vez mayor de alfabetización mediática e informacional, y de que la economía digital presenta importantes desafíos sociales, económicos y ambientales y ofrece oportunidades de compartir los beneficios, especialmente para los países de ingreso mediano bajo, incluidos, entre otros, los países menos adelantados (PMA), los países en desarrollo sin litoral (PDSL) y los pequeños Estados insulares en desarrollo (PEID), que requieren el reconocimiento, la protección y la promoción de las

culturas, los valores y los conocimientos endógenos a fin de desarrollar economías digitales sostenibles,

**Reconociendo además** que las tecnologías de la IA pueden ser beneficiosas para el medio ambiente y los ecosistemas y que, para que esos beneficios se materialicen, no deberían pasarse por alto, sino tenerse en cuenta, los posibles daños y las repercusiones negativas que pueden ocasionar en el medio ambiente y los ecosistemas,

**Observando** que el hecho de tener en cuenta los riesgos y las preocupaciones éticas no debería obstaculizar la innovación y el desarrollo, sino más bien ofrecer nuevas oportunidades y estimular una investigación y una innovación realizadas de manera ética que afiancen las tecnologías de la IA en los derechos humanos y las libertades fundamentales, los valores, los principios y la reflexión moral y ética,

**Recordando también** que en noviembre de 2019 aprobó, en su 40ª reunión, la resolución 40 C/37, en la que encargó a la Directora General que “[preparara] un instrumento normativo internacional sobre la ética de la inteligencia artificial (IA) en forma de recomendación”, que se le presentaría en su 41ª reunión en 2021,

**Reconociendo** que el desarrollo de las tecnologías de la IA requiere un incremento proporcional de la educación en materia de datos y de la alfabetización mediática e informacional, así como el acceso a fuentes de información independientes, pluralistas y fidedignas, en particular en el marco de los esfuerzos destinados a atenuar los riesgos de información errónea, desinformación y discurso de odio, así como los daños causados por el uso indebido de los datos personales,

**Observando también** que el marco normativo para las tecnologías de la IA y sus implicaciones sociales se fundamenta en los marcos jurídicos internacionales y nacionales, los derechos humanos y las libertades fundamentales, la ética, la necesidad de acceder a los datos, la información y los conocimientos, la libertad de investigación e innovación y el bienestar de los seres humanos, del medio ambiente y de los ecosistemas, y conecta los valores y principios éticos con los retos y oportunidades vinculados a las tecnologías de la IA, sobre la base de un entendimiento común y unos objetivos compartidos,

**Reconociendo también** que los valores y principios éticos pueden ayudar a elaborar y aplicar medidas de política y normas jurídicas basadas en los derechos, proporcionando orientación con miras al rápido desarrollo tecnológico,

**Convencida también** de que las normas éticas aceptadas mundialmente para las tecnologías de la IA, que respetan plenamente el derecho internacional, en particular el derecho de los derechos humanos, pueden desempeñar una función esencial en la elaboración de normas relacionadas con la IA en todo el mundo,

**Teniendo presentes** la Declaración Universal de Derechos Humanos (1948), los instrumentos del marco internacional de derechos humanos, entre ellos la Convención sobre el Estatuto de los Refugiados (1951), el Convenio sobre la Discriminación (Empleo y Ocupación) (1958), la Convención Internacional sobre la Eliminación de Todas las Formas de Discriminación Racial (1965), el Pacto Internacional de Derechos Civiles y Políticos (1966), el Pacto Internacional de Derechos Económicos, Sociales y Culturales (1966), la Convención sobre la Eliminación de Todas las Formas de Discriminación contra la Mujer (1979), la Convención sobre los Derechos del Niño (1989), la Convención sobre los Derechos de las Personas con Discapacidad (2006), la Convención relativa a la Lucha contra las Discriminaciones en la Esfera de la Enseñanza (1960) y la Convención sobre la Protección y Promoción de la Diversidad de las Expresiones Culturales (2005), así como cualesquiera otros instrumentos, recomendaciones y declaraciones internacionales pertinentes,

**Tomando nota** de la Declaración sobre el Derecho al Desarrollo (1986); de la Declaración sobre las Responsabilidades de las Generaciones Actuales para con las Generaciones Futuras (1997);

de la Declaración Universal sobre Bioética y Derechos Humanos (2005); de la Declaración de las Naciones Unidas sobre los Derechos de los Pueblos Indígenas (2007); de la resolución de la Asamblea General de las Naciones Unidas sobre el examen de la Cumbre Mundial sobre la Sociedad de la Información (A/RES/70/125) (2015); de la resolución de la Asamblea General de las Naciones Unidas titulada “Transformar nuestro mundo: la Agenda 2030 para el Desarrollo Sostenible” (A/RES/70/1) (2015); de la Recomendación relativa a la Preservación del Patrimonio Documental, comprendido el Patrimonio Digital, y el Acceso al mismo (2015); de la Declaración de Principios Éticos en relación con el Cambio Climático (2017); de la Recomendación sobre la Ciencia y los Investigadores Científicos (2017); de los indicadores sobre la universalidad de Internet (aprobados en 2018 por el Programa Internacional para el Desarrollo de la Comunicación de la UNESCO), incluidos los principios ROAM (aprobados por la Conferencia General de la UNESCO en 2015); de la resolución del Consejo de Derechos Humanos sobre “El derecho a la privacidad en la era digital” (A/HRC/RES/42/15) (2019); y de la resolución del Consejo de Derechos Humanos titulada “Las tecnologías digitales nuevas y emergentes y los derechos humanos” (A/HRC/RES/41/11) (2019),

**Haciendo hincapié** en que debe prestarse atención específica a los países de ingreso mediano bajo, incluidos, entre otros, los PMA, los PDSL y los PEID, ya que, aunque tienen su propia capacidad, han estado insuficientemente representados en el debate sobre la ética de la IA, lo que suscita la preocupación de que se descuiden los conocimientos locales, el pluralismo cultural, los sistemas de valores y las exigencias de equidad mundial para gestionar las repercusiones positivas y negativas de las tecnologías de la IA,

**Consciente también** de que existen numerosas políticas en el plano nacional, así como otros marcos e iniciativas elaborados por las entidades pertinentes de las Naciones Unidas, por organizaciones intergubernamentales, incluidas organizaciones regionales, así como las del sector privado, y por organizaciones profesionales, organizaciones no gubernamentales y de la comunidad científica, relacionados con la ética y la regulación de las tecnologías de la IA,

**Convencida además** de que las tecnologías de la IA pueden aportar importantes beneficios, pero que su consecución también puede hacer aumentar la tensión en lo que respecta a la innovación, el acceso asimétrico a los conocimientos y las tecnologías, en particular la insuficiente educación digital y cívica que limita la capacidad del público para participar en los temas relacionados con la IA, así como las barreras al acceso a la información y las brechas en cuanto a las capacidades humanas e institucionales, los obstáculos al acceso a la innovación tecnológica y la falta de infraestructura física y digital y de marcos reguladores adecuados, en particular los relativos a los datos, problemas todos ellos que deben ser abordados,

**Subrayando** que es necesario fortalecer la cooperación y la solidaridad mundiales, en particular mediante el multilateralismo, para facilitar el acceso equitativo a las tecnologías de la IA y hacer frente a los desafíos que estas plantean en culturas y sistemas éticos diversos e interconectados, para reducir la posibilidad de usos indebidos, para aprovechar todo el potencial que la IA puede aportar, especialmente en el ámbito del desarrollo, y para garantizar que las estrategias nacionales en materia de IA se guíen por principios éticos,

**Teniendo plenamente en cuenta** que el rápido desarrollo de las tecnologías de la IA complica su aplicación y gobernanza éticas, así como el respeto y la protección de la diversidad cultural, y puede perturbar las normas y los valores éticos locales y regionales,

1. **Aprueba**, en este día veintitrés de noviembre de dos mil veintiuno, la presente Recomendación sobre la Ética de la Inteligencia Artificial;
2. **Recomienda** que los Estados Miembros apliquen, de manera voluntaria, las disposiciones de la presente Recomendación mediante la adopción de las medidas adecuadas, en particular las medidas legislativas o de otra índole que puedan ser necesarias, de acuerdo con la práctica constitucional y las estructuras de gobierno de cada Estado, con el fin de dar efecto

en sus respectivas jurisdicciones a los principios y normas enunciados en la Recomendación, de conformidad con el derecho internacional, incluido el derecho internacional de los derechos humanos;

3. **Recomienda también** a los Estados Miembros que hagan partícipes a todas las partes interesadas, incluidas las empresas, para asegurarse de que desempeñan sus respectivas funciones en la aplicación de la presente Recomendación y que señalen la Recomendación a la atención de las autoridades, organismos, organizaciones universitarias y de investigación, instituciones y organizaciones de los sectores público, privado y de la sociedad civil que participan en las tecnologías de la IA, para que el desarrollo y la utilización de esas tecnologías se guíen tanto por una investigación científica sólida como por un análisis y una evaluación éticos.

## I. ÁMBITO DE APLICACIÓN

1. La presente Recomendación trata de las cuestiones éticas relacionadas con el ámbito de la inteligencia artificial en la medida en que competen al mandato de la UNESCO. Aborda la ética de la IA como una reflexión normativa sistemática, basada en un marco integral, global, multicultural y evolutivo de valores, principios y acciones interdependientes, que puede guiar a las sociedades a la hora de afrontar de manera responsable los efectos conocidos y desconocidos de las tecnologías de la IA en los seres humanos, las sociedades y el medio ambiente y los ecosistemas, y les ofrece una base para aceptar o rechazar las tecnologías de la IA. Considera la ética como una base dinámica para la evaluación y la orientación normativas de las tecnologías de la IA, tomando como referencia la dignidad humana, el bienestar y la prevención de daños y apoyándose en la ética de la ciencia y la tecnología.
2. La presente Recomendación no pretende proporcionar una única definición de la IA, ya que tal definición tendría que cambiar con el tiempo en función de los avances tecnológicos. Su objetivo es más bien abordar las características de los sistemas de IA que tienen una importancia ética central. Por tanto, la presente Recomendación considera los sistemas de IA sistemas capaces de procesar datos e información de una manera que se asemeja a un comportamiento inteligente, y abarca generalmente aspectos de razonamiento, aprendizaje, percepción, predicción, planificación o control. Tres elementos ocupan un lugar central en este enfoque:
  - a) los sistemas de IA son tecnologías de procesamiento de la información que integran modelos y algoritmos que producen una capacidad para aprender y realizar tareas cognitivas, dando lugar a resultados como la predicción y la adopción de decisiones en entornos materiales y virtuales. Los sistemas de IA están diseñados para funcionar con diferentes grados de autonomía, mediante la modelización y representación del conocimiento y la explotación de datos y el cálculo de correlaciones. Pueden incluir varios métodos, como, por ejemplo, aunque no exclusivamente:
    - i) el aprendizaje automático, incluido el aprendizaje profundo y el aprendizaje de refuerzo;
    - ii) el razonamiento automático, incluidas la planificación, la programación, la representación del conocimiento y el razonamiento, la búsqueda y la optimización.

Los sistemas de IA pueden utilizarse en los sistemas ciberfísicos, incluidos la Internet de las cosas, los sistemas robóticos, la robótica social y las interfaces entre seres humanos y ordenadores, que comportan el control, la percepción, el procesamiento de los datos recogidos por sensores y el funcionamiento de los actuadores en el entorno en que operan los sistemas de IA;

- b) las cuestiones éticas relativas a los sistemas de IA atañen a todas las etapas del ciclo de vida de estos sistemas, que aquí se entiende que van desde la investigación, la concepción y el desarrollo hasta el despliegue y la utilización, pasando por el mantenimiento, el funcionamiento, la comercialización, la financiación, el seguimiento y la evaluación, la validación, el fin de la utilización, el desmontaje y la terminación. Además, los actores de la IA pueden definirse como todo actor que participe en al menos una etapa del ciclo de vida del sistema de IA y pueden ser tanto personas físicas como jurídicas, por ejemplo, investigadores, programadores, ingenieros, especialistas en datos, usuarios finales, empresas, universidades y entidades públicas y privadas, entre otros;
  - c) los sistemas de IA plantean nuevos tipos de cuestiones éticas que incluyen, aunque no exclusivamente, su impacto en la adopción de decisiones, el empleo y el trabajo, la interacción social, la atención de la salud, la educación, los medios de comunicación, el acceso a la información, la brecha digital, la protección del consumidor y de los datos personales, el medio ambiente, la democracia, el estado de derecho, la seguridad y el mantenimiento del orden, el doble uso y los derechos humanos y las libertades fundamentales, incluidas la libertad de expresión, la privacidad y la no discriminación. Además, surgen nuevos desafíos éticos por el potencial de los algoritmos de la IA para reproducir y reforzar los sesgos existentes, lo que puede exacerbar las formas ya existentes de discriminación, los prejuicios y los estereotipos. Algunas de estas cuestiones tienen que ver con la capacidad de los sistemas de IA para realizar tareas que antes solo podían hacer los seres vivos y que, en algunos casos, incluso se limitaban solo a los seres humanos. Estas características otorgan a los sistemas de IA una función nueva y determinante en las prácticas y la sociedad humanas, así como en su relación con el medio ambiente y los ecosistemas, creando un nuevo contexto para que los niños y los jóvenes crezcan, desarrollen una comprensión del mundo y de sí mismos, comprendan críticamente los medios de comunicación y la información y aprendan a tomar decisiones. A largo plazo, los sistemas de IA podrían disputar al ser humano el sentido especial de la experiencia y la capacidad de actuar que le son propios, lo que plantearía nuevas inquietudes sobre la autocomprensión, la interacción social, cultural y ambiental, la autonomía, la capacidad de actuar, el valor y la dignidad del ser humano, entre otras.
3. En la presente Recomendación se presta especial atención a las repercusiones éticas más amplias de los sistemas de IA en las principales esferas de competencia de la UNESCO —la educación, la ciencia, la cultura y la comunicación y la información—, examinadas en el estudio preliminar sobre la ética de la inteligencia artificial elaborado en 2019 por la Comisión Mundial de Ética del Conocimiento Científico y la Tecnología (COMEST) de la UNESCO:
- a) la educación, porque vivir en sociedades en proceso de digitalización exige nuevas prácticas educativas, una reflexión ética, un pensamiento crítico, prácticas de concepción responsables y nuevas competencias, dadas las implicaciones para el mercado laboral, la empleabilidad y la participación cívica;
  - b) la ciencia, en el sentido más amplio, que abarca todos los ámbitos académicos desde las ciencias exactas y naturales y las ciencias médicas hasta las ciencias sociales y humanas, ya que las tecnologías de la IA aportan nuevas capacidades y enfoques de investigación, influyen en nuestra concepción de la comprensión y la explicación científicas y crean una nueva base para la adopción de decisiones;
  - c) la identidad y la diversidad culturales, ya que las tecnologías de la IA pueden enriquecer las industrias culturales y creativas, pero también pueden dar lugar a una mayor concentración de la oferta de contenidos, los datos, los mercados y los ingresos de la cultura en manos de unos pocos actores, lo que puede tener consecuencias negativas

para la diversidad y el pluralismo de las lenguas, los medios de comunicación, las expresiones culturales, la participación y la igualdad;

- d) la comunicación y la información, ya que las tecnologías de la IA desempeñan una función cada vez más importante en el procesamiento, la estructuración y el suministro de información; las cuestiones del periodismo automatizado y del suministro algorítmico de noticias y la moderación y la conservación de contenidos en los medios sociales y los buscadores son solo algunos ejemplos que plantean cuestiones relacionadas con el acceso a la información, la desinformación, la información errónea, el discurso de odio, la aparición de nuevas formas de narrativa social, la discriminación, la libertad de expresión, la privacidad y la alfabetización mediática e informacional, entre otras.

- 4. La presente Recomendación se dirige a los Estados Miembros, tanto en su calidad de actores de la IA como de autoridades responsables de la elaboración de marcos jurídicos y reguladores a lo largo de todo el ciclo de vida de los sistemas de IA, así como de la promoción de la responsabilidad empresarial. También proporciona orientación ética a todos los actores de la IA, incluidos los sectores público y privado, al sentar las bases para una evaluación del impacto ético de los sistemas de IA a lo largo de su ciclo de vida.

## **II. FINES Y OBJETIVOS**

- 5. La presente Recomendación tiene por objeto servir de base para poner los sistemas de IA al servicio de la humanidad, las personas, las sociedades y el medio ambiente y los ecosistemas, así como para prevenir daños. Aspira también a estimular la utilización de los sistemas de IA con fines pacíficos.
- 6. Además de los marcos éticos relativos a la IA existentes en todo el mundo, la presente Recomendación pretende aportar un instrumento normativo aceptado mundialmente que no solo se centre en la articulación de valores y principios, sino también en su aplicación práctica, mediante recomendaciones de política concretas, haciendo especial hincapié en la inclusión, las cuestiones de igualdad de género y la protección del medio ambiente y los ecosistemas.
- 7. Dado que la complejidad de las cuestiones éticas que rodean a la IA exige la cooperación de múltiples partes interesadas en los diversos niveles y sectores de las comunidades internacionales, regionales y nacionales, la presente Recomendación tiene por objeto permitir que las partes interesadas asuman una responsabilidad compartida basada en un diálogo mundial e intercultural.
- 8. Los objetivos de la presente Recomendación son los siguientes:
  - a) proporcionar un marco universal de valores, principios y acciones para orientar a los Estados en la formulación de sus leyes, políticas u otros instrumentos relativos a la IA, de conformidad con el derecho internacional;
  - b) orientar las acciones de las personas, los grupos, las comunidades, las instituciones y las empresas del sector privado a fin de asegurar la incorporación de la ética en todas las etapas del ciclo de vida de los sistemas de IA;
  - c) proteger, promover y respetar los derechos humanos y las libertades fundamentales, la dignidad humana y la igualdad, incluida la igualdad de género; salvaguardar los intereses de las generaciones presentes y futuras; preservar el medio ambiente, la biodiversidad y los ecosistemas; y respetar la diversidad cultural en todas las etapas del ciclo de vida de los sistemas de IA;
  - d) fomentar el diálogo multidisciplinario y pluralista entre múltiples partes interesadas y la concertación sobre cuestiones éticas relacionadas con los sistemas de IA;

- e) promover el acceso equitativo a los avances y los conocimientos en el ámbito de la IA y el aprovechamiento compartido de los beneficios, prestando especial atención a las necesidades y contribuciones de los países de ingreso mediano bajo, incluidos los PMA, los PDSL y los PEID.

### **III. VALORES Y PRINCIPIOS**

9. Los valores y principios que figuran a continuación deberían ser respetados por todos los actores durante el ciclo de vida de los sistemas de IA, en primer lugar, y, cuando resulte necesario y conveniente, ser promovidos mediante modificaciones de las leyes, los reglamentos y las directrices empresariales existentes y la elaboración de otros nuevos. Todo ello debe ajustarse al derecho internacional, en particular la Carta de las Naciones Unidas y las obligaciones de los Estados Miembros en materia de derechos humanos, y estar en consonancia con los objetivos de sostenibilidad social, política, ambiental, educativa, científica y económica acordados internacionalmente, como los Objetivos de Desarrollo Sostenible (ODS) de las Naciones Unidas.
10. Los valores desempeñan una importante función como ideales que motivan la orientación de las medidas de política y las normas jurídicas. Mientras que el conjunto de valores que se enuncian a continuación inspira, por tanto, un comportamiento deseable y representa los fundamentos de los principios, los principios, por su parte, revelan los valores subyacentes de manera más concreta, de modo que estos últimos puedan aplicarse más fácilmente en las declaraciones de política y las acciones.
11. Si bien todos los valores y principios que se enuncian a continuación son deseables en sí mismos, en cualquier contexto práctico, puede haber tensiones entre ellos. En cualquier situación, será necesaria una evaluación del contexto para gestionar estas posibles tensiones, teniendo en cuenta el principio de proporcionalidad y de conformidad con los derechos humanos y las libertades fundamentales. En todos los casos, toda posible limitación de los derechos humanos y las libertades fundamentales ha de tener una base jurídica y ser razonable, necesaria y proporcional, así como conforme a las obligaciones de los Estados con arreglo al derecho internacional. Para elegir de manera juiciosa entre esas posibilidades, será necesario por lo general colaborar con una amplia gama de partes interesadas apropiadas, haciendo uso del diálogo social, así como de la deliberación ética, la diligencia debida y la evaluación del impacto.
12. La fiabilidad y la integridad del ciclo de vida de los sistemas de IA son esenciales para velar por que las tecnologías de la IA estén al servicio del bien de la humanidad, las personas, las sociedades y el medio ambiente y los ecosistemas, y encarnen los valores y principios enunciados en la presente Recomendación. Las personas deberían tener buenas razones para confiar en que los sistemas de IA pueden aportar beneficios individuales y compartidos, al tiempo que se adoptan medidas adecuadas para atenuar los riesgos. Un requisito esencial para la fiabilidad es que, a lo largo de su ciclo de vida, los sistemas de IA estén sujetos a un seguimiento exhaustivo por las partes interesadas pertinentes, según corresponda. Dado que la fiabilidad resulta de la aplicación de los principios expuestos en este documento, todas las medidas políticas propuestas en la presente Recomendación están dirigidas a promover la fiabilidad en todas las etapas del ciclo de vida de los sistemas de IA.

#### **III.1 VALORES**

##### **Respeto, protección y promoción de los derechos humanos, las libertades fundamentales y la dignidad humana**

13. La dignidad inviolable e intrínseca de cada ser humano constituye la base del sistema universal, indivisible, inalienable, interdependiente e interrelacionado de derechos humanos y libertades fundamentales. Por consiguiente, el respeto, la protección y la promoción de la

dignidad humana y de los derechos establecidos por el derecho internacional, en particular el derecho internacional de los derechos humanos, son esenciales a lo largo del ciclo de vida de los sistemas de IA. La dignidad humana tiene que ver con el reconocimiento del valor intrínseco e igual de cada ser humano, con independencia de su raza, color, ascendencia, género, edad, idioma, religión, opiniones políticas, origen nacional, étnico o social, condición económica o social de nacimiento, discapacidad o cualquier otro motivo.

14. Ningún ser humano ni comunidad humana debería sufrir daños o sometimiento, ya sean de carácter físico, económico, social, político, cultural o mental, durante ninguna etapa del ciclo de vida de los sistemas de IA. A lo largo de su ciclo de vida, los sistemas de IA deberían mejorar la calidad de vida de los seres humanos, dejando a las personas o los grupos la tarea de definir el concepto de “calidad de vida”, siempre que como resultado de esa definición no se produzca ninguna violación o abuso de los derechos humanos y las libertades fundamentales ni de la dignidad de las personas.
15. Las personas pueden interactuar con los sistemas de IA a lo largo de su ciclo de vida y recibir su asistencia, por ejemplo, para el cuidado de las personas vulnerables o en situación de vulnerabilidad, incluidos, entre otros, los niños, las personas de edad, las personas con discapacidad o los enfermos. En el marco de esas interacciones, las personas nunca deberían ser cosificadas, su dignidad no debería ser menoscabada de ninguna otra manera, y sus derechos humanos y libertades fundamentales nunca deberían ser objeto de violación o abusos.
16. Los derechos humanos y las libertades fundamentales han de ser respetados, protegidos y promovidos a lo largo del ciclo de vida de los sistemas de IA. Los gobiernos, el sector privado, la sociedad civil, las organizaciones internacionales, las comunidades técnicas y las universidades deben respetar los instrumentos y marcos de derechos humanos en sus intervenciones en los procesos que rodean el ciclo de vida de los sistemas de IA. Es necesario que las nuevas tecnologías proporcionen nuevos medios para promover, defender y ejercer los derechos humanos, y no para vulnerarlos.

### **Prosperidad del medio ambiente y los ecosistemas**

17. La prosperidad del medio ambiente y los ecosistemas debería ser reconocida, protegida y promovida a lo largo del ciclo de vida de los sistemas de IA. Además, el medio ambiente y los ecosistemas son una necesidad existencial para que la humanidad y los demás seres vivos puedan disfrutar de los beneficios derivados de los avances de la IA.
18. Todos los actores que participan en el ciclo de vida de los sistemas de IA deben respetar el derecho internacional y las leyes, normas y prácticas nacionales aplicables, como la precaución, concebidas para la protección y la restauración del medio ambiente y los ecosistemas y para el desarrollo sostenible. Deberían reducir el impacto ambiental de los sistemas de IA, en particular, aunque no exclusivamente, su huella de carbono, para asegurar la minimización del cambio climático y los factores de riesgo ambiental, y prevenir la explotación, la utilización y la transformación no sostenibles de los recursos naturales que contribuyen al deterioro del medio ambiente y a la degradación de los ecosistemas.

### **Garantizar la diversidad y la inclusión**

19. El respeto, la protección y la promoción de la diversidad y la inclusión deberían garantizarse a lo largo del ciclo de vida de los sistemas de IA, de conformidad con el derecho internacional, en particular el derecho de los derechos humanos. Para ello se podría promover la participación activa de todas las personas o grupos, con independencia de su raza, color, ascendencia, género, edad, idioma, religión, opiniones políticas, origen nacional, étnico o social, condición económica o social de nacimiento, discapacidad o cualquier otro motivo.

20. La diversidad de las elecciones de estilo de vida, creencias, opiniones, expresiones o experiencias personales, incluida la utilización opcional de sistemas de IA y la concepción conjunta de estas arquitecturas, no debería restringirse durante ninguna etapa del ciclo de vida de dichos sistemas.
21. Además, habría que esforzarse, principalmente mediante la cooperación internacional, por paliar la falta de infraestructura, educación y competencias tecnológicas necesarias, así como de marcos jurídicos, que afecta a algunas comunidades, en particular en los países de ingreso mediano bajo, los PMA, los PDSL y los PEID, y no aprovecharse nunca de esa situación.

### **Vivir en sociedades pacíficas, justas e interconectadas**

22. Los actores de la IA deberían propiciar sociedades pacíficas y justas, sobre la base de un futuro interconectado en beneficio de todos, compatibles con los derechos humanos y las libertades fundamentales, y participar en su construcción. El valor de vivir en sociedades pacíficas y justas apunta al potencial de los sistemas de IA para contribuir a lo largo de su ciclo de vida a la interconexión de todas las criaturas vivas entre sí y con el medio natural.
23. La noción de interconexión de los seres humanos se basa en el conocimiento de que cada uno de ellos pertenece a un todo más amplio, que prospera cuando todas las partes que lo constituyen pueden progresar. Vivir en sociedades pacíficas, justas e interconectadas requiere un vínculo orgánico, inmediato y no calculado de solidaridad, caracterizado por una búsqueda permanente de relaciones pacíficas, tendentes al cuidado de los demás y del medio natural en el sentido más amplio del término.
24. Este valor exige que se promuevan la paz, la inclusión y la justicia, la equidad y la interconexión durante el ciclo de vida de los sistemas de IA, en la medida en que los procesos de dicho ciclo de vida no deberían segregar ni cosificar a los seres humanos y las comunidades ni mermar su libertad, su autonomía de decisión y su seguridad, así como tampoco dividir y enfrentar entre sí a las personas y los grupos ni amenazar la coexistencia entre los seres humanos, los demás seres vivos y el medio natural.

## **III.2 PRINCIPIOS**

### **Proporcionalidad e inocuidad**

25. Debería reconocerse que las tecnologías de la IA no garantizan necesariamente, por sí mismas, la prosperidad de los seres humanos ni del medio ambiente y los ecosistemas. Además, ninguno de los procesos relacionados con el ciclo de vida de los sistemas de IA podrá ir más allá de lo necesario para lograr propósitos u objetivos legítimos, y esos procesos deberían ser adecuados al contexto. En caso de que pueda producirse cualquier daño para los seres humanos, los derechos humanos y las libertades fundamentales, las comunidades y la sociedad en general, o para el medio ambiente y los ecosistemas, debería garantizarse la aplicación de procedimientos de evaluación de riesgos y la adopción de medidas para impedir que ese daño se produzca.
26. La decisión de utilizar sistemas de IA y la elección del método de IA deberían justificarse de las siguientes maneras: a) el método de IA elegido debería ser adecuado y proporcional para lograr un objetivo legítimo determinado; b) el método de IA elegido no debería vulnerar los valores fundamentales enunciados en el presente documento, en particular, su utilización no debe constituir una violación o un abuso de los derechos humanos; y c) el método de IA elegido debería ser adecuado al contexto y basarse en fundamentos científicos rigurosos. En los casos en que se entienda que las decisiones tienen un impacto irreversible o difícil de revertir o que pueden implicar decisiones de vida o muerte, la decisión final debería ser adoptada por un ser humano. En particular, los sistemas de IA no deberían utilizarse con fines de calificación social o vigilancia masiva.

## **Seguridad y protección**

27. Los daños no deseados (riesgos de seguridad) y las vulnerabilidades a los ataques (riesgos de protección) deberían ser evitados y deberían tenerse en cuenta, prevenirse y eliminarse a lo largo del ciclo de vida de los sistemas de IA para garantizar la seguridad y la protección de los seres humanos, del medio ambiente y de los ecosistemas. La seguridad y la protección de la IA se propiciarán mediante el desarrollo de marcos de acceso a los datos que sean sostenibles, respeten la privacidad y fomenten un mejor entrenamiento y validación de los modelos de IA que utilicen datos de calidad.

## **Equidad y no discriminación**

28. Los actores de la IA deberían promover la justicia social, salvaguardar la equidad y luchar contra todo tipo de discriminación, de conformidad con el derecho internacional. Ello supone adoptar un enfoque inclusivo para garantizar que los beneficios de las tecnologías de la IA estén disponibles y sean accesibles para todos, teniendo en cuenta las necesidades específicas de los diferentes grupos de edad, los sistemas culturales, los diferentes grupos lingüísticos, las personas con discapacidad, las niñas y las mujeres y las personas desfavorecidas, marginadas y vulnerables o en situación de vulnerabilidad. Los Estados Miembros deberían esforzarse por promover un acceso inclusivo para todos, incluidas las comunidades locales, a sistemas de IA con contenidos y servicios adaptados al contexto local, y respetando el multilingüismo y la diversidad cultural. Los Estados Miembros deberían esforzarse por reducir las brechas digitales y garantizar el acceso inclusivo al desarrollo de la IA y la participación en él. En el plano nacional, los Estados Miembros deberían promover la equidad entre las zonas rurales y urbanas y entre todas las personas, con independencia de su raza, color, ascendencia, género, edad, idioma, religión, opiniones políticas, origen nacional, étnico o social, condición económica o social de nacimiento, discapacidad o cualquier otro motivo, en lo que respecta al acceso al ciclo de vida de los sistemas de IA y la participación en él. En el plano internacional, los países más avanzados tecnológicamente tienen la responsabilidad de ser solidarios con los menos avanzados para garantizar que los beneficios de las tecnologías de la IA se compartan de manera que, para estos últimos, el acceso al ciclo de vida de los sistemas de IA y la participación en él contribuyan a un orden mundial más equitativo en lo que respecta a la información, la comunicación, la cultura, la educación, la investigación y la estabilidad socioeconómica y política.
29. Los actores de la IA deberían hacer todo lo razonablemente posible por reducir al mínimo y evitar reforzar o perpetuar aplicaciones y resultados discriminatorios o sesgados a lo largo del ciclo de vida de los sistemas de IA, a fin de garantizar la equidad de dichos sistemas. Debería disponerse de un recurso efectivo contra la discriminación y la determinación algorítmica sesgada.
30. Además, es necesario abordar las brechas digital y de conocimientos dentro de los países y entre ellos a lo largo del ciclo de vida de los sistemas de IA, en particular en lo que respecta al acceso y la calidad del acceso a la tecnología y los datos, de conformidad con los marcos jurídicos nacionales, regionales e internacionales pertinentes, así como en lo referente a la conectividad, los conocimientos y las competencias y a la participación significativa de las comunidades afectadas, de manera que todas las personas sean tratadas equitativamente.

## **Sostenibilidad**

31. El desarrollo de sociedades sostenibles depende del logro de un complejo conjunto de objetivos relacionados con distintas dimensiones humanas, sociales, culturales, económicas y ambientales. La llegada de las tecnologías de la IA puede beneficiar los objetivos de sostenibilidad o dificultar su consecución, dependiendo de la forma en que se apliquen en países con diferentes niveles de desarrollo. Por consiguiente, la evaluación continua de los efectos humanos, sociales, culturales, económicos y ambientales de las tecnologías de la IA

debería llevarse a cabo con pleno conocimiento de las repercusiones de dichas tecnologías en la sostenibilidad como un conjunto de metas en constante evolución en toda una serie de dimensiones, como las que se definen actualmente en los Objetivos de Desarrollo Sostenible (ODS) de las Naciones Unidas.

### **Derecho a la intimidad y protección de datos**

32. La privacidad, que constituye un derecho esencial para la protección de la dignidad, la autonomía y la capacidad de actuar de los seres humanos, debe ser respetada, protegida y promovida a lo largo del ciclo de vida de los sistemas de IA. Es importante que los datos para los sistemas de IA se recopilen, utilicen, compartan, archiven y supriman de forma coherente con el derecho internacional y acorde con los valores y principios enunciados en la presente Recomendación, respetando al mismo tiempo los marcos jurídicos nacionales, regionales e internacionales pertinentes.
33. Deberían establecerse en los planos nacional o internacional, de acuerdo con un enfoque de múltiples partes interesadas, marcos de protección de datos y mecanismos de gobernanza adecuados, protegidos por los sistemas judiciales y aplicados a lo largo del ciclo de vida de los sistemas de IA. Los marcos de protección de datos y todo mecanismo conexo deberían tomar como referencia los principios y normas internacionales de protección de datos relativos a la recopilación, la utilización y la divulgación de datos personales y al ejercicio de sus derechos por parte de los interesados, garantizando al mismo tiempo un objetivo legítimo y una base jurídica válida para el tratamiento de los datos personales, incluido el consentimiento informado.
34. Los sistemas algorítmicos requieren evaluaciones adecuadas del impacto en la privacidad, que incluyan también consideraciones sociales y éticas de su utilización y un empleo innovador del enfoque de privacidad desde la etapa de concepción. Los actores de la IA deben asumir la responsabilidad de la concepción y la aplicación de los sistemas de IA de manera que se garantice la protección de la información personal durante todo el ciclo de vida del sistema de IA.

### **Supervisión y decisión humanas**

35. Los Estados Miembros deberían velar por que siempre sea posible atribuir la responsabilidad ética y jurídica, en cualquier etapa del ciclo de vida de los sistemas de IA, así como en los casos de recurso relacionados con sistemas de IA, a personas físicas o a entidades jurídicas existentes. La supervisión humana se refiere, por tanto, no solo a la supervisión humana individual, sino también a la supervisión pública inclusiva, según corresponda.
36. Puede ocurrir que, en algunas ocasiones, los seres humanos decidan depender de los sistemas de IA por razones de eficacia, pero la decisión de ceder el control en contextos limitados seguirá recayendo en los seres humanos, ya que estos pueden recurrir a los sistemas de IA en la adopción de decisiones y en la ejecución de tareas, pero un sistema de IA nunca podrá reemplazar la responsabilidad final de los seres humanos y su obligación de rendir cuentas. Por regla general, las decisiones de vida o muerte no deberían cederse a los sistemas de IA.

### **Transparencia y explicabilidad**

37. La transparencia y la explicabilidad de los sistemas de IA suelen ser condiciones previas fundamentales para garantizar el respeto, la protección y la promoción de los derechos humanos, las libertades fundamentales y los principios éticos. La transparencia es necesaria para que los regímenes nacionales e internacionales pertinentes en materia de responsabilidad funcionen eficazmente. La falta de transparencia también podría mermar la posibilidad de impugnar eficazmente las decisiones basadas en resultados producidos por los

sistemas de IA y, por lo tanto, podría vulnerar el derecho a un juicio imparcial y a un recurso efectivo, y limita los ámbitos en los que estos sistemas pueden utilizarse legalmente.

38. Si bien hay que hacer todo lo posible por aumentar la transparencia y la explicabilidad de los sistemas de IA, incluidos los que tienen un efecto extraterritorial, a lo largo de su ciclo de vida para respaldar la gobernanza democrática, el grado de transparencia y explicabilidad debería ser siempre adecuado al contexto y al efecto, ya que puede ser necesario encontrar un equilibrio entre la transparencia y la explicabilidad y otros principios como la privacidad, la seguridad y la protección. Las personas deberían estar plenamente informadas cuando una decisión se basa en algoritmos de IA o se toma a partir de ellos, en particular cuando afecta a su seguridad o a sus derechos humanos; en esas circunstancias, deberían tener la oportunidad de solicitar explicaciones e información al actor de la IA o a las instituciones del sector público correspondientes. Además, las personas deberían poder conocer los motivos por los que se ha tomado una decisión que afecta a sus derechos y libertades y tener la posibilidad de presentar alegaciones a un miembro del personal de la empresa del sector privado o de la institución del sector público habilitado para revisar y enmendar la decisión. Los actores de la IA deberían informar a los usuarios cuando un producto o servicio se proporcione directamente o con la ayuda de sistemas de IA de manera adecuada y oportuna.
39. Desde un punto de vista sociotécnico, una mayor transparencia contribuye a crear sociedades más pacíficas, justas, democráticas e inclusivas. Posibilita un escrutinio público que puede reducir la corrupción y la discriminación, y también puede ayudar a detectar y prevenir los efectos negativos sobre los derechos humanos. La transparencia tiene como objetivo proporcionar información adecuada a los respectivos destinatarios para permitir su comprensión y fomentar la confianza. En el caso específico de los sistemas de IA, la transparencia puede permitir a las personas comprender cómo se implementa cada etapa de un sistema de IA, en función del contexto y la sensibilidad del sistema en cuestión. También puede proporcionar información sobre los factores que influyen en una predicción o decisión específicas, y sobre la existencia o no de garantías adecuadas (como medidas de seguridad o de equidad). En los casos de amenazas graves con repercusiones adversas para los derechos humanos, la transparencia puede requerir también que se compartan códigos o conjuntos de datos.
40. La explicabilidad supone hacer inteligibles los resultados de los sistemas de IA y facilitar información sobre ellos. La explicabilidad de los sistemas de IA también se refiere a la inteligibilidad de la entrada, salida y funcionamiento de cada componente algorítmico y la forma en que contribuye a los resultados de los sistemas. Así pues, la explicabilidad está estrechamente relacionada con la transparencia, ya que los resultados y los subprocesos que conducen a ellos deberían aspirar a ser comprensibles y trazables, apropiados al contexto. Los actores de la IA deberían comprometerse a velar por que los algoritmos desarrollados sean explicables. En el caso de las aplicaciones de IA cuyo impacto en el usuario final no es temporal, fácilmente reversible o de bajo riesgo, debería garantizarse que se proporcione una explicación satisfactoria con toda decisión que haya dado lugar a la acción tomada, a fin de que el resultado se considere transparente.
41. La transparencia y la explicabilidad están estrechamente relacionadas con las medidas adecuadas de responsabilidad y rendición de cuentas, así como con la fiabilidad de los sistemas de IA.

### **Responsabilidad y rendición de cuentas**

42. Los actores de la IA y los Estados Miembros deberían respetar, proteger y promover los derechos humanos y las libertades fundamentales, y deberían también fomentar la protección del medio ambiente y los ecosistemas, asumiendo su responsabilidad ética y jurídica respectiva, de conformidad con el derecho nacional e internacional, en particular las obligaciones de los Estados Miembros en materia de derechos humanos, y con las directrices

éticas establecidas durante todo el ciclo de vida de los sistemas de IA, incluso con respecto a los actores de la IA dentro de su territorio y bajo su control efectivos. La responsabilidad ética y la obligación de rendir cuentas de las decisiones y las acciones basadas de alguna manera en un sistema de IA siempre deberían ser atribuibles, en última instancia, a los actores de la IA conforme a la función que tengan en el ciclo de vida del sistema de IA.

43. Deberían elaborarse mecanismos adecuados de supervisión, evaluación del impacto, auditoría y diligencia debida, incluso en lo que se refiere a la protección de los denunciantes de irregularidades, para garantizar la rendición de cuentas respecto de los sistemas de IA y de su impacto a lo largo de su ciclo de vida. Dispositivos tanto técnicos como institucionales deberían garantizar la auditabilidad y la trazabilidad (del funcionamiento) de los sistemas de IA, en particular para intentar solucionar cualquier conflicto con las normas relativas a los derechos humanos y las amenazas al bienestar del medio ambiente y los ecosistemas.

### **Sensibilización y educación**

44. La sensibilización y la comprensión del público respecto de las tecnologías de la IA y el valor de los datos deberían promoverse mediante una educación abierta y accesible, la participación cívica, las competencias digitales y la capacitación en materia de ética de la IA, la alfabetización mediática e informacional y la capacitación dirigida conjuntamente por los gobiernos, las organizaciones intergubernamentales, la sociedad civil, las universidades, los medios de comunicación, los dirigentes comunitarios y el sector privado, y teniendo en cuenta la diversidad lingüística, social y cultural existente, a fin de garantizar una participación pública efectiva, de modo que todos los miembros de la sociedad puedan adoptar decisiones informadas sobre su utilización de los sistemas de IA y estén protegidos de influencias indebidas.
45. El aprendizaje sobre el impacto de los sistemas de IA debería incluir el aprendizaje sobre los derechos humanos y las libertades fundamentales, a través de ellos y para ellos, lo que significa que el enfoque y la comprensión de los sistemas de IA deberían basarse en el impacto de estos sistemas en los derechos humanos y el acceso a esos derechos, así como en el medio ambiente y los ecosistemas.

### **Gobernanza y colaboración adaptativas y de múltiples partes interesadas**

46. En la utilización de datos deben respetarse el derecho internacional y la soberanía nacional. Esto significa que los Estados, en cumplimiento del derecho internacional, pueden regular los datos generados dentro de sus territorios o que pasan por ellos y adoptar medidas para la regulación efectiva de los datos, en particular su protección, sobre la base del respeto del derecho a la privacidad, de conformidad con el derecho internacional y otras normas relativas a los derechos humanos.
47. La participación de las diferentes partes interesadas a lo largo del ciclo de vida de los sistemas de IA es necesaria para garantizar enfoques inclusivos de la gobernanza de la IA, de modo que los beneficios puedan ser compartidos por todos, y para contribuir al desarrollo sostenible. Entre las partes interesadas figuran, entre otros, los gobiernos, las organizaciones intergubernamentales, la comunidad técnica, la sociedad civil, los investigadores y los círculos universitarios, los medios de comunicación, los responsables de la educación, los encargados de formular políticas, las empresas del sector privado, las instituciones de derechos humanos y los organismos de fomento de la igualdad, los órganos de vigilancia de la lucha contra la discriminación y los grupos de jóvenes y niños. Convendría adoptar normas abiertas y garantizar la interoperabilidad para facilitar la colaboración. Deberían adoptarse medidas para tener en cuenta los cambios en las tecnologías y la aparición de nuevos grupos de partes interesadas y para permitir una participación significativa de las personas, las comunidades y los grupos marginados y, si procede, en el caso de los pueblos indígenas, el respeto de su autonomía en la gestión de sus datos.

#### **IV. ÁMBITOS DE ACCIÓN POLÍTICA**

48. Las acciones políticas que se describen en los siguientes ámbitos de actuación ponen en práctica los valores y principios enunciados en la presente Recomendación. La principal acción consiste en que los Estados Miembros establezcan medidas eficaces, por ejemplo marcos o mecanismos normativos, y velen por que otras partes interesadas, como las empresas del sector privado, las instituciones universitarias y de investigación y la sociedad civil, se adhieran a ellas, sobre todo alentando a todas las partes interesadas a que elaboren instrumentos de evaluación del impacto en los derechos humanos, el estado de derecho, la democracia y la ética, así como instrumentos de diligencia debida, de conformidad con las orientaciones, incluidos los Principios Rectores sobre las Empresas y los Derechos Humanos de las Naciones Unidas. El proceso de elaboración de esas políticas o mecanismos debería incluir a todas las partes interesadas y tener en cuenta las circunstancias y prioridades de cada Estado Miembro. La UNESCO puede ser un asociado y apoyar a los Estados Miembros en la elaboración, así como en el seguimiento y la evaluación, de los mecanismos de política.
49. La UNESCO reconoce que los Estados Miembros se encontrarán en diferentes estadios de preparación para aplicar la presente Recomendación, desde los puntos de vista científico, tecnológico, económico, educativo, jurídico, regulador, de infraestructura, social y cultural, entre otros. Cabe señalar que, aquí, el “estadio de preparación” es un estado dinámico. Por consiguiente, a fin de posibilitar la aplicación efectiva de la presente Recomendación, la UNESCO: 1) elaborará una metodología de evaluación del estadio de preparación para ayudar a los Estados Miembros interesados a determinar su situación en momentos concretos de su trayectoria de preparación a través de un conjunto de dimensiones; y 2) garantizará el apoyo a los Estados Miembros interesados en lo que respecta a la elaboración de una metodología de la UNESCO para la evaluación del impacto ético de las tecnologías de la IA y el intercambio de mejores prácticas, directrices de evaluación y otros mecanismos y trabajo analítico.

#### **ÁMBITO DE ACTUACIÓN 1: EVALUACIÓN DEL IMPACTO ÉTICO**

50. Los Estados Miembros deberían establecer marcos de evaluación del impacto, como evaluaciones del impacto ético, para determinar y analizar los beneficios, los problemas y los riesgos de los sistemas de IA, así como medidas adecuadas de prevención, atenuación y seguimiento de los riesgos, entre otros mecanismos de garantía. Esas evaluaciones del impacto deberían revelar las repercusiones en los derechos humanos y las libertades fundamentales, en particular, aunque no exclusivamente, los derechos de las personas marginadas y vulnerables o en situación de vulnerabilidad, los derechos laborales, el medio ambiente y los ecosistemas, así como las consecuencias éticas y sociales, y facilitar la participación ciudadana, de conformidad con los valores y principios enunciados en la presente Recomendación.
51. Los Estados Miembros y las empresas del sector privado deberían desarrollar mecanismos de diligencia debida y supervisión para determinar, prevenir y atenuar los riesgos y rendir cuentas de la forma en que abordan el impacto de los sistemas de IA en el respeto de los derechos humanos, el estado de derecho y las sociedades inclusivas. Los Estados Miembros deberían también poder evaluar los efectos socioeconómicos de los sistemas de IA en la pobreza y velar por que la brecha entre los ricos y los pobres, así como la brecha digital entre los países y dentro de ellos, no aumenten con la adopción masiva de tecnologías de la IA en la actualidad y en el futuro. Para ello, en particular, deberían aplicarse protocolos de transparencia ejecutables, que correspondan al acceso a la información, incluida la información de interés público en poder de entidades privadas. Los Estados Miembros, las empresas del sector privado y la sociedad civil deberían investigar los efectos sociológicos y psicológicos de las recomendaciones basadas en la IA sobre los seres humanos en lo que respecta a su autonomía de decisión. Los sistemas de IA considerados riesgos potenciales para los derechos humanos deberían ser ampliamente probados por los actores de la IA,

incluso en condiciones reales si es necesario, en el marco de la evaluación del impacto ético, antes de sacarlos al mercado.

52. Los Estados Miembros y las empresas deberían aplicar medidas adecuadas para vigilar todas las etapas del ciclo de vida de los sistemas de IA en el marco de la evaluación del impacto ético, incluidos el funcionamiento de los algoritmos utilizados para la adopción de decisiones, los datos y los actores de la IA que participan en el proceso, especialmente en los servicios públicos y en los casos en que se necesita una interacción directa con el usuario final. Las obligaciones de los Estados Miembros en materia de derechos humanos deberían formar parte de los aspectos éticos de las evaluaciones de los sistemas de IA.
53. Los gobiernos deberían adoptar un marco regulador que establezca un procedimiento para que las autoridades públicas, en particular, lleven a cabo evaluaciones del impacto ético de los sistemas de IA a fin de anticipar las repercusiones, atenuar los riesgos, evitar las consecuencias perjudiciales, facilitar la participación de los ciudadanos y hacer frente a los desafíos sociales. La evaluación también debería establecer mecanismos de supervisión adecuados, como la auditabilidad, la trazabilidad y la explicabilidad, que permitan evaluar los algoritmos, los datos y los procesos de concepción, así como incluir un examen externo de los sistemas de IA. Las evaluaciones del impacto ético deberían ser transparentes y abiertas al público, cuando proceda. También deberían ser multidisciplinarias, multiculturales, pluralistas e inclusivas y contar con múltiples partes interesadas. Se debería exigir a las autoridades públicas que supervisen los sistemas de IA que hayan implantado o desplegado, mediante la introducción de mecanismos e instrumentos adecuados.

## **ÁMBITO DE ACTUACIÓN 2: GOBERNANZA Y ADMINISTRACIÓN ÉTICAS**

54. Los Estados Miembros deberían velar por que los mecanismos de gobernanza de la IA sean inclusivos, transparentes, multidisciplinarios y multilaterales (lo que incluye la posibilidad de atenuar y reparar daños más allá de las fronteras) y cuenten con múltiples partes interesadas. En particular, la gobernanza debería incluir aspectos de previsión y dispositivos eficaces de protección, seguimiento de los efectos, aplicación y reparación.
55. Los Estados Miembros deberían velar por que se investiguen y reparen los daños causados mediante sistemas de IA, estableciendo mecanismos de aplicación estrictos y medidas correctivas, a fin de asegurarse de que los derechos humanos, las libertades fundamentales y el estado de derecho son respetados en el mundo digital y en el mundo físico. Entre esos mecanismos y medidas deberían figurar mecanismos de reparación aportados por empresas de los sectores público y privado. Con ese fin, debería promoverse la auditabilidad y la trazabilidad de los sistemas de IA. Además, los Estados Miembros deberían reforzar sus capacidades institucionales para cumplir con este compromiso y deberían colaborar con investigadores y otras partes interesadas para investigar, prevenir y minimizar todo uso potencialmente malicioso de los sistemas de IA.
56. Se alienta a los Estados Miembros a que elaboren estrategias nacionales y regionales en materia de IA y consideren la posibilidad de adoptar formas de gobernanza “blanda”, por ejemplo, un mecanismo de certificación para los sistemas de IA y el reconocimiento mutuo de su certificación, con arreglo a la sensibilidad del ámbito de aplicación y al impacto previsto en los derechos humanos, el medio ambiente y los ecosistemas, así como otras consideraciones éticas establecidas en la presente Recomendación. Dicho mecanismo podría incluir diferentes niveles de auditoría de los sistemas, los datos y el cumplimiento de las directrices éticas y de los requisitos de procedimiento teniendo en cuenta los aspectos éticos. Al mismo tiempo, no debería obstaculizar la innovación ni poner en situación de desventaja a las pequeñas y medianas empresas o las empresas incipientes, la sociedad civil y las organizaciones científicas y de investigación como resultado de una carga administrativa excesiva. Ese mecanismo también debería incluir un componente de seguimiento periódico para garantizar la solidez del sistema de IA y el mantenimiento de su integridad y su cumplimiento de las

directrices éticas durante todo su ciclo de vida, exigiendo una nueva certificación si fuera necesario.

57. Los Estados Miembros y las autoridades públicas deberían realizar una autoevaluación transparente de los sistemas de IA existentes y propuestos, en la cual se debería analizar, en particular, si la adopción de la IA es apropiada y, en caso afirmativo, realizar una nueva evaluación para establecer cuál es el método adecuado, así como una evaluación para determinar si dicha adopción daría lugar a violaciones o abusos de las obligaciones de los Estados Miembros en materia de derechos humanos y, si así fuera, prohibir su utilización.
58. Los Estados Miembros deberían alentar a las entidades públicas, las empresas del sector privado y las organizaciones de la sociedad civil a que incorporen a diferentes partes interesadas a su gobernanza en materia de IA y consideren la posibilidad de añadir una función de responsable independiente de la ética de la IA o algún otro mecanismo para supervisar las actividades relacionadas con la evaluación del impacto ético, las auditorías y el seguimiento continuo, así como para garantizar la orientación ética de los sistemas de IA. Se alienta a los Estados Miembros, las empresas del sector privado y las organizaciones de la sociedad civil a que, con el respaldo de la UNESCO, creen una red de responsables independientes de la ética de la IA para apoyar este proceso en los planos nacional, regional e internacional.
59. Los Estados Miembros deberían fomentar el desarrollo y la accesibilidad de un ecosistema digital para el desarrollo ético e inclusivo de los sistemas de IA en el plano nacional, en particular con miras a reducir las diferencias de acceso durante el ciclo de vida de los sistemas de IA, contribuyendo al mismo tiempo a la colaboración internacional. Ese ecosistema incluiría, en particular, tecnologías e infraestructuras digitales y mecanismos para compartir los conocimientos en materia de IA, según proceda.
60. Los Estados Miembros deberían establecer mecanismos, en colaboración con las organizaciones internacionales, las empresas transnacionales, las instituciones universitarias y la sociedad civil, para garantizar la participación activa de todos los Estados Miembros, especialmente los países de ingreso mediano bajo, en particular los PMA, los PDSL y los PEID, en los debates internacionales sobre la gobernanza de la IA. Esto puede hacerse mediante la provisión de fondos, garantizando la participación regional en condiciones de igualdad, o mediante cualquier otro mecanismo. Además, para velar por que los foros sobre la IA sean inclusivos, los Estados Miembros deberían facilitar los desplazamientos de los actores de la IA dentro y fuera de su territorio, especialmente los de los países de ingreso mediano bajo, en particular los PMA, los PDSL y los PEID, para que puedan participar en esos foros.
61. Las modificaciones de la legislación nacional existente o la elaboración de una nueva legislación nacional en materia de sistemas de IA deben ajustarse a las obligaciones de los Estados Miembros en materia de derechos humanos y promover los derechos humanos y las libertades fundamentales a lo largo del ciclo de vida de esos sistemas. La promoción de los derechos humanos y las libertades fundamentales también debería adoptar la forma de iniciativas de gobernanza, buenos ejemplos de prácticas de colaboración en relación con los sistemas de IA y directrices técnicas y metodológicas nacionales e internacionales a medida que avancen las tecnologías de la IA. En sus prácticas relativas a los sistemas de IA, diversos sectores, incluido el privado, deben respetar, proteger y promover los derechos humanos y las libertades fundamentales utilizando los instrumentos existentes y nuevos en combinación con la presente Recomendación.
62. Los Estados Miembros que adquieran sistemas de IA para casos de utilización sensible en materia de derechos humanos, como la aplicación de la ley, la asistencia social, el empleo, los medios de comunicación y los proveedores de información, la atención de la salud y el sistema judicial independiente, deberían prever mecanismos para vigilar el impacto social y

económico de dichos sistemas mediante autoridades de supervisión adecuadas, como autoridades independientes de protección de datos, una supervisión sectorial y organismos públicos encargados de la supervisión.

63. Los Estados Miembros deberían reforzar la capacidad del poder judicial para adoptar decisiones relacionadas con los sistemas de IA en el marco del estado de derecho y de conformidad con el derecho y las normas internacionales, en particular en lo que respecta a la utilización de los sistemas de IA en sus deliberaciones, velando al mismo tiempo por que se respete el principio de la supervisión humana. En caso de que los sistemas de IA sean utilizados por el poder judicial, se necesitan suficientes salvaguardias para garantizar, entre otras cosas, la protección de los derechos humanos fundamentales, el estado de derecho, la independencia judicial y el principio de supervisión humana, así como para asegurar un desarrollo y una utilización de los sistemas de IA en el poder judicial que sean fiables, orientados al interés público y centrados en el ser humano.
64. Los Estados Miembros deberían velar por que los gobiernos y las organizaciones multilaterales desempeñen una función de liderazgo para garantizar la seguridad y la protección de los sistemas de IA, con la participación de múltiples partes interesadas. En concreto, los Estados Miembros, las organizaciones internacionales y otros órganos pertinentes deberían elaborar normas internacionales que describan niveles de seguridad y transparencia mensurables y comprobables, de modo que se puedan evaluar objetivamente los sistemas y determinar los niveles de cumplimiento. Además, los Estados Miembros y las empresas deberían apoyar continuamente la investigación estratégica sobre los posibles riesgos de seguridad y protección de las tecnologías de la IA y deberían alentar la investigación sobre la transparencia y la explicabilidad, la inclusión y los conocimientos de esas tecnologías asignando fondos adicionales a esas esferas para diferentes ámbitos y en diferentes niveles, como el lenguaje técnico y natural.
65. Los Estados Miembros deberían aplicar políticas destinadas a garantizar que las acciones de los actores de la IA se ajusten al derecho, las normas y los principios internacionales de derechos humanos durante todo el ciclo de vida de los sistemas de IA, tomando plenamente en consideración al mismo tiempo las diversidades culturales y sociales existentes, en particular las costumbres locales y las tradiciones religiosas, con el debido respeto a la primacía y la universalidad de los derechos humanos.
66. Los Estados Miembros deberían establecer mecanismos para exigir a los actores de la IA que revelen y combatan toda clase de estereotipos en los resultados de los sistemas y datos de IA, ya sean deliberados o por negligencia, y que velen por que los conjuntos de datos de entrenamiento para los sistemas de IA no fomenten las desigualdades culturales, económicas o sociales, los prejuicios, la propagación de desinformación y de información falsa y la vulneración de la libertad de expresión y del acceso a la información. Debería prestarse especial atención a las regiones donde los datos son escasos.
67. Los Estados Miembros deberían aplicar políticas para promover y aumentar la diversidad y la inclusión en los equipos de desarrollo de la IA y los conjuntos de datos de entrenamiento, de modo que estos sean un reflejo de su población, así como para velar por la igualdad de acceso a las tecnologías de la IA y sus beneficios, en particular para los grupos marginados, tanto de zonas rurales como urbanas.
68. Los Estados Miembros deberían elaborar, examinar y adaptar, según proceda, marcos reguladores para alcanzar la rendición de cuentas y la responsabilidad por el contenido y los resultados de los sistemas de IA en las diferentes etapas de su ciclo de vida. Cuando sea necesario, los Estados Miembros deberían introducir marcos de responsabilidad o aclarar la interpretación de los marcos existentes para garantizar la atribución de la responsabilidad por los resultados y el funcionamiento de los sistemas de IA. Además, al elaborar los marcos reguladores, los Estados Miembros deberían tener en cuenta, en particular, que la

responsabilidad y la rendición de cuentas deben recaer siempre en última instancia en personas físicas o jurídicas y que no se debe otorgar personalidad jurídica a los propios sistemas de IA. Para lograrlo, esos marcos reguladores deberían ajustarse al principio de la supervisión humana y establecer un enfoque global centrado en los actores de la IA y los procesos tecnológicos que intervienen en las diferentes etapas del ciclo de vida de los sistemas de IA.

69. A fin de establecer normas cuando no las haya, o de adaptar los marcos jurídicos existentes, los Estados Miembros deberían contar con todos los actores de la IA (incluidos, entre otros, investigadores, representantes de la sociedad civil y de los organismos encargados de hacer cumplir la ley, aseguradores, inversores, fabricantes, ingenieros, abogados y usuarios). Las normas pueden evolucionar hasta convertirse en mejores prácticas, leyes y reglamentos. Se alienta además a los Estados Miembros a que utilicen mecanismos como los prototipos de políticas y los entornos de pruebas reguladores para acelerar la formulación de leyes, reglamentos y políticas, incluidas sus revisiones periódicas, acuerdos con el rápido desarrollo de las nuevas tecnologías y garantizar que las leyes y los reglamentos se puedan poner a prueba en un entorno seguro antes de su aprobación oficial. Los Estados Miembros deberían apoyar a los gobiernos locales en la formulación de políticas, reglamentos y leyes locales que estén en consonancia con los marcos jurídicos nacionales e internacionales.
70. Los Estados Miembros deberían establecer requisitos claros de transparencia y explicabilidad de los sistemas de IA para ayudar a garantizar la fiabilidad de dichos sistemas durante todo su ciclo de vida. Esos requisitos deberían abarcar la concepción y la aplicación de mecanismos de evaluación del impacto que tengan en cuenta la naturaleza del ámbito de aplicación, la utilización prevista, los destinatarios y la viabilidad de cada sistema de IA en particular.

### **ÁMBITO DE ACTUACIÓN 3: POLÍTICA DE DATOS**

71. Los Estados Miembros deberían procurar elaborar estrategias de gobernanza de datos que garanticen la evaluación continua de la calidad de los datos de entrenamiento para los sistemas de IA, en particular la idoneidad de los procesos de recopilación y selección de datos, y que prevean medidas adecuadas de seguridad y protección de los datos, así como mecanismos de retroalimentación para aprender de los errores y compartir las mejores prácticas entre todos los actores de la IA.
72. Los Estados Miembros deberían establecer salvaguardias adecuadas para proteger el derecho a la privacidad, de conformidad con el derecho internacional, en particular respondiendo a preocupaciones como la vigilancia. Los Estados Miembros deberían, entre otras medidas, adoptar o aplicar marcos legislativos que proporcionen una protección adecuada, conforme al derecho internacional. Los Estados Miembros deberían alentar enérgicamente a todos los actores de la IA, incluidas las empresas, a que cumplan las normas internacionales vigentes y, en particular, a que realicen evaluaciones adecuadas del impacto en la privacidad, como parte de las evaluaciones del impacto ético, que tengan en cuenta las repercusiones socioeconómicas más amplias del tratamiento previsto de los datos, y a que apliquen el principio de protección de la privacidad desde la concepción de sus sistemas. La privacidad debería ser respetada, protegida y promovida a lo largo del ciclo de vida de los sistemas de IA.
73. Los Estados Miembros deberían velar por que las personas conserven sus derechos sobre sus datos personales y estén protegidas por un marco que prevea, en particular: la transparencia; salvaguardias adecuadas para el tratamiento de datos sensibles; un nivel adecuado de protección de los datos; planes y mecanismos de rendición de cuentas eficaces y significativos; el pleno disfrute de los derechos de los interesados y la posibilidad de acceder a los datos personales en los sistemas de IA y de borrarlos, excepto en determinadas circunstancias, de conformidad con el derecho internacional; un nivel adecuado de protección

que cumpla plenamente la legislación en materia de protección de datos cuando estos se utilicen con fines comerciales —por ejemplo, para permitir la publicidad con fines específicos— o sean transferidos al extranjero; y una supervisión eficaz e independiente en el marco de un mecanismo de gobernanza de los datos que permita a las personas mantener el control de sus datos personales y promueva los beneficios de la libre circulación de información a nivel internacional, incluido el acceso a los datos.

74. Los Estados Miembros deberían establecer sus políticas de datos o marcos equivalentes, o reforzar las políticas y marcos existentes, para garantizar la seguridad total de los datos personales y los datos sensibles que, de ser divulgados, puedan causar daños, lesiones o dificultades excepcionales a las personas. Cabe citar como ejemplos los datos relativos a infracciones, procesos penales y condenas, así como a las medidas de seguridad conexas; los datos biométricos, genéticos y de salud; y los datos personales como los relativos a la raza, el color, la ascendencia, el género, la edad, el idioma, la religión, las opiniones políticas, el origen nacional, étnico o social, la condición económica o social de nacimiento, la discapacidad o cualquier otra característica.
75. Los Estados Miembros deberían promover los datos abiertos. A este respecto, los Estados Miembros deberían considerar la posibilidad de revisar sus políticas y marcos reguladores, particularmente en lo que respecta al acceso a la información y el gobierno abierto, para reflejar los requisitos específicos de la IA y promover mecanismos, como repositorios abiertos de datos y códigos fuente públicos o de financiación pública y fideicomisos de datos, a fin de apoyar el intercambio seguro, equitativo, legal y ético de datos, entre otros.
76. Los Estados Miembros deberían promover y facilitar el uso de conjuntos de datos sólidos y de calidad para el entrenamiento, el desarrollo y la utilización de los sistemas de IA y ejercer vigilancia al supervisar su recopilación y utilización. Esto podría suponer, si es posible y factible, invertir en la creación de conjuntos de datos de excelencia, incluidos conjuntos de datos abiertos y fiables, que sean diversos, establecidos sobre una base jurídica válida, en particular con el consentimiento de los interesados, cuando así lo exija la ley. Debería alentarse la elaboración de normas para anotar los conjuntos de datos, incluido el desglose de datos por género y otros criterios, de manera que pueda determinarse fácilmente cómo se recopila un conjunto de datos y qué propiedades tiene.
77. Como también se sugiere en el informe del Panel de Alto Nivel sobre la Cooperación Digital del Secretario General de las Naciones Unidas, los Estados Miembros, con el apoyo de las Naciones Unidas y la UNESCO, deberían adoptar un enfoque de patrimonio digital común respecto a los datos, cuando proceda, aumentar la interoperabilidad de los instrumentos y conjuntos de datos, así como las interfaces de los sistemas que albergan datos, y alentar a las empresas del sector privado a que compartan con todas las partes interesadas los datos que recopilan, en beneficio de la investigación, la innovación o el interés público, según proceda. También deberían promover las iniciativas públicas y privadas para crear plataformas de colaboración destinadas a compartir datos de calidad en espacios de datos fiables y seguros.

#### **ÁMBITO DE ACTUACIÓN 4: DESARROLLO Y COOPERACIÓN INTERNACIONAL**

78. Los Estados Miembros y las empresas transnacionales deberían dar prioridad a la ética de la IA, incluyendo debates sobre cuestiones éticas relacionadas con la IA en los foros internacionales, intergubernamentales y de múltiples partes interesadas pertinentes.
79. Los Estados Miembros deberían velar por que la utilización de la IA en esferas relacionadas con el desarrollo, como la educación, la ciencia, la cultura, la comunicación y la información, la atención sanitaria, la agricultura y el suministro de alimentos, el medio ambiente, la gestión de recursos naturales y de infraestructuras y la planificación y el crecimiento económicos, entre otras, respete los valores y principios enunciados en la presente Recomendación.

80. Los Estados Miembros deberían procurar, por conducto de organizaciones internacionales, establecer plataformas de cooperación internacional en el ámbito de la IA para el desarrollo, en particular aportando competencias técnicas, financiación, datos, conocimientos del sector e infraestructura y facilitando la colaboración entre múltiples partes interesadas para hacer frente a los problemas complejos en materia de desarrollo, especialmente para los países de ingreso mediano bajo, en particular los PMA, los PDSL y los PEID.
81. Los Estados Miembros deberían procurar promover la colaboración internacional en materia de investigación e innovación en IA, especialmente en centros y redes de investigación e innovación que promuevan una mayor participación y liderazgo de los investigadores procedentes de países de ingreso mediano bajo y otros países, en particular de PMA, PDSL y PEID.
82. Los Estados Miembros deberían promover las investigaciones sobre la ética de la IA recurriendo a organizaciones internacionales e instituciones de investigación, así como a empresas transnacionales, que puedan servir de base para la utilización ética de los sistemas de IA por entidades públicas y privadas, incluidas las investigaciones sobre la aplicabilidad de marcos éticos específicos en culturas y contextos concretos y sobre las posibilidades de elaborar soluciones tecnológicamente viables de conformidad con esos marcos.
83. Los Estados Miembros deberían alentar la cooperación y la colaboración internacionales en el ámbito de la IA para salvar las divisiones geotecnológicas. Deberían realizarse intercambios y consultas de carácter tecnológico entre los Estados Miembros y su población, entre los sectores público y privado y entre los países más y menos avanzados tecnológicamente, respetando plenamente el derecho internacional.

#### **ÁMBITO DE ACTUACIÓN 5: MEDIO AMBIENTE Y ECOSISTEMAS**

84. Los Estados Miembros y las empresas deberían evaluar el impacto ambiental directo e indirecto de los sistemas de IA a lo largo de su ciclo de vida, en particular, aunque no exclusivamente, su huella de carbono, su consumo de energía y el impacto ambiental de la extracción de las materias primas necesarias para la fabricación de tecnologías de la IA, y reducir el impacto ambiental de los sistemas de IA y las infraestructuras de datos. Los Estados Miembros deberían asegurar el cumplimiento de las leyes, políticas y prácticas ambientales por parte de todos los actores de la IA.
85. Los Estados Miembros deberían establecer incentivos, cuando sea necesario y apropiado, para garantizar la elaboración y adopción de soluciones basadas en los derechos y en la ética de la IA en favor de la resiliencia ante el riesgo de desastres; la vigilancia, protección y regeneración del medio ambiente y los ecosistemas; y la preservación del planeta. Esos sistemas de IA deberían contar, durante todo su ciclo de vida, con la participación de las comunidades locales e indígenas y apoyar enfoques del tipo de economía circular y modalidades de consumo y producción sostenibles. Por ejemplo, los sistemas de IA podrían utilizarse, cuando sea necesario y apropiado, para:
  - a) apoyar la protección, el seguimiento y la gestión de los recursos naturales;
  - b) apoyar la predicción, la prevención, el control y la atenuación de los problemas relacionados con el clima;
  - c) favorecer un ecosistema alimentario más eficiente y sostenible;
  - d) contribuir a acelerar el acceso a la energía sostenible y su adopción a gran escala;
  - e) facilitar y promover la incorporación de infraestructuras sostenibles, modelos empresariales sostenibles y financiación sostenible al servicio del desarrollo sostenible;

f) detectar los contaminantes o prever los niveles de contaminación y, de ese modo, ayudar a las partes interesadas pertinentes a definir, planificar y poner en marcha intervenciones específicas para prevenir y reducir la contaminación y la exposición.

86. Al elegir un método de IA, dada la gran necesidad potencial de datos o recursos de algunos de ellos y su consiguiente impacto en el medio ambiente, los Estados Miembros deberían velar por que los actores de la IA, de conformidad con el principio de proporcionalidad, favorezcan los métodos de IA eficientes en cuanto a datos, energía y recursos. Deberían establecerse requisitos destinados a garantizar que se disponga de pruebas adecuadas para demostrar que una aplicación de IA tendrá el efecto deseado, o que dicha aplicación cuenta con garantías que permitan justificar su utilización. Si esto no es posible, deberá privilegiarse el principio de precaución y, en los casos en que haya impactos negativos desproporcionados en el medio ambiente, no debería utilizarse la IA.

## **ÁMBITO DE ACTUACIÓN 6: GÉNERO**

87. Los Estados Miembros deberían velar por que se optimice plenamente el potencial de las tecnologías digitales y la inteligencia artificial para contribuir a lograr la igualdad de género, y han de asegurarse de que no se conculquen los derechos humanos y las libertades fundamentales de las niñas y las mujeres, ni su seguridad e integridad, en ninguna etapa del ciclo de vida de los sistemas de IA. Además, la evaluación del impacto ético debería incluir una perspectiva transversal de género.

88. Los Estados Miembros deberían asignar fondos específicos de sus presupuestos públicos a la financiación de planes con perspectiva de género, velar por que las políticas digitales nacionales incluyan un plan de acción en materia de género y elaborar políticas pertinentes, por ejemplo, sobre la educación laboral, destinadas a apoyar a las niñas y las mujeres para que no queden excluidas de la economía digital impulsada por la IA. Deberían considerarse y realizarse inversiones especiales para ofrecer programas específicos y un lenguaje que respete la igualdad de género, a fin de ampliar las oportunidades de participación de las niñas y las mujeres en la ciencia, la tecnología, la ingeniería y las matemáticas (CTIM), incluidas las disciplinas relacionadas con las tecnologías de la información y la comunicación (TIC), y mejorar la preparación, la empleabilidad, el desarrollo de las perspectivas de carrera en condiciones de igualdad y el crecimiento profesional de las niñas y las mujeres.

89. Los Estados Miembros deberían velar por que se aproveche el potencial de los sistemas de IA para impulsar el logro de la igualdad de género. Deberían asegurarse de que estas tecnologías no exacerben las ya amplias brechas que existen entre los géneros en varios ámbitos del mundo analógico, sino que, al contrario, las eliminen. Entre estas brechas cabe citar la disparidad salarial entre hombres y mujeres; su representación desigual en ciertas profesiones y actividades; la falta de representación en los puestos directivos superiores, las juntas directivas o los equipos de investigación en el campo de la IA; la brecha educativa; las desigualdades en el acceso, la adopción, la utilización y la asequibilidad de la tecnología digital y de la IA; y la distribución desigual del trabajo no remunerado y de las responsabilidades de cuidado en nuestras sociedades.

90. Los Estados Miembros deberían velar por que los estereotipos de género y los sesgos discriminatorios no se trasladen a los sistemas de IA, sino que se detecten y corrijan de manera proactiva. Es preciso esforzarse por evitar el efecto negativo combinado de las brechas tecnológicas para lograr la igualdad de género y prevenir la violencia contra las niñas y las mujeres, así como contra los grupos insuficientemente representados, que puede manifestarse en forma de hostigamiento, acoso o trata, incluso en línea.

91. Los Estados Miembros deberían alentar la iniciativa empresarial, la participación y el compromiso de las mujeres en todas las etapas del ciclo de vida de los sistemas de IA, ofreciendo y promoviendo incentivos económicos y reguladores, entre otros estímulos y

planes de apoyo, así como políticas encaminadas a lograr una participación equilibrada de hombres y mujeres en la investigación en IA en las universidades y la representación de género en los puestos directivos superiores, las juntas directivas y los equipos de investigación de las empresas digitales y de IA. Los Estados Miembros deberían garantizar que los fondos públicos (para innovación, investigación y tecnologías) se destinen a programas y empresas inclusivos, con un claro equilibrio de género, y que se alienten igualmente los fondos privados mediante principios de acción afirmativa. Deberían elaborarse y aplicarse políticas que favorezcan los entornos libres de acoso, alentando al mismo tiempo la transferencia de las mejores prácticas sobre la forma de promover la diversidad durante el ciclo de vida de los sistemas de IA.

92. Los Estados Miembros deberían promover la diversidad de género en la investigación en IA en el mundo universitario y la industria, ofreciendo incentivos a las niñas y las mujeres para que se incorporen a este campo, estableciendo mecanismos para luchar contra los estereotipos de género y el acoso en la comunidad de investigadores en IA y alentando a las entidades universitarias y privadas a que compartan las mejores prácticas sobre la forma de potenciar la diversidad de género.
93. La UNESCO puede ayudar a compilar un repositorio de mejores prácticas para incentivar la participación de las niñas, las mujeres y los grupos insuficientemente representados en todas las etapas del ciclo de vida de los sistemas de IA.

#### **ÁMBITO DE ACTUACIÓN 7: CULTURA**

94. Se alienta a los Estados Miembros a que incorporen sistemas de IA, cuando proceda, a la preservación, el enriquecimiento, la comprensión, la promoción, la gestión y la accesibilidad del patrimonio cultural material, documental e inmaterial, incluidas las lenguas en peligro y las lenguas y conocimientos indígenas, por ejemplo, introduciendo o actualizando programas educativos relacionados con la aplicación de los sistemas de IA en esas esferas, cuando sea apropiado, y asegurando un enfoque participativo, dirigidos a las instituciones y al público.
95. Se alienta a los Estados Miembros a que examinen y aborden las repercusiones de los sistemas de IA en la cultura, especialmente de las aplicaciones de procesamiento del lenguaje natural (PLN), como la traducción automática y los asistentes de voz, en los matices del lenguaje y la expresión humanos. Esas evaluaciones deberían contribuir a la elaboración y aplicación de estrategias que maximicen los beneficios de esos sistemas reduciendo las desigualdades culturales y mejorando la comprensión humana, así como haciendo frente a las repercusiones negativas, como una menor utilización, que podría conducir a la desaparición de lenguas en peligro, dialectos locales y variaciones tonales y culturales asociadas con el lenguaje y la expresión humanos.
96. Los Estados Miembros deberían promover la educación en IA y la capacitación digital de los artistas y los profesionales creativos, a fin de que puedan evaluar la idoneidad de las tecnologías de la IA para utilizarlas en su profesión y contribuir a la concepción y aplicación de tecnologías de la IA adecuadas, ya que estas tecnologías se emplean actualmente para crear, producir, distribuir, difundir y consumir una variedad de bienes y servicios culturales, teniendo en cuenta la importancia de preservar el patrimonio cultural, la diversidad y la libertad artística.
97. Los Estados Miembros deberían promover el conocimiento y la evaluación de los instrumentos de IA entre las industrias culturales locales y las pequeñas y medianas empresas que trabajan en el ámbito de la cultura, a fin de evitar el riesgo de concentración en el mercado cultural.
98. Los Estados Miembros deberían contar con las empresas tecnológicas y otras partes interesadas para promover una oferta diversa de expresiones culturales y un acceso plural a

ellas y, en particular, para garantizar que la recomendación algorítmica aumente la notoriedad de los contenidos locales y la posibilidad de descubrirlos.

99. Los Estados Miembros deberían promover nuevas investigaciones en la intersección entre la IA y la propiedad intelectual, por ejemplo, para determinar si hay que proteger con derechos de propiedad intelectual las obras creadas mediante tecnologías de la IA y la manera de hacerlo. Los Estados Miembros también deberían evaluar cómo afectan las tecnologías de la IA a los derechos o los intereses de los titulares de derechos de propiedad intelectual cuyas obras se utilizan para investigar, desarrollar, entrenar o implantar aplicaciones de IA.
100. Los Estados Miembros deberían alentar a los museos, las galerías, las bibliotecas y los archivos de ámbito nacional a que utilicen sistemas de IA para realzar sus colecciones y enriquecer sus bibliotecas, bases de datos y bases de conocimientos, ofreciendo al mismo tiempo acceso a sus usuarios.

### **ÁMBITO DE ACTUACIÓN 8: EDUCACIÓN E INVESTIGACIÓN**

101. Los Estados Miembros deberían colaborar con organizaciones internacionales, instituciones educativas y entidades privadas y no gubernamentales para impartir al público de todos los países, a todos los niveles, conocimientos adecuados en materia de IA, a fin de empoderar a la población y reducir las brechas digitales y las desigualdades en el acceso a la tecnología digital resultantes de la adopción a gran escala de sistemas de IA.
102. Los Estados Miembros deberían promover la adquisición de “competencias previas” para la educación en materia de IA, como la alfabetización básica, la aritmética elemental, las competencias digitales y de codificación y la alfabetización mediática e informacional, así como el pensamiento crítico y creativo, el trabajo en equipo, la comunicación, las aptitudes socioemocionales y las competencias en materia de ética de la IA, especialmente en los países y en las regiones o zonas dentro de los países en que existen lagunas notables en la enseñanza de esas competencias.
103. Los Estados Miembros deberían promover programas generales de sensibilización sobre los avances de la IA, en particular sobre los datos y las oportunidades que ofrecen y los retos que plantean las tecnologías de la IA, el impacto de los sistemas de IA en los derechos humanos, incluidos los derechos de los niños, y sus repercusiones. Estos programas deberían ser accesibles tanto a los grupos técnicos como a los no técnicos.
104. Los Estados Miembros deberían alentar las iniciativas de investigación sobre la utilización responsable y ética de las tecnologías de la IA en la enseñanza, la formación de docentes y el aprendizaje electrónico, entre otras cuestiones, a fin de aumentar las oportunidades y atenuar los problemas y los riesgos existentes en este ámbito. Esas iniciativas deberían ir acompañadas de una evaluación adecuada de la calidad de la educación y de las repercusiones que la utilización de las tecnologías de la IA tiene para los educandos y los docentes. Los Estados Miembros deberían también velar por que las tecnologías de la IA empoderen a los educandos y los docentes y mejoren su experiencia, teniendo presente que los aspectos relacionales y sociales y el valor de las formas tradicionales de educación son fundamentales en las relaciones entre docentes y educandos y entre los propios educandos y deberían tenerse en cuenta al examinar la adopción de las tecnologías de la IA en la educación. Los sistemas de IA utilizados en el aprendizaje deberían estar sujetos a requisitos estrictos en materia de supervisión, evaluación de las capacidades o predicción de los comportamientos de los educandos. La IA debería apoyar el proceso de aprendizaje sin reducir las capacidades cognitivas y sin recabar información sensible, respetando las normas pertinentes en materia de protección de los datos personales. Los datos facilitados para adquirir conocimientos, recopilados durante las interacciones del educando con el sistema de IA, no deben ser objeto de uso indebido, apropiación indebida o explotación delictiva, incluidos los fines comerciales.

105. Los Estados Miembros deberían promover la participación y el liderazgo de las niñas y las mujeres, las personas de diversos orígenes étnicos y culturas, las personas con discapacidad, las personas marginadas y vulnerables o en situación de vulnerabilidad y las minorías, así como de todas aquellas personas que no gocen plenamente de los beneficios de la inclusión digital, en los programas de educación en materia de IA en todos los niveles, así como el seguimiento y el intercambio con otros Estados Miembros de las mejores prácticas en este ámbito.
106. Los Estados Miembros deberían elaborar, de conformidad con sus tradiciones y programas de educación nacionales, planes de estudios sobre la ética de la IA para todos los niveles y promover la colaboración cruzada entre la enseñanza de competencias técnicas de IA y los aspectos humanísticos, éticos y sociales de la educación en IA. Deberían elaborarse cursos en línea y recursos digitales de enseñanza de la ética de la IA en las lenguas locales, incluidas las lenguas indígenas, y tenerse en cuenta la diversidad de los entornos, velando especialmente por la accesibilidad de los formatos para las personas con discapacidad.
107. Los Estados Miembros deberían promover y apoyar las investigaciones sobre la IA, en particular las investigaciones sobre la ética de la IA, por ejemplo mediante inversiones en este tipo de investigaciones o creando incentivos para que los sectores público y privado inviertan en este ámbito, reconociendo que las investigaciones contribuyen de manera significativa al desarrollo y la mejora posteriores de las tecnologías de la IA con miras a promover el derecho internacional y los valores y principios enunciados en la presente Recomendación. Los Estados Miembros también deberían promover públicamente las mejores prácticas de los investigadores y las empresas que desarrollan la IA de forma ética y cooperar con ellos.
108. Los Estados Miembros deberían velar por que los investigadores en IA reciban formación en ética de la investigación y exigirles que tengan en cuenta consideraciones éticas en sus concepciones, productos y publicaciones, especialmente en los análisis de los conjuntos de datos que utilizan, la forma en que estos se anotan y la calidad y el alcance de los resultados, así como las posibles aplicaciones.
109. Los Estados Miembros deberían alentar a las empresas del sector privado a que faciliten el acceso de la comunidad científica a sus datos para la investigación, especialmente en los países de ingreso mediano bajo, en particular en los PMA, los PDSL y los PEID. Este acceso debería ser conforme a las normas pertinentes en materia de respeto de la privacidad y protección de datos.
110. Para asegurar una evaluación crítica de las investigaciones en IA y un seguimiento adecuado de los posibles usos indebidos o efectos adversos, los Estados Miembros deberían velar por que cualquier evolución futura relacionada con las tecnologías de la IA se base en investigaciones científicas rigurosas e independientes y promover la investigación interdisciplinaria en IA mediante la inclusión de disciplinas distintas de la ciencia, la tecnología, la ingeniería y las matemáticas (CTIM), como los estudios culturales, la educación, la ética, las relaciones internacionales, el derecho, la lingüística, la filosofía, las ciencias políticas, la sociología y la psicología.
111. Reconociendo que las tecnologías de la IA ofrecen grandes oportunidades para contribuir al avance de los conocimientos y las prácticas científicos, especialmente en las disciplinas tradicionalmente basadas en modelos, los Estados Miembros deberían alentar a las comunidades científicas a ser conscientes de los beneficios, los límites y los riesgos de su utilización, por ejemplo, intentando garantizar que las conclusiones extraídas de los enfoques, modelos y tratamientos basados en los datos sean sólidas y consistentes. Además, los Estados Miembros deberían celebrar y apoyar la función de la comunidad científica en la contribución a las políticas y en la concienciación respecto de los puntos fuertes y débiles de las tecnologías de la IA.

## **ÁMBITO DE ACTUACIÓN 9: COMUNICACIÓN E INFORMACIÓN**

112. Los Estados Miembros deberían utilizar los sistemas de IA para mejorar el acceso a la información y el conocimiento. Para ello, pueden, por ejemplo, apoyar a los investigadores, las universidades, los periodistas, el público en general y los desarrolladores, a fin de mejorar la libertad de expresión, las libertades académicas y científicas y el acceso a la información, así como aumentar la divulgación proactiva de los datos y la información oficiales.
113. Los Estados Miembros deberían garantizar que los actores de la IA respeten y promuevan la libertad de expresión y el acceso a la información en lo que respecta a la generación, moderación y conservación automáticas de contenidos. Marcos adecuados, incluso reglamentarios, deberían propiciar la transparencia de los operadores de comunicación e información en línea, velar por que los usuarios tengan acceso a puntos de vista diversos y prever procesos de notificación rápida a los usuarios sobre los motivos de la eliminación u otro tratamiento de los contenidos, así como mecanismos de recurso que permitan a los usuarios solicitar reparación.
114. Los Estados Miembros deberían invertir en competencias digitales y de alfabetización mediática e informacional y promoverlas, a fin de reforzar el pensamiento crítico y las competencias necesarias para comprender el uso y las implicaciones de los sistemas de IA, con miras a atenuar y contrarrestar la desinformación, la información errónea y el discurso de odio. Una mejor comprensión y evaluación de los efectos tanto positivos como potencialmente perjudiciales de los sistemas de recomendación debería formar parte de esos esfuerzos.
115. Los Estados Miembros deberían crear entornos propicios para que los medios de comunicación tengan los derechos y recursos necesarios para informar eficazmente sobre las ventajas y los inconvenientes de los sistemas de IA, y también alentar a los medios de comunicación a que hagan un uso ético de estos sistemas en su trabajo.

## **ÁMBITO DE ACTUACIÓN 10: ECONOMÍA Y TRABAJO**

116. Los Estados Miembros deberían evaluar y abordar el impacto de los sistemas de IA en los mercados de trabajo y sus consecuencias en las necesidades educativas en todos los países y, más concretamente, en los países cuya economía requiere mucha mano de obra. Para ello puede ser preciso introducir una gama más amplia de competencias “básicas” e interdisciplinarias en todos los niveles educativos, a fin de dar a los trabajadores actuales y a las nuevas generaciones una oportunidad equitativa de encontrar empleo en un mercado en rápida evolución y para asegurar que sean conscientes de los aspectos éticos de los sistemas de IA. Junto a las competencias técnicas especializadas, así como a las tareas poco especializadas, deberían enseñarse competencias como “aprender a aprender”, comunicación, pensamiento crítico, trabajo en equipo, empatía y la capacidad de transferir los conocimientos propios a diversos ámbitos. Es fundamental actuar con transparencia con respecto a las competencias de las que existe demanda y actualizar los planes de estudios en torno a ellas.
117. Los Estados Miembros deberían apoyar los acuerdos de colaboración entre los gobiernos, las instituciones universitarias, las instituciones de enseñanza y formación profesional, la industria, las organizaciones de trabajadores y la sociedad civil a fin de reducir la brecha en cuanto a las competencias exigidas para adecuar los programas y estrategias de capacitación a las futuras implicaciones del trabajo y a las necesidades de la industria, incluidas las pequeñas y medianas empresas. Deberían promoverse enfoques de enseñanza y aprendizaje de la IA basados en proyectos, facilitando las asociaciones de colaboración entre las instituciones públicas, las empresas del sector privado, las universidades y los centros de investigación.
118. Los Estados Miembros deberían colaborar con empresas del sector privado, organizaciones de la sociedad civil y otras partes interesadas, incluidos trabajadores y sindicatos, para

garantizar una transición equitativa a los empleados en situación de riesgo. Esto supone poner en marcha programas de perfeccionamiento y reconversión profesional, encontrar mecanismos eficaces para retener a los empleados durante esos periodos de transición y explorar programas de protección social para aquellos que no puedan reconvertirse. Los Estados Miembros deberían elaborar y aplicar programas para analizar los problemas detectados y darles respuesta, entre los que podrían figurar el perfeccionamiento y la reconversión profesional, el fortalecimiento de la protección social, la aplicación de políticas e intervenciones sectoriales proactivas y la introducción de ventajas fiscales y nuevas formas de tributación. Los Estados Miembros deberían garantizar que haya suficiente financiación pública para apoyar estos programas. Las reglamentaciones pertinentes, como los regímenes fiscales, deberían examinarse cuidadosamente y modificarse, si es necesario, para contrarrestar las consecuencias del desempleo causado por la automatización basada en la IA.

119. Los Estados Miembros deberían alentar y apoyar a los investigadores para que analicen el impacto de los sistemas de IA en el entorno laboral local con miras a anticipar las tendencias y los desafíos futuros. Estos estudios deberían tener un enfoque interdisciplinario y examinar el impacto de los sistemas de IA en los sectores económico, social y geográfico, así como en las interacciones entre seres humanos y robots y entre los propios seres humanos, a fin de asesorar sobre las mejores prácticas de reconversión y reasignación profesionales.
120. Los Estados Miembros deberían adoptar las medidas adecuadas para garantizar la competitividad de los mercados y la protección de los consumidores, considerando posibles medidas y mecanismos en los planos nacional, regional e internacional, a fin de impedir los abusos de posición dominante en el mercado, incluidos los monopolios, en relación con los sistemas de IA durante su ciclo de vida, ya se trate de datos, investigación, tecnología o mercados. Los Estados Miembros deberían prevenir las desigualdades resultantes, evaluar los mercados correspondientes y promover mercados competitivos. Se debería prestar la debida atención a los países de ingreso mediano bajo, en particular a los PMA, los PDSL y los PEID, que están más expuestos y son más vulnerables a la posibilidad de que se produzcan abusos de posición dominante en el mercado, como consecuencia de la falta de infraestructuras, capacidad humana y reglamentación, entre otros factores. Los actores de la IA que desarrollen sistemas de IA en países que hayan establecido o adoptado normas éticas en materia de IA deberían respetar estas normas cuando exporten estos productos, desarrollen sus sistemas de IA o los apliquen en países donde no existan dichas normas, respetando al mismo tiempo el derecho internacional y las leyes, normas y prácticas nacionales aplicables de estos países.

## **ÁMBITO DE ACTUACIÓN 11: SALUD Y BIENESTAR SOCIAL**

121. Los Estados Miembros deberían esforzarse por emplear sistemas eficaces de IA para mejorar la salud humana y proteger el derecho a la vida, en particular atenuando los brotes de enfermedades, al tiempo que desarrollan y mantienen la solidaridad internacional para hacer frente a los riesgos e incertidumbres relacionados con la salud en el plano mundial, y garantizar que su despliegue de sistemas de IA en el ámbito de la atención de la salud sea conforme al derecho internacional y a sus obligaciones en materia de derechos humanos. Los Estados Miembros deberían velar por que los actores que participan en los sistemas de IA relacionados con la atención de la salud tengan en cuenta la importancia de las relaciones del paciente con su familia y con el personal sanitario.
122. Los Estados Miembros deberían garantizar que el desarrollo y el despliegue de los sistemas de IA relacionados con la salud en general y con la salud mental en particular —prestando la debida atención a los niños y los jóvenes— estén regulados, de modo que esos sistemas sean seguros, eficaces, eficientes y probados desde el punto de vista científico y médico y faciliten la innovación y el progreso médico con base empírica. Además, en el ámbito conexo de las intervenciones de salud digital, se alienta encarecidamente a los Estados Miembros a que

hagan participar activamente a los pacientes y sus representantes en todas las etapas pertinentes del desarrollo del sistema.

123. Los Estados Miembros deberían prestar particular atención a la regulación de las soluciones de predicción, detección y tratamiento médicos en las aplicaciones de la IA, mediante:
  - a) la supervisión para minimizar y atenuar los sesgos;
  - b) la inclusión del profesional, el paciente, el cuidador o el usuario del servicio en el equipo en calidad de "experto en la materia" en todas las etapas pertinentes al elaborar los algoritmos;
  - c) una debida atención a la privacidad, dado que quizá sea necesaria una vigilancia médica, y el cumplimiento de todos los requisitos nacionales e internacionales pertinentes en materia de protección de datos;
  - d) mecanismos eficaces para que las personas cuyos datos personales se están analizando sepan de la utilización y el análisis de sus datos y den su consentimiento informado al respecto, sin impedir el acceso a la atención de la salud;
  - e) la garantía de que el cuidado humano y la decisión final sobre el diagnóstico y el tratamiento correspondan siempre a seres humanos, reconociendo al mismo tiempo que los sistemas de IA también pueden ayudarlos en su trabajo;
  - f) el examen, cuando sea necesario, de los sistemas de IA por un comité de investigación ética antes de su uso clínico.
124. Los Estados Miembros deberían realizar investigaciones sobre los efectos y la regulación de los posibles daños de los sistemas de IA para la salud mental, tales como un aumento de la depresión, la ansiedad, el aislamiento social, el desarrollo de adicciones, el tráfico, la radicalización y la información errónea, entre otros.
125. Los Estados Miembros deberían elaborar directrices sobre las interacciones entre seres humanos y robots y sus repercusiones en las relaciones entre seres humanos, basadas en la investigación y orientadas al desarrollo futuro de robots, y prestando especial atención a la salud mental y física de los seres humanos. Debería prestarse particular atención al uso de robots en la atención de la salud, en la atención a las personas de edad y las personas con discapacidad y en el ámbito de la educación, así como a los robots para uso infantil y para usos lúdicos, conversacionales y de compañía para niños y adultos. Además, deberían utilizarse las tecnologías de la IA para mejorar la seguridad y el uso ergonómico de los robots, en particular en entornos de trabajo en los que intervienen robots y seres humanos. Debería prestarse especial atención a la posibilidad de utilizar la IA para manipular los sesgos cognitivos humanos y hacer un mal uso de ellos.
126. Los Estados Miembros deberían velar por que las interacciones entre seres humanos y robots se ajusten a los mismos valores y principios que se aplican a cualquier otro sistema de IA, lo que incluye los derechos humanos y las libertades fundamentales, la promoción de la diversidad y la protección de las personas vulnerables o en situación de vulnerabilidad. Las cuestiones éticas relativas a los sistemas basados en la IA utilizados en las neurotecnologías y las interfaces cerebro-ordenador deberían tenerse en cuenta a fin de preservar la dignidad y la autonomía humanas.
127. Los Estados Miembros deberían velar por que los usuarios puedan determinar fácilmente si interactúan con un ser vivo o con un sistema de IA que imita las características humanas o animales y puedan rechazar eficazmente dicha interacción y solicitar la intervención humana.

128. Los Estados Miembros deberían aplicar políticas de sensibilización sobre la antropomorfización de las tecnologías de la IA y las tecnologías que reconocen e imitan las emociones humanas, especialmente en el lenguaje utilizado para referirse a ellas, y evaluar las manifestaciones, las implicaciones éticas y las posibles limitaciones de esa antropomorfización, en particular en el contexto de la interacción entre robots y seres humanos y, especialmente, cuando se trate de niños.
129. Los Estados Miembros deberían alentar y promover la investigación colaborativa sobre los efectos de la interacción a largo plazo de las personas con los sistemas de IA, prestando especial atención a las consecuencias psicológicas y cognitivas que estos sistemas pueden tener en los niños y los jóvenes. Para ello deberían utilizarse múltiples normas, principios, protocolos, enfoques disciplinarios y un análisis de la modificación de las conductas y los hábitos, así como una cuidadosa evaluación de los impactos culturales y sociales posteriores. Además, los Estados Miembros deberían alentar la investigación sobre el efecto de las tecnologías de la IA en el desempeño del sistema sanitario y los resultados en materia de salud.
130. Los Estados Miembros, así como todas las partes interesadas, deberían establecer mecanismos para hacer participar de manera significativa a los niños y los jóvenes en las conversaciones, los debates y la adopción de decisiones sobre las repercusiones de los sistemas de IA en sus vidas y su futuro.

## **V. SEGUIMIENTO Y EVALUACIÓN**

131. Los Estados Miembros, de acuerdo con sus circunstancias, estructuras de gobierno y disposiciones constitucionales específicas, deberían velar por el seguimiento y la evaluación de las políticas, los programas y los mecanismos relativos a la ética de la IA de forma creíble y transparente mediante una combinación de enfoques cuantitativos y cualitativos. La UNESCO puede contribuir a apoyar a los Estados Miembros mediante:
  - a) la elaboración de una metodología de la UNESCO de evaluación del impacto ético de las tecnologías de la IA basada en una investigación científica rigurosa y fundamentada en el derecho internacional de los derechos humanos, orientaciones para su aplicación en todas las etapas del ciclo de vida de los sistemas de IA y materiales de desarrollo de las capacidades para apoyar la labor de los Estados Miembros dirigida a capacitar a funcionarios públicos, responsables de formular políticas y otros actores pertinentes de la IA en la metodología de la evaluación del impacto ético;
  - b) la elaboración de una metodología de la UNESCO de evaluación del estadio de preparación para ayudar a los Estados Miembros a determinar su situación en momentos concretos de su trayectoria de preparación a través de un conjunto de dimensiones;
  - c) la elaboración de una metodología de la UNESCO para evaluar *ex ante* y *ex post* la eficacia y la eficiencia de las políticas y los incentivos relacionados con la ética de la IA con respecto a objetivos definidos;
  - d) el fortalecimiento del análisis, basado en investigaciones y pruebas, de las políticas relativas a la ética de la IA y la presentación de informes a este respecto;
  - e) la recopilación y difusión de información sobre los avances y las innovaciones, informes de investigación, publicaciones científicas, datos y estadísticas relativos a las políticas sobre la ética de la IA, en particular mediante las iniciativas existentes, a fin de apoyar el intercambio de mejores prácticas y el aprendizaje mutuo y de impulsar la aplicación de la presente Recomendación.

132. Los procesos de seguimiento y evaluación deberían asegurar una amplia participación de todas las partes interesadas, entre ellas, aunque no exclusivamente, las personas vulnerables o en situación de vulnerabilidad. Se debería garantizar la diversidad social, cultural y de género, con miras a mejorar los procesos de aprendizaje y fortalecer los nexos entre las conclusiones, la adopción de decisiones, la transparencia y la rendición de cuentas sobre los resultados.
133. A fin de promover las mejores políticas y prácticas relacionadas con la ética de la IA, deberían elaborarse instrumentos e indicadores adecuados para evaluar su eficacia y eficiencia en función de normas, prioridades y objetivos acordados, incluidos objetivos específicos para las personas pertenecientes a poblaciones desfavorecidas y marginadas y personas vulnerables o en situación de vulnerabilidad, así como el impacto de los sistemas de IA en los planos individual y social. El seguimiento y la evaluación del impacto de los sistemas de IA y de las políticas y prácticas conexas relativas a la ética de la IA deberían realizarse de manera continua y sistemática proporcionalmente a los riesgos correspondientes. Este proceso debería basarse en marcos acordados internacionalmente e ir acompañado de evaluaciones de instituciones, proveedores y programas públicos y privados, incluidas autoevaluaciones, así como de estudios de seguimiento y la elaboración de conjuntos de indicadores. La recopilación y el procesamiento de datos deberían realizarse de conformidad con el derecho internacional, la legislación nacional en materia de protección y confidencialidad de datos y los valores y principios enunciados en la presente Recomendación.
134. En particular, los Estados Miembros podrían considerar posibles mecanismos de seguimiento y evaluación, como una comisión de ética, un observatorio de ética de la IA, un repositorio que abarque el desarrollo ético y conforme a los derechos humanos de los sistemas de IA, o contribuciones a las iniciativas existentes para reforzar la conformidad a los principios éticos en todas las esferas de competencia de la UNESCO, un mecanismo de intercambio de experiencias, entornos de pruebas reguladores de la IA y una guía de evaluación para que todos los actores de la IA determinen en qué medida cumplen las recomendaciones de actuación mencionadas en el presente documento.

## **VI. UTILIZACIÓN Y APLICACIÓN DE LA PRESENTE RECOMENDACIÓN**

135. Los Estados Miembros y todas las demás partes interesadas que se indican en la presente Recomendación deberían respetar, promover y proteger los valores, principios y normas éticos relativos a la IA que se establecen en esta Recomendación y adoptar todas las medidas posibles para dar efecto a sus recomendaciones de actuación.
136. Los Estados Miembros deberían esforzarse por ampliar y complementar su propia acción en lo que respecta a la presente Recomendación, cooperando con todas las organizaciones gubernamentales y no gubernamentales nacionales e internacionales pertinentes, así como con las empresas transnacionales y las organizaciones científicas, cuyas actividades correspondan al ámbito de aplicación y a los objetivos de la presente Recomendación. La elaboración de una metodología de la UNESCO de evaluación del impacto ético y el establecimiento de comisiones nacionales de ética de la IA pueden ser instrumentos importantes a este respecto.

## **VII. PROMOCIÓN DE LA PRESENTE RECOMENDACIÓN**

137. La UNESCO tiene la vocación de ser el principal organismo de las Naciones Unidas encargado de promover y difundir la presente Recomendación y, en consecuencia, trabajará en colaboración con otras entidades pertinentes de las Naciones Unidas, respetando al mismo tiempo su mandato y evitando la duplicación de esfuerzos.
138. La UNESCO, incluidos algunos de sus órganos, como la Comisión Mundial de Ética del Conocimiento Científico y la Tecnología (COMEST), el Comité Internacional de Bioética (CIB)

y el Comité Intergubernamental de Bioética (CIGB), trabajará también en colaboración con otras organizaciones gubernamentales y no gubernamentales internacionales, regionales y subregionales.

139. Aunque en la UNESCO el mandato de promover y proteger es competencia de los gobiernos y los organismos intergubernamentales, la sociedad civil desempeñará una función importante en la defensa de los intereses del sector público y, por lo tanto, la UNESCO debe garantizar y promover su legitimidad.

## **VIII. DISPOSICIONES FINALES**

140. La presente Recomendación debe entenderse como un todo, y los valores y principios fundamentales deben considerarse complementarios y relacionados entre sí.
141. Ninguna disposición de la presente Recomendación podrá interpretarse como si reemplazara, modificara o menoscabara de cualquier otra manera las obligaciones o los derechos de los Estados con arreglo al derecho internacional o como si autorizara a un Estado, otro agente político, económico o social, grupo o individuo a emprender actividades o realizar actos que sean contrarios a los derechos humanos, las libertades fundamentales, la dignidad humana y el respeto del medio ambiente y los ecosistemas, tanto vivos como no vivos.