



United Nations
Educational, Scientific and
Cultural Organization

Organisation
des Nations Unies
pour l'éducation,
la science et la culture

Organización
de las Naciones Unidas
para la Educación,
la Ciencia y la Cultura

Организация
Объединенных Наций по
вопросам образования,
науки и культуры

منظمة الأمم المتحدة
للتربية والعلم والثقافة

联合国教育、
科学及文化组织

Reunión intergubernamental de expertos (categoría II) relativa al proyecto de recomendación sobre la ética de la inteligencia artificial

En línea, 26-30 de abril de 2021

Distribución: limitada

SHS/IGM-AIETHICS/2021/APR/INF.1

31 de marzo de 2021

Original: inglés y francés

INFORME FINAL SOBRE EL PROYECTO DE RECOMENDACIÓN SOBRE LA ÉTICA DE LA INTELIGENCIA ARTIFICIAL

De conformidad con lo establecido en la Constitución de la UNESCO y en el Reglamento sobre las recomendaciones a los Estados Miembros y las convenciones internacionales previstas en el párrafo 4 del Artículo IV de la Constitución, el informe definitivo con el proyecto de Recomendación sobre la ética de la inteligencia artificial se envió a los Estados Miembros de la UNESCO en marzo de 2021 ([CL4347](#)). Se presentará a un comité especial compuesto por expertos técnicos y jurídicos nombrados por los Estados Miembros el cual se reunirá en dos fases, la primera del 26 al 30 de abril de 2021, y la segunda del 21 al 25 de junio 2021, tal y como comunicado en la carta circular ([CL4337](#)) enviada en enero de 2021.

INFORME FINAL RELATIVO AL PROYECTO DE RECOMENDACIÓN SOBRE LA ÉTICA DE LA INTELIGENCIA ARTIFICIAL

RESUMEN

El presente documento se preparó en cumplimiento de la resolución 40 C/37 de la Conferencia General de la UNESCO, en la que se encargaba a la Directora General que preparara un instrumento normativo internacional sobre la ética de la inteligencia artificial (IA) en forma de recomendación, que se presentaría a la Conferencia General, en su 41^a reunión, en 2021.

El informe final se presenta como seguimiento del “Estudio preliminar sobre la ética de la inteligencia artificial” elaborado por la Comisión Mundial de Ética del Conocimiento Científico (COMEST), así como del estudio preliminar sobre los aspectos técnicos y jurídicos relativos a la conveniencia de disponer de un instrumento normativo sobre la ética de la inteligencia artificial (IA) (documentos [206 EX/42](#) y [40 C/67](#)), que fueron examinados por el Consejo Ejecutivo y la Conferencia General de la UNESCO de conformidad con el artículo 3 del [Reglamento sobre las recomendaciones a los Estados Miembros y las convenciones internacionales previstas en el párrafo 4 del artículo IV de la Constitución](#). El informe final también se basa en el informe preliminar de la Directora General que se transmitió a los Estados Miembros en septiembre de 2020, junto con la primera versión revisada del proyecto de recomendación sobre la ética de la inteligencia artificial, preparada por el Grupo Especial de Expertos (GEE) sobre la base de amplias consultas con múltiples partes interesadas.

El informe final, que contiene un proyecto de recomendación sobre la ética de la inteligencia artificial, se elaboró de conformidad con el párrafo 3 del artículo 10 del Reglamento. Plasma los comentarios y observaciones recibidos de los Estados Miembros y los observadores entre septiembre y diciembre de 2020. El informe final se envía a los Estados Miembros mediante una circular para que lo examine el Comité Especial de expertos técnicos y jurídicos, que se reunirá en abril y junio de 2021 para finalizar el proyecto de recomendación sobre la ética de la inteligencia artificial y presentarlo a la Conferencia General en su 41^a reunión.

INTRODUCCIÓN

1. La inteligencia artificial (IA) es uno de los temas centrales de la era de las tecnologías convergentes, que tiene profundas repercusiones para la humanidad, las culturas, las sociedades y el medio ambiente. Las consecuencias de la tecnología dependerán de la forma en que la humanidad la enmarque y la domine, y de cuánta prioridad dé al objetivo de no dejar a nadie atrás. La pandemia de COVID-19 ha servido de ejemplo de la ayuda que pueden prestar las tecnologías de la IA, pero también ha puesto de relieve los riesgos, las consecuencias no deseadas y el uso malintencionado de la IA.

2. Por ello, la investigación sobre la IA y su desarrollo y utilización deben ir acompañados de una reflexión ética. Las tecnologías de la IA no son valorativamente neutras, sino sesgadas, entre otras cosas, debido a los datos sobre cuya base se capacitan, y a las opciones adoptadas durante la capacitación respecto a los datos. También influye el hecho de que las decisiones de la IA, en particular las basadas en el aprendizaje automático, no pueden ser totalmente predecibles ni comprendidas. Por otra parte, como la IA es una tecnología distribuida, cuya gobernanza actual la ejercen al tiempo numerosas instituciones, organizaciones y empresas, la reflexión sobre su buena gobernanza requiere un enfoque pluralista, multidisciplinario, multicultural y multipartito, que plantea interrogantes sobre el tipo de futuro que queremos para la humanidad. Esta reflexión debe abordar los principales desafíos derivados del desarrollo y uso de tecnologías de la IA en cuanto a los sesgos incorporados en los datos y los algoritmos, la protección de la privacidad de las personas y los datos personales, la falta de diversidad e inclusión en el modelo de negocio digital, las cuestiones

vinculadas a la distribución justa de los beneficios y los riesgos, la rendición de cuentas, la responsabilidad, las consecuencias para el empleo y el futuro del trabajo, los derechos humanos, la seguridad y los riesgos del doble uso.

3. Hay muchas organizaciones e iniciativas en todo el mundo que proponen enfoques para abordar los desafíos relacionados con la IA. La UNESCO tiene una perspectiva única para fortalecer los esfuerzos en curso orientados a garantizar una IA beneficiosa, dada su gran ventaja comparativa derivada de su composición universal y sus competencias multidisciplinarias. En este sentido, la Organización puede proporcionar una verdadera plataforma mundial y pluralista para el diálogo sobre la ética de la IA, reuniendo a países desarrollados y en desarrollo, diferentes puntos de vista culturales, así como diversas partes interesadas de los sectores público y privado. Por lo tanto, además de las numerosas directrices y marcos éticos que han elaborado los gobiernos, las empresas y las organizaciones de la sociedad civil, la UNESCO puede contribuir al desarrollo de la IA en beneficio de toda la humanidad, el desarrollo sostenible y la paz. Con este fin, la Organización actúa como puente entre, por un lado, los Estados Miembros, que han destacado en repetidas ocasiones su apoyo a la labor de la UNESCO en materia de IA, y, por otro, la sociedad civil, la comunidad técnica, el mundo académico y el sector privado, incluidas las industrias culturales y creativas. La UNESCO aprovecha también su larga trayectoria de elaboración de marcos éticos sobre ciencia, incluida la bioética, así como su experiencia en cuanto a consultas entre múltiples interesados y búsqueda de consenso. El mandato conferido por los Estados Miembros a la Organización es de gran importancia ya que, aunque hay numerosos esfuerzos encaminados a asegurar la buena gobernanza del desarrollo de la IA, el actual proyecto de recomendación de la UNESCO es el único instrumento normativo que expresa claramente que esta gobernanza debe reposar en consideraciones éticas y en un firme respeto de los derechos humanos y la dignidad humana.

4. Así pues, este instrumento normativo internacional contribuirá a alcanzar en el mundo un consenso ético sobre los retos más acuciantes que plantea la IA, que no es solo un punto de inflexión tecnológico, sino también una perturbación antropológica. Se trata de un paso necesario hacia un pensamiento más proactivo que sobreponga los enfoques jurídicos tradicionales, que se van quedando a la zaga. La recomendación propuesta deberá convertirse en una brújula ética y un pilar normativo que permitan fomentar un sólido respeto por el estado de derecho en el mundo digital, reconociendo al mismo tiempo su rica diversidad de culturas y contextos nacionales.

EL PROCESO

5. En cumplimiento de la [decisión 206 EX/42](#) y la [resolución 40 C/37](#), se determinó que una recomendación era el tipo de instrumento normativo internacional más adecuado en este ámbito. Por su parte, los Estados Miembros invitaron a la UNESCO a presentar a la Conferencia General, en su 41^a reunión, el proyecto de recomendación. En consonancia con la hoja de ruta que figura en el [anexo III del documento 40 C/67](#), en marzo de 2020 se creó el [Grupo Especial de Expertos](#) (GEE) (categoría VI) para elaborar el anteproyecto de recomendación. Éste estaba integrado por 24 expertos independientes, nombrados con arreglo a un principio de equilibrio geográfico entre todas las regiones de la UNESCO. De marzo a mayo de 2020, debido a la pandemia de COVID-19, el GEE llevó a cabo su labor utilizando modalidades en línea o virtuales a lo largo de seis semanas, y elaboró una primera versión del proyecto de recomendación.

6. Entre junio y agosto de 2020 se emprendió un proceso de consulta con múltiples interesados sobre la primera versión del proyecto de recomendación. A la luz de los desafíos que plantea la pandemia, un cierto número de [asociados](#) de todo el mundo colaboró con la UNESCO para apoyar este proceso de consulta basado en tres componentes: i) una consulta pública en línea; ii) consultas virtuales regionales y subregionales organizadas conjuntamente con las instituciones o países anfitriones en todas las regiones de la UNESCO; y iii) talleres de deliberación abiertos con la participación de múltiples interesados y ciudadanos, organizados por los asociados. La consulta pública en línea recibió respuestas completas de más de 600 participantes, lo que generó más de 50 000 comentarios. Además, se celebraron 11 consultas virtuales regionales y subregionales, dos

de ellas destinadas a la juventud, en todas las regiones de la UNESCO, en las que tomaron parte más de 500 personas. La mayoría de las consultas fueron realizadas a un alto nivel por la Subdirectora General de Ciencias Sociales y Humanas (SHS). Por último, en los talleres de deliberación abiertos, de múltiples interesados y de ciudadanos, participaron alrededor de 500 personas. Se desplegaron asimismo esfuerzos para consultar específicamente al sector privado, en particular las principales plataformas digitales, que también aportaron numerosas sugerencias. Se organizó una sesión especial con la Cámara de Comercio Internacional, que reunió a un amplio grupo de empresas.

7. También se celebraron consultas sobre la primera versión del proyecto de recomendación con el sistema de las Naciones Unidas en el contexto de la reunión entre períodos de sesiones del Comité de Alto Nivel sobre Programas, celebrada el 29 de julio de 2020. Todas las entidades expresaron su pleno apoyo a la labor de la UNESCO para elaborar esta recomendación, reconociendo su importancia y su pertinencia para el sistema de las Naciones Unidas, los Estados Miembros y otros interesados. En general, se hizo hincapié en que el texto podía y debía ser más ambicioso, poniendo más de relieve las características innovadoras de la recomendación y proponiendo un lenguaje más rotundo y recomendaciones de mayor alcance respecto a cuestiones como la gobernanza, los derechos humanos y la aplicación tras la adopción. Esto es necesario para garantizar un establecimiento eficaz de enfoques éticos de la IA en todo el mundo, distinguir a la recomendación de otros documentos en esta esfera, y ayudar a convertir a la UNESCO en una referencia en materia de desarrollo de capacidades de los Estados Miembros en el ámbito de la ética de la IA. Las observaciones de los asociados de las Naciones Unidas se transmitieron debidamente al GEE.

8. A continuación, el GEE revisó la primera versión del proyecto de recomendación desde mediados de agosto hasta principios de septiembre de 2020, tomando como base la información obtenida en las consultas y sus debates posteriores. Con arreglo a lo dispuesto en el párrafo 2 del artículo 10 del Reglamento sobre las recomendaciones a los Estados Miembros y las convenciones internacionales previstas en el párrafo 4 del artículo IV de la Constitución, el anteproyecto de recomendación se envió a los Estados Miembros junto con el informe preliminar en septiembre de 2020, pidiéndoles que formularan sus comentarios y observaciones antes del 31 de diciembre de 2020.

9. En total, 48 Estados Miembros y dos observadores presentaron sus comentarios y observaciones, incluidas propuestas concretas para modificar el texto. De conformidad con el párrafo 3 del artículo 10 del Reglamento antes mencionado, el proyecto de texto de recomendación y el informe preliminar fueron revisados de acuerdo con los comentarios recibidos. La Sección de Bioética y Ética de la Ciencia y la Tecnología (SHS/BIO) dedicó los meses de enero y febrero de 2021 a examinar e integrar los comentarios y a llevar a cabo un proceso consultivo, inclusivo y exhaustivo, en el que participó el Grupo de trabajo de puntos focales sobre la ética de la IA del equipo de tareas sobre la IA de la Secretaría de la UNESCO. El Grupo de trabajo comprendía puntos focales de ED, SC, CLT, GEN, AFR, BSP y CAB, y el proceso lo lideró SHS/BIO bajo la supervisión de la ADG/SHS. El resultado fueron el informe final y el proyecto de recomendación sobre la ética de la inteligencia artificial, que se adjunta.

10. En su 210^a reunión, el Consejo Ejecutivo ([210 EX/35](#)) aprobó las propuestas relativas a las invitaciones a las reuniones intergubernamentales de expertos (categoría II) que se celebrarían en el marco de esta recomendación. Mediante una circular, se invitó a los Estados Miembros a que designaran a los expertos que deseaban enviar a la reunión de expertos del Comité Especial Intergubernamental en enero de 2021.

11. Con arreglo a lo dispuesto en el párrafo 4 del artículo 10 del Reglamento mencionado, el informe final se presentará en la reunión de expertos del Comité Especial Intergubernamental, que tendrá lugar en dos fases: la primera del 26 al 30 de abril de 2021, y la segunda del 21 al 25 de junio de 2021. En la reunión intergubernamental de expertos se examinará el proyecto de recomendación. De conformidad con el párrafo 5 del artículo 10 del Reglamento en cuestión, el

proyecto aprobado por el Comité Especial se transmitirá a los Estados Miembros, para que sea examinado por la Conferencia General, en su 41^a reunión, en virtud de la resolución 40 C/37.

COMENTARIOS Y OBSERVACIONES RECIBIDOS

12. Ha sido alentador comprobar que la mayoría de los Estados Miembros y observadores que formularon comentarios expresaron su firme apoyo al anteproyecto de recomendación, incluida su estructura, ámbito de aplicación, formulación y exhaustividad. Muchos Estados Miembros reconocieron la importancia de los esfuerzos emprendidos por la UNESCO para elaborar esta recomendación.

13. La mayor parte de los comentarios y propuestas recibidos tenía por finalidad mejorar la formulación, aclarar términos, o añadir o suprimir párrafos o cambiar su orden. No obstante, algunos comentarios han incluido también reflexiones sobre cómo abordar determinadas cuestiones básicas, que se han analizado y tomado en cuenta debidamente. Al estudiar los comentarios, la Secretaría adoptó un enfoque consistente en solo introducir cambios cuando los Estados Miembros lo solicitaran específicamente, teniendo en consideración al mismo tiempo las opiniones y posturas de los demás Estados Miembros sobre los mismos temas y ajustando el texto para mejorar la claridad y la coherencia en todo el documento. A continuación se presenta un resumen de algunas de las principales cuestiones.

14. Los Estados Miembros consideraron que el hecho de tener un ámbito de aplicación amplio y exhaustivo, sin sobrepasar los límites del mandato de la UNESCO, habida cuenta del papel que desempeña la Organización en materia de ética y su objetivo global de promover la paz, era ciertamente un punto fuerte del instrumento. Algunos Estados Miembros señalaron la importancia de centrarse en las competencias básicas de la UNESCO, mientras que otros estimaron necesario plasmar el amplio alcance de las repercusiones éticas, tal como se reflejan en los ámbitos de actuación del proyecto de recomendación. La incorporación en el ámbito de aplicación del proyecto de recomendación de las implicaciones éticas para todas las partes interesadas de la IA en todas las etapas del ciclo de vida de los sistemas de IA fue acogida con particular satisfacción. Varias sugerencias permitieron definir con mayor precisión el ámbito de aplicación del instrumento para que se ajustara mejor a las competencias de la Organización.

15. Se hizo hincapié en particular en el ámbito de aplicación civil del instrumento. Aunque algunos Estados Miembros sugirieron que se considerara la utilización como arma de la IA como parte del mandato de la UNESCO, y que se establecieran directrices éticas específicas al respecto, la mayoría de los comentarios sobre esta cuestión era favorable a indicar claramente que la recomendación tenía un ámbito de aplicación civil y que, por tanto, no abarcaba las cuestiones del uso con fines militares de la IA, incluidos temas como las armas autónomas letales, la utilización como arma y el derecho humanitario. Se aceptaron las revisiones correspondientes, en particular la supresión de la referencia a la utilización como arma. Sin embargo, el texto de la recomendación afirma que un sistema de IA nunca podrá reemplazar la responsabilidad y la rendición de cuentas final por parte de un ser humano, y que en los casos que implican decisiones de vida o muerte, la decisión final debería ser adoptada por un ser humano.

16. Asimismo, se aclaró el enfoque del instrumento sobre la ética de la IA. Otras tecnologías, como la Internet de las cosas, la robótica y las interfaces entre seres humanos y ordenadores, se mencionan únicamente como ejemplos de sistemas y entornos convergentes en los que puede encontrarse la IA.

17. Los Estados Miembros acogieron con satisfacción la estructura de la recomendación, en especial la parte sobre los valores, que presenta un panorama normativo general; la parte sobre los principios, que establece un marco ético más concreto para guiar el desarrollo y la aplicación de la tecnología; y el conjunto completo de recomendaciones, clasificadas en ámbitos de actuación claramente definidos, a fin de orientar a las partes interesadas de la IA para hacer realidad estos valores y principios. Unos cuantos países señalaron que la distinción entre valores y principios no era lo suficientemente clara, aunque para la mayoría resultaba adecuada; así pues, los valores

fundamentales cumplen la función de condiciones y requisitos previos y necesarios para que los principios funcionen, mientras que los principios se basan en estos valores fundamentales y los desglosan de forma más detallada de acuerdo con cuestiones éticas específicas, garantizando al mismo tiempo la adecuación entre los valores y los principios. Se apreció especialmente la definición de ámbitos de actuación concretos, que además de ser audaz, innovadora y orientada a las repercusiones, permitía al instrumento no centrarse únicamente en lo que es la ética, sino también en sus modalidades de aplicación. No obstante, se hicieron varias sugerencias, que se aceptaron con miras a reforzar y racionalizar la estructura del documento.

18. En particular, la revisión tuvo en cuenta las sugerencias de varios Estados Miembros de establecer un ámbito de actuación específico sobre la comunicación y la información, con la finalidad de destacar las consecuencias éticas de la IA en la libertad de expresión, el discurso de odio y la desinformación y deformación de la realidad. Los Estados Miembros reconocieron que ciertos elementos pertinentes de este ámbito ya figuraban en distintas partes del documento, pero se pedía que se agruparan en un ámbito de actuación específico. Por lo tanto, se añadió al instrumento un nuevo ámbito de actuación sobre la comunicación y la información, en el que se integraron los elementos existentes referentes a este tema. A nivel de la Secretaría, el contenido del ámbito de actuación de la comunicación y la información fue objeto de debates entre los equipos de SHS y CI, en los que también participó el Grupo de trabajo de puntos focales, y el texto acordado por estas entidades se presentó a los Estados Miembros para que lo examinaran.

19. Si bien algunos Estados Miembros opinaron que el documento era demasiado largo, otros deseaban que se tratara más a fondo el contenido, y ofrecieron propuestas concretas para que se tuvieran en cuenta en el proyecto de la recomendación. En respuesta a las sugerencias de varios Estados Miembros, se mejoró el uso de los ejemplos. En particular, se eliminaron los ejemplos de ciertas partes en las que los Estados Miembros consideraban que su uso limitaba la interpretación y aplicación más amplias de los principios y medidas políticas pertinentes. No obstante, se siguieron utilizando ejemplos en aquellas partes en las que ayudaban a explicar o reforzar un principio concreto o un curso de acción recomendado, sin presentar listas extensas ni análisis en profundidad, dada la naturaleza de instrumento jurídico internacional de la recomendación (véase, por ejemplo, el párrafo 94).

20. En cuanto a los destinatarios, algunos Estados Miembros pidieron más precisiones en cuanto a las obligaciones del sector privado y otros interesados. En el proyecto de recomendación se reconoce que, dadas las repercusiones de gran alcance de la IA, los Estados Miembros no pueden participar solos en este esfuerzo. Por lo tanto, el proyecto tiene en cuenta las consecuencias éticas para todos los interesados de la IA en todas las etapas del ciclo de vida de los sistemas de IA dentro de su ámbito de aplicación, y hace hincapié en el principio de “gobernanza y colaboración adaptativas y de múltiples partes interesadas”. Se anima a los Estados Miembros a que promuevan mecanismos de gobernanza inclusiva de la IA y a que establezcan marcos adecuados para que las demás partes interesadas, en particular las empresas del sector privado, los apliquen. Por otro lado, aunque la recomendación se dirige específicamente a los Estados Miembros, también puede servir de guía para todos los países, incluidos los que no son miembros, y todos los interesados de la IA, a fin de asegurar la incorporación de la ética en todas las etapas del ciclo de vida de los sistemas de IA.

21. Ciertos Estados Miembros sugirieron que se alineara el proyecto de recomendación a los marcos existentes relacionados con la ética de la IA elaborados por otras organizaciones, como la OCDE y el G20. En realidad, los marcos existentes se tuvieron en cuenta durante el proceso de elaboración del proyecto de recomendación, y se mencionan específicamente en el preámbulo. Además, el proyecto se ajusta a muchos de estos marcos, con algunos de cuyos principios crea sinergias, siendo a la vez más amplio y de mayor alcance. Ciertamente, el valor añadido de la recomendación es su carácter mundial, ya que los miembros de la UNESCO proceden de todos los países, y su firme base ética. Como la ética cobra cada vez más importancia para la toma de decisiones en diversos ámbitos, no sirve únicamente para la reflexión, sino que además es una guía moral práctica y menos rígida que una reglamentación. Por otra parte, la recomendación va más

allá de una declaración de valores y principios éticos, habida cuenta de que contiene medidas políticas concretas para su aplicación. Se prestó especial atención a la igualdad de género, que es una de las dos prioridades globales de la UNESCO, así como al medio ambiente y los ecosistemas, ya que estas esferas tienden a descuidarse en otros marcos similares. Debido a su carácter mundial, la recomendación da prioridad a los países de ingreso bajo y mediano, para garantizar que no se queden atrás en los avances en el ámbito de la IA.

22. Algunos Estados Miembros propusieron añadir definiciones de diversos términos y conceptos en el proyecto de recomendación, incluida la definición de ética. El enfoque inicial consistía en no dar definiciones en este instrumento jurídico, dado que en muchos casos no había consenso a pesar de los extensos debates académicos, y con miras a que el instrumento pudiera adaptarse a la evolución de ciertos conceptos. En cambio, se describieron con bastante detalle determinadas nociones esenciales a los efectos del proyecto de recomendación, incluida la ética de la IA en el párrafo 1, y los sistemas de IA, el ciclo de vida de los sistemas de IA y los interesados de la IA en el párrafo 2. Además, a los efectos del proyecto de recomendación, y teniendo en cuenta la práctica habitual, algunos términos se utilizan indistintamente a lo largo del texto, como sistemas de IA y tecnologías de IA, Estados Miembros y gobiernos, o empresas comerciales y empresas del sector privado.

23. También se plantearon preocupaciones específicas en cuanto a la elección de palabras. En consonancia con el enfoque adoptado por la UNESCO en sus recomendaciones anteriores, el verbo “deber” se emplea en condicional en todo el proyecto de recomendación al sugerir las medidas que podrían emprender los Estados Miembros. Sin embargo, dadas las numerosas preocupaciones de los Estados Miembros sobre el hecho de que las obligaciones en virtud del derecho internacional, incluido el derecho de los derechos humanos, no son voluntarias, en estos casos se emplea el verbo “deber” en indicativo. Asimismo, se escogió la palabra “mundial” en lugar de “común”, como propusieron algunos países, en lo que se refiere a la aceptación de este instrumento, así como a las herramientas y metodologías que describe, a fin de indicar su aplicación universal, habida cuenta de la composición internacional de la UNESCO, y su pertinencia en todos los contextos nacionales y culturales.

24. Conviene responder a las preocupaciones y comentarios planteados por algunos Estados Miembros sobre la definición de la ética, los sistemas éticos, la situación de la ética en relación con los derechos humanos y el riesgo de caer en el relativismo moral. El mandato recibido de los Estados Miembros consistía en preparar un documento sobre la ética de la IA. Se acordó que sería uno de los principales valores añadidos y la contribución de la UNESCO a las iniciativas, ya numerosas, sobre la IA, ya que la ética es un elemento esencial del mandato de la Organización, y que ésta tiene más de 25 años de experiencia al respecto. El proyecto de recomendación considera que la ética de la IA es una base dinámica de reflexión que puede guiar a las sociedades a la hora de afrontar de manera responsable los efectos conocidos y desconocidos de la IA, con el objetivo de garantizar la protección de la dignidad humana, los derechos humanos, el bienestar y la prevención de daños. Las tecnologías en general, y las de IA en particular, evolucionan muy rápidamente y no siempre es fácil anticipar sus consecuencias. Es por esta razón que no basta con disponer de un conjunto fijo de principios o códigos éticos. Debe emprenderse continuamente una reflexión ética inclusiva para garantizar que estas tecnologías sigan utilizándose sin atentar contra la dignidad humana, los derechos humanos y las libertades fundamentales. El hecho de considerar la ética como un proceso dinámico no implica un cambio de valores. Por el contrario, se trata de velar por que se respeten siempre los valores y de adaptar los marcos existentes, incluidos los jurídicos, cuando sea necesario, a fin de garantizar la protección de la dignidad humana, los derechos humanos y las libertades fundamentales, al tiempo que las tecnologías avanzan. La ética guarda relación con los valores y es la fuente de toda decisión y marco jurídico; explícita o implícitamente, sustenta las políticas y las leyes. Los derechos humanos son una forma de codificar estos principios mediante la ley. Sin embargo, como sabemos, no todas las cuestiones éticas pueden abordarse únicamente mediante la ley, ni codificarse o reglamentarse, y podría ser necesario tomar decisiones para asegurar el pleno respeto del marco de los derechos humanos en

casos específicos donde no exista ninguna reglamentación. La ética es entonces la única herramienta disponible.

25. En el proyecto de recomendación se reconoce la existencia de diferentes sistemas éticos, así como la diversidad cultural. Al mismo tiempo, el objetivo que persigue la UNESCO con este proyecto de recomendación, al igual que con todos sus instrumentos normativos, consiste en encontrar un pilar ético común en el que pueda reposar la comunidad internacional para lograr una utilización responsable de las tecnologías de la IA en el marco de los derechos humanos de las Naciones Unidas. Esto explica que se reconozcan los diferentes sistemas éticos y que sea necesario dialogar con todos ellos. En otras palabras, la idea es garantizar el pluralismo, procurando evitar el relativismo, y establecer una base común en materia de ética.

26. Se expresó la preocupación general de que los derechos humanos, tal como se mencionan en el anteproyecto de recomendación, pudieran interpretarse como valores éticos o conceptos más amplios, en lugar de obligaciones jurídicas claras en virtud del derecho internacional. El proyecto se enmendó para reforzar el carácter obligatorio de las normas sobre derechos humanos. Esto supone garantizar que los derechos humanos no sean objeto de ningún compromiso. En general, la Secretaría ha escogido ceñirse al marco internacional de los derechos humanos y no ofrecer nuevas interpretaciones, incluso cuando se trata de posibles repercusiones negativas de las tecnologías de la IA para los derechos humanos. Además, se afianzaron las referencias a las obligaciones derivadas del derecho internacional, en particular la cláusula general que figura al final del proyecto de recomendación, y que establece que este instrumento no va en perjuicio del derecho internacional y las obligaciones referentes a los derechos humanos.

27. Muchos Estados Miembros acogieron con satisfacción la atención específica que se presta a la igualdad de género en el proyecto de recomendación, y algunos proponen integrarla en todo el documento. En cambio, varios Estados Miembros manifestaron su preocupación por el gran énfasis que se da el género, en detrimento de otros conceptos como la justicia social y la igualdad en general. El género se aborda en un ámbito de actuación específico en el proyecto de recomendación, debido a que es una de las dos prioridades globales de la UNESCO y a las preocupaciones específicas que han surgido a este respecto en la esfera de la IA. Aunque se ponga de relieve el género en un ámbito de actuación específico, esto no implica que el proyecto de recomendación solo aborde la cuestión de la igualdad con un enfoque de género. El proyecto de recomendación también incorpora el valor de la “Diversidad e inclusión” y el principio de “Equidad y no discriminación”, así como acciones políticas específicas para aumentar la diversidad y combatir los estereotipos de todo tipo. Ciertos Estados Miembros también propusieron alternativas para sustituir los términos “género” e “igualdad de género”. Pero como estos términos ya figuran en diversos documentos e instrumentos de las Naciones Unidas, se mantuvieron en aras de la coherencia. Muchos Estados Miembros deseaban, por su parte, fortalecer el ámbito de actuación sobre el género y ampliar su alcance.

28. Algunos Estados Miembros expresaron su preocupación por el lugar que ocupa la protección de los derechos de propiedad intelectual en el proyecto de recomendación, incluidas las cuestiones relativas a los derechos de autor. Es importante señalar que el proyecto se refiere al derecho internacional, que comprende las leyes de propiedad intelectual. Además, la propiedad intelectual se menciona explícitamente en varias ocasiones. En particular, en el párrafo 99 se insta a los Estados Miembros a “promover nuevas investigaciones en la intersección entre la IA y la propiedad intelectual”. Estas investigaciones son necesarias para garantizar que se cumplan los requisitos de transparencia, que pueden entrar en conflicto con los secretos comerciales. Esto no significa que haya que revelar siempre los secretos comerciales, ya que su protección se reconoce y acepta ampliamente. Existen muchas maneras de asegurar la transparencia, y no todas ellas exigen mostrar la caja negra completamente. El objetivo principal, no obstante, es evitar abusar de la protección del secreto comercial e introducir cláusulas de propiedad que impidan incluso al organismo que compra el algoritmo entender cómo funciona. Esto es especialmente importante en los ámbitos sensibles, como se indica en el párrafo 39 de la recomendación: “En los casos en que

se prevean graves repercusiones adversas para los derechos humanos, la transparencia puede requerir también que se comparten códigos o conjuntos de datos específicos”.

29. Varios Estados Miembros se mostraron partidarios de una reglamentación más eficaz, tal y como se especificaba en los ámbitos de actuación del anteproyecto de recomendación, mientras que otros expresaron su preocupación por el hecho de que se propusieran medidas reguladoras. Cabe recordar que una recomendación de la UNESCO es un instrumento que formula los principios y normas destinados a reglamentar internacionalmente cuestiones específicas, e invita a los Estados Miembros a adoptar cualesquiera medidas legislativas o de otro carácter que sean necesarias según las particularidades de las cuestiones de que se trate y las disposiciones constitucionales de cada Estado para aplicar en sus respectivos territorios los principios y normas formulados. Esto significa que el proyecto de recomendación no impone obligaciones a los Estados Miembros, sino que enuncia metas y objetivos comunes que deben alcanzarse, y propone formas de conseguirlos, aunque los medios para ello se deciden en el plano nacional. Las propuestas formuladas en el proyecto no pretenden impedir la innovación, sino velar por que se realice de forma ética.

30. El valor “Vivir en armonía y paz”, propuesto en el anteproyecto de la recomendación, fue objeto de un número considerable de comentarios y propuestas de modificación por parte de los Estados Miembros. Las principales preocupaciones se referían a la falta de claridad del significado del término “armonía”, sus consecuencias para los derechos humanos y el hecho de que se preste a diferentes interpretaciones y, posiblemente, a malentendidos. Por lo tanto, se revisó en gran medida el contenido de este valor, teniendo en cuenta las sugerencias específicas formuladas por varios Estados Miembros, y procurando mejorar la claridad y reforzar el compromiso con las obligaciones en materia de derechos humanos y su cumplimiento. Se cambió el título por “Vivir en sociedades pacíficas, justas e interconectadas”.

31. A algunos Estados Miembros les preocupaba que el principio de “Proporcionalidad e inocuidad” diera a entender que las tecnologías de IA son intrínsecamente malas y solo deberían utilizarse cuando fuera absolutamente necesario, y que, de manera más general, este principio fuera difícil de aplicar en la práctica. Sin embargo, la prueba de necesidad, que forma parte del principio general de proporcionalidad común a muchos ordenamientos jurídicos, no debería interpretarse de forma tan restrictiva, sino más bien como favorable para crear condiciones propicias. La aplicación de este principio se verá facilitada por la herramienta de evaluación del impacto ético propuesta en el proyecto de recomendación.

32. Como se indica en el principio “Supervisión y decisión humanas”, la postura que adopta el proyecto de recomendación es que el control debe estar siempre en manos del ser humano, y que “un sistema de IA nunca podrá reemplazar la responsabilidad y la rendición de cuentas final por parte de un ser humano” (párrafo 36). Para ello se deberá fortalecer la capacidad humana paralelamente al desarrollo de los sistemas de IA, a fin de garantizar la supervisión humana. Pero se reconoce que puede haber situaciones en las que la decisión humana no sea posible, aunque la decisión de ceder el control en contextos limitados seguirá recayendo en los seres humanos. Estas cuestiones deberían abordarse en el marco de la metodología de evaluación del estadio de preparación.

33. Como se menciona en los párrafos 46 y 73 del proyecto de recomendación, se deberá respetar la soberanía de los datos en su tratamiento. Este concepto se mantuvo debido a su importancia. Su finalidad principal consiste en proteger a las personas y mejorar su capacidad de mantener el control de sus datos personales. Como reiteró el GEE, también busca proteger a los Estados de los abusos que puedan producirse en el tratamiento de los datos de sus ciudadanos en otros países, lo que también podría poner en peligro la seguridad nacional. Al defender este concepto, el objetivo es animar a los Estados Miembros, en colaboración con otras partes interesadas, a establecer la infraestructura necesaria, incluidos centros de datos para el tratamiento de los datos de sus ciudadanos de acuerdo con las leyes nacionales. Es preciso destacar que el concepto de soberanía de los datos no incide en el intercambio y los flujos de datos. Han empezado a aparecer diversos

marcos para el control de los datos, así como mecanismos para compartirlos, centrados esencialmente en poder separar los diferentes componentes de los datos y tener un control total sobre lo que se comparte y para qué fines.

34. Para ayudar a los Estados Miembros a aplicar la recomendación y definir las medidas concretas que deberán adoptarse, el proyecto de recomendación precisa que la UNESCO elaborará determinadas herramientas, como la evaluación del impacto ético y la metodología de evaluación del estadio de preparación. Aunque la Organización reconoce que no es fácil elaborar tales mecanismos, que abarcan una amplia gama de circunstancias, se trata de una etapa de seguimiento importante que respaldará la recomendación, confiriéndole una dimensión más práctica. La UNESCO está dispuesta a ayudar a los Estados Miembros a realizar dichas evaluaciones, en función de las solicitudes. Desde hace mucho tiempo prepara metodologías para respaldar la aplicación de instrumentos internacionales, y participa en un proceso inclusivo que tendrá en cuenta las diversas perspectivas de los Estados Miembros, así como de las diferentes partes interesadas.

35. Varios Estados Miembros hicieron hincapié en la necesidad de reforzar la autonomía personal en el texto. Dado que la autonomía es inherente a la dignidad humana y a los derechos humanos y las libertades fundamentales, estos aspectos se afianzaron en todo el proyecto de recomendación. El texto también pone énfasis en la autonomía al abarcarla de diversas maneras, especialmente garantizando la capacidad de actuar de los seres humanos, el control humano, el pensamiento crítico, etc.

36. Un cierto número de Estados Miembros propuso que la recomendación incorporara un enfoque de la IA basado en los riesgos. Aunque este enfoque tiene sus ventajas, plantea asimismo la cuestión de definir cuáles son los criterios que determinan el riesgo y quién debería hacerlo. Además, es imperativo que se tengan en cuenta consideraciones éticas para todos los sistemas de IA, independientemente de sus niveles de riesgo. Por otra parte, aunque tomar en cuenta el riesgo es importante, otros elementos también lo son, como los beneficios y las repercusiones, tal como sugirieron ciertos Estados Miembros. Así pues, habida cuenta de estas consideraciones, el proyecto de recomendación no adoptó un enfoque basado en los riesgos.

37. Por último, pero no menos importante, algunos Estados Miembros destacaron que la recomendación no debía interferir en los asuntos internos de los países. En respuesta a esta preocupación, la Secretaría desea subrayar el carácter no vinculante de la recomendación. Además, en el párrafo 48 del proyecto de recomendación se indica que “El proceso de elaboración de esas políticas o mecanismos debería [...] tener en cuenta las circunstancias y prioridades de cada Estado Miembro”. La recomendación no pretende en modo alguno interferir en los asuntos internos de los Estados Miembros, sino más bien proporcionar un marco universal de valores, principios y acciones para orientar a los Estados en la formulación de sus leyes, políticas u otros instrumentos pertinentes en esta esfera. Aunque la recomendación no es vinculante, la UNESCO desea poner de relieve, en el contexto de un marco multilateral, la importancia de la cooperación para lograr nuestro objetivo común de hacer que los sistemas de IA estén al servicio de la humanidad.

OBSERVACIONES FINALES

38. Se prevé que, para ser eficaz, la recomendación deberá ir acompañada de un firme apoyo en materia de políticas y desarrollo de capacidades para su aplicación y transformación en marcos jurídicos, y de prácticas específicas de IA en los sectores público y privado. Por ello, la UNESCO aportará sus competencias multidisciplinarias a fin de ayudar a los Estados Miembros a cumplir las disposiciones de la recomendación, una vez adoptada. La recomendación también proporcionará una base sólida para todo el sistema de las Naciones Unidas en su respuesta colectiva a los desafíos éticos que plantean las tecnologías de la IA en diversos ámbitos, en los que la UNESCO desempeña un papel de liderazgo. A este respecto, la Organización tiene más de 25 años de experiencia en la elaboración y aplicación de trabajos normativos mediante asesoramiento sobre políticas y actividades de creación de capacidad en materia de bioética y ética de la ciencia y la tecnología.

39. El proyecto de texto adjunto se presenta a la reunión intergubernamental del comité especial de expertos técnicos y jurídicos designados por los Estados Miembros, para su finalización. El proyecto de recomendación, aprobado por el comité especial, se transmitirá a los Estados Miembros para que la Conferencia General lo apruebe en su 41^a reunión.

40. Los comentarios y observaciones recibidos confirman la necesidad de disponer de este instrumento normativo en forma de recomendación. Dada la importancia de lo que está en juego y de los desafíos que nos esperan, y a pesar de las posibles divergencias, es fundamental alcanzar un consenso en este momento crucial de la historia de la humanidad.