



Ética e inteligencia artificial en el Sistema Universal de Derechos Humanos

Ethics and artificial intelligence in the Universal System of Human Rights

RESUMEN

Una pregunta que resulta factible formularse frente al acelerado desarrollo de la inteligencia artificial (IA) sería sobre los aspectos que merecen destacarse cuando ponemos frente a frente las variables disrupción tecnológica, derechos humanos y **ética**. En este artículo se plantea como *objetivo* describir la dirección que se ha tomado en materia de protección de derechos humanos en la sede de la Organización de las Naciones Unidas (ONU) y el Sistema ONU. Para la consecución de los objetivos, se hizo un análisis crítico de los instrumentos internacionales recientemente aprobados por los diversos órganos, materia prima de la técnica documental aquí utilizada. Se revisaron las acciones significativas a efecto de destacar los factores más relevantes que han permitido confeccionar una ética para la denominada cibernsiedad en el contexto del presente siglo. Uno de los hallazgos que saltan a la vista son los retos a los que se enfrenta el órgano supranacional al momento de dar cauce a este tipo de herramientas digitales: regulación, validez, amenazas, etcétera, en virtud de que los instrumentos adoptados son predominantemente de naturaleza *soft law*, lo que para algunos podría implicar un alcance limitado. Se concluye sobre el papel determinante de la ONU y sus organismos subsidiarios frente a las amenazas de los derechos humanos que implica la IA, sobre todo en la formulación de criterios éticos que marquen el rumbo a seguir en el seno de la comunidad internacional.

Palabras clave: derechos humanos; ética; inteligencia artificial; ONU.

1 Investigador y profesor de tiempo completo en la Facultad de Derecho de la Universidad Autónoma de Querétaro (uaq), México. Correo electrónico: raul.canizales@uaq.mx; orcid: <https://orcid.org/0000-0001-8428-3711>.

2 Investigador y profesor de tiempo completo en la Facultad de Derecho de la uaq, México. Correo electrónico: javier-rascado@hotmail.com; orcid: <https://orcid.org/0000-0001-5589-0619>.

ÉTICA E INTELIGENCIA

ARTIFICIAL EN EL SISTEMA UNIVERSAL DE DERECHOS HUMANOS

ABSTRACT

A question that is feasible to ask in the face of the accelerated development of artificial intelligence (AI) would be about the aspects that deserve to be highlighted when we put the variables of technological disruption, human rights and ethics face to face. This article aims to describe the direction that has been taken regarding the protection of human rights at the headquarters of the United Nations (UN) and the UN System. To achieve the objectives, a critical analysis was made of the international instruments recently approved by the various bodies, raw material for the documentary technique used here. The significant actions were reviewed in order to highlight the most relevant factors that have allowed the development of ethics for the so-called cybersociety in the context of the present century. One of the findings that stand out are the challenges that the supranational body faces when giving channel to this type of digital tools: regulation, validity, threats, etc., given that the instruments adopted are predominantly soft law in nature, which for some could imply a limited scope. It concludes on the determining role of the UN and its subsidiary organizations in the face of threats to human rights that AI implies, especially in the formulation of ethical criteria that mark the course to follow within the international community.

Keywords: artificial intelligence; ethics; human rights; UN.

INTRODUCCIÓN

La inteligencia artificial (IA) tiene décadas de ser parte de la vida cotidiana, aunque el grado de interactividad siempre ha dependido del tipo de sociedad y de su avance tecnológico. El acelerado desarrollo en los últimos años no es casual, ya que a partir de la experiencia de la pandemia por el COVID-19 lo que se observó fue, en realidad, un impulso a pasos agigantados en el mundo de las tecnologías (TIC), resultado del fenómeno denominado migración digital. Este tránsito —acelerado y forzado— hacia

el mundo de la virtualidad no se presentó aislado de aspectos problemáticos en una nutrida lista de aristas. Dentro del amplio espectro de innovaciones tecnológicas que se aceleraron durante y después de la pandemia, uno de los más inquietantes —pero también fascinantes— ha sido el de la IA. Sin duda constituye un aspecto fundamental de la sociedad del conocimiento y de las denominadas sociedades digitales, en las que la carrera por la innovación en este tipo de tecnologías ha estado acompañada también de mitos, incertidumbres, narrativas, temores y no sin razón.

En toda sociedad que experimenta estos contextos disruptivos siempre permea en ellas una especie de narrativa, mediante la cual se despliega un escenario de incertidumbres, lo que significa, a su vez, que se extiende un binomio inevitable: bondades y amenazas del desarrollo tecnológico, y en este caso específico aquellos asignados a la IA. Es claro que cuando únicamente se difunden, en esa guerra de opuestos (beneficios-amenazas), los posibles daños colaterales, la sociedad se coloca inevitablemente en visiones segmentadas. La suma de todo lo anterior ha generado una especie de desconfianza social hacia la IA y a esto ha contribuido su incipiente marco jurídico regulatorio (Aneja, 2021, p. 29), así como la tardía definición —desde el marco supranacional— de los linderos éticos que orientaran este tipo de herramientas.

Desde hace apenas algunos años a la fecha en la Asamblea General (AG) de la ONU (y su Sistema), se han aprobado un conjunto de instrumentos internacionales enfocados a orientar las pautas éticas en la materia que aquí se aborda. Normalmente se les denomina “instrumentos jurídicos internacionales”, los cuales, dependiendo del órgano subsidiario que los emite, tienen asignada una nomenclatura propia de la entidad supranacional. Los más comunes son las resoluciones de la AG, mediante las cuales se publicitan informes, estudios, acuerdos, etcétera, y cuyas fuentes de las que emanan van desde los órganos de tratados hasta los propios relatores especiales, expertos independientes, procedimientos especiales, etcétera.

Los instrumentos que aquí se abordan pertenecen al ámbito extraconvencional. En este esquema no se depende de un tratado internacional (TI) y, por ende, predominan los instrumentos de naturaleza denominada *soft law*. Más allá de la discusión que se ha generado en torno a estos documentos y su vinculatoriedad (Laporta, 2014; Escudero, 2012; Restrepo, 2018; Benavides, 2021), resulta oportuno recordar que detrás de ellos están implicados órganos subsidiarios que la ONU ha ido creando para la protección de los derechos humanos en virtud de cada uno de los TI. Son órganos basados en la Carta de Derechos Humanos (*Charter-Based*) o Carta Internacional de Derechos Humanos.

Ahora bien, en cuanto al tema de la vinculatoriedad de los documentos que se describirán, ello no es óbice para adjudicarles valor alguno, si no jurídico, sí lo adquieren en términos de su función y la dimensión donde están colocados: el campo de la ética, lo que significa, a su vez, un espacio donde la materia prima son instrumentos que ofrecen los ingredientes para un futuro marco normativo y, por tanto, se erigen como agregadores de normas (Rodrigo, 2022, pp. 74-75).

A partir de las anotaciones previas, el *objetivo* que aquí se persigue es presentar una descripción lineal cronológica del trayecto tomado en torno a la IA y la protección de los derechos humanos —principalmente en la jurisdicción de la ONU— y otros organismos subsidiarios del Sistema ONU. Esta revisión se verificó a partir de la aprobación y adopción de una serie de instrumentos internacionales, a través de los cuales este órgano supranacional y los demás subsidiarios pretenden arrojar luz respecto de los aspectos éticos que están en juego. Mediante un análisis histórico-cronológico, se presenta un enfoque descriptivo a efecto de revisar los principales instrumentos internacionales aprobados por los diversos órganos, en los que se destacan los factores más relevantes que permitan confeccionar una posible ética para la denominada cibernación en el contexto del siglo XXI.

MARCO TEÓRICO

Desde hace décadas las cuestiones relacionadas con el desarrollo tecnológico —en general— han constituido un tema central en la agenda del organismo supranacional objeto de estudio en esta disertación, pero también del propio subsistema, es decir, los organismos subsidiarios cuyas acciones se encuentran coordinadas por aquel. Merece destacar que, por tratarse de una entidad de la naturaleza que la caracteriza, requiere la inversión de enormes esfuerzos encaminados a proporcionar rutas a seguir y que estas encuentren eco en el interior de las naciones que la integran, sobre todo en lo que respecta a la protección de los derechos humanos. El caso específico de la IA no es la excepción y, como asignatura, se le ha incluido en el catálogo de los denominados derechos humanos de tercera generación o derechos de la era tecnológica como resultado de una reformulación de nuevos derechos o una actualización y adecuación a los nuevos desafíos a los que se enfrentan los mecanismos de garantías de derechos preexistentes (Pérez, 2018, p. 144).

Ha sido la propia ONU y su Sistema quien ha asumido un papel protagónico en la tarea de marcar los senderos que dicho órgano convoca a transitar en este campo. Coordinar esfuerzos en un área que advierte cada día un ritmo acelerado de desarrollo e innovación no es tarea fácil y, para

esto, se requiere identificar el mapa completo de este cuerpo de conocimientos, lo que a su vez exige congregarse los perfiles de expertos que permitan una revisión multidisciplinaria de las virtudes y amenazas derivadas de la puesta en práctica de esta tecnología. Una vez identificado el mapa completo, gracias a la convergencia de saberes, toca trazar esa ruta encaminada a formular los principios, criterios, amenazas, reglas de juego y, entre otras, las coordenadas éticas que orienten las acciones humanas al momento de diseñar herramientas tecnológicas cuya función esté ligada total o parcialmente a la IA.

En esta disertación se fijó como principal *objetivo* compartir una descripción respecto del rumbo, la dirección que recientemente ha adoptado la ONU y en el Sistema ONU en torno a la IA frente a los desafíos y promesas para la protección de los derechos humanos. Para lo anterior, se desplegó un breve análisis crítico de los instrumentos internacionales recientemente aprobados por los diversos órganos. Asimismo, se ofrece una revisión de las acciones con impacto directo en la temática, a fin de resaltar las cuestiones más relevantes que permitan formular una posible ética para la cibernética del siglo XXI.

LA ÉTICA Y LA INTELIGENCIA ARTIFICIAL EN LA JURISDICCIÓN DE LA ONU Y EL SISTEMA ONU

Sería una tarea imposible abordar en este artículo el conjunto completo de los instrumentos internacionales aprobados en el seno de la Asamblea General de la ONU, así como los propios de los organismos subsidiarios que conforman el denominado Sistema ONU, es decir, comisiones orgánicas, oficinas, departamentos, grupos de expertos, etcétera. Se trata de una multiplicidad de entidades imposibles de abarcar en esta ocasión. Este amplio abanico de áreas obliga a enfocarse tanto en las resoluciones como en las acciones más actuales y de mayor impacto en la línea de reflexión que aquí se propone. Por lo que el análisis se limitará a dos de ellas: 1) La Comisión de Ciencia y Tecnología para el Desarrollo, perteneciente al Consejo Económico y Social (Ecosoc); y 2) la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (Unesco), organismo especializado según el propio organigrama de la ONU.

Comisión de Ciencia y Tecnología para el Desarrollo

a) La Resolución E/CN.16/2024/3: *Cooperación mundial en ciencia, tecnología e innovación para el desarrollo* (Consejo Económico y Social, 2024), aprobada en el 27.º periodo de sesiones de esta

Comisión (del 15 al 19 de abril de 2024), dentro del Tema 3b de su programa: Ciencia y tecnología para el desarrollo: Temas prioritarios.

Se trata de la más reciente. A manera de resumen, se comparten los siguientes puntos esenciales. En este documento el secretario general (SG) de la ONU hace una revisión respecto de los mecanismos de cooperación mundial en lo concerniente a la ciencia, tecnología e innovación (CTI) para el desarrollo. Son cuatro rubros en los que se despliega la revisión: 1) Planificación estratégica; 2) Facilitadores de CTI; 3) Investigación y desarrollo; y 4) Innovación. En estos cuatro incisos recae la supervisión mencionada, en la que se destacan todos los mecanismos de colaboración, progresos, etcétera, así como las experiencias obtenidas y, por tanto, las denominadas buenas prácticas.

En el texto también se pueden leer las recomendaciones para que las organizaciones de CTI estén en condiciones de mejorar sus mecanismos de colaboración y cooperación, con el propósito de ampliar el impacto positivo de las tecnologías, pero siempre a partir de una serie de alianzas equitativas. De acuerdo con el SG de la ONU, lo anterior posibilitaría capacitar a los países que advierten menos desarrollo o, si se quiere decir así, advierten un retraso en materia de CTI. Una consecuencia deseable derivada del escenario anterior

es que las naciones puedan estar al día ante un contexto de desarrollo tecnológico acelerado. En este mismo sentido, en esta resolución se hace alusión al fenómeno de la cada vez más compleja realidad de las nuevas tecnologías y de su vertiginosa evolución, pero también se alude a las profundas transformaciones que han derivado de las recientes expresiones de innovación, circunstancia que para el SG amerita adoptar un enfoque colaborativo en CTI, que se traduce en una estrategia prometedora para cumplir el compromiso asumido por la comunidad internacional: no dejar atrás a nadie. Merece destacar que el enfoque mencionado constantemente ha sido invocado al interior de los organismos de la ONU a los que se les ha asignado la tarea de abocarse a todo lo relacionado con el desarrollo, la innovación tecnológica, etcétera, pero sobre todo merece subrayar que se trata de una perspectiva con la que se anhela contribuir al mayor disfrute de los derechos humanos.

En esta Resolución E/CN.16/2024/3: Cooperación mundial en ciencia, tecnología e innovación para el desarrollo, particularmente en el cuerpo del Informe, se localiza el apartado intitulado 1B relativo a los bienes públicos digitales. A este tipo de bienes, se les asigna la cualidad de primordiales de cara a la concreción de los Objetivos de Desarrollo Sostenible (ODS), pero se hace énfasis en el caso particular de aquellos países que representan bajos

o medianos ingresos. En efecto, el papel de la ética y la IA, en el caso específico de los ODS, ha tenido una importancia singular, pues más allá del innegable impacto favorable de la segunda en estos, para González y Ballesteros (2024, p. 42) lo que está en juego es canalizar los avances de la IA de la mano de la ética, a efecto de que los propios ODS se beneficien de la transferencia de lo sabido.

En un ejercicio de repaso histórico, el SG recuerda que el internet, en sus inicios, surgió como una red gestionada públicamente en sincronía con una ética de código abierto; esto último, según se lee en el Informe, fomentó un esquema de colaboración entre los involucrados, además de que detonó las condiciones favorables para la experimentación en la materia. Desafortunadamente —se advierte—, con el paso de las décadas, el internet de código abierto al público prácticamente es muy inferior respecto del internet en manos de empresas privadas, que operan bajo la lógica del máximo rédito. Esta situación imposibilita a los más afectados el acceso a una significativa parte de información útil que se encuentra en la red.

b) El Informe de la Comisión de Ciencia y Tecnología para el Desarrollo (Informe sobre el 26.º Periodo de Sesiones, del 27 al 31 de marzo de 2023) (Consejo Económico y Social, 2023a).

Del contenido de este documento merece resaltar el capítulo I, inciso A, fracción I, intitulado “Evaluación de los progresos realizados en la aplicación y el seguimiento de los resultados de la Cumbre Mundial sobre la Sociedad de la Información”, publicado después bajo el mismo título como resolución E/RES/2023/3, aprobada por el Ecosoc el 7 de junio de 2023 (Consejo Económico y Social, 2023b).

Se replantean algunas posturas del Ecosoc y una de ellas tiene que ver con la persistencia de las llamadas brechas digitales verificadas tanto entre las naciones como al interior de ellas, además de la que se presenta entre hombres y mujeres dentro de sus comunidades. Lo que plantea el Ecosoc es que estas brechas tienen que resolverse a través del fortalecimiento de los sistemas jurídicos propicios, pero también mediante la cooperación internacional, con el propósito de concretar, en la mayor medida posible, el acceso, la educación, la generación de capacidades, la asequibilidad, el multilingüismo, la conservación de la cultura, el financiamiento adecuado y la necesaria inversión por parte de los gobiernos. No obstante, reconoce que sigue presente una brecha digital entre los géneros que forman parte de estas, por lo que exhorta a todos los interesados a que velen por que las mujeres y las niñas participen plenamente en la sociedad de la información, y dispongan de acceso a las nuevas tecnologías, sobre todo las tecnologías de

la información y las comunicaciones (TIC) para el desarrollo, combatiendo tanto la violencia de género facilitada por la tecnología como la propia explotación, el acoso y los abusos contra las mujeres y las niñas (Consejo Económico y Social, 2023a, p. 9).

Asimismo, el Ecosoc continúa insistiendo en las cuestiones éticas que surgen en el terreno particular de la aplicación de la IA, pues admite que el cúmulo de todos estos progresos tecnológicos —a los que constantemente alude— siempre ha estado acompañado de conflictos éticos, sin omitir la clara amenaza que para este organismo representa en términos de derechos humanos. En virtud de lo anterior acoge con beneplácito la Recomendación sobre la Ética de la Inteligencia Artificial (REIA), de la Unesco, aprobada el 23 de noviembre de 2021, de la cual nos ocuparemos más adelante (Consejo Económico y Social, 2023a, p. 14).

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

La Unesco se enfocará a tres fuentes: a) Recomendación sobre la Ética de la Inteligencia Artificial (REIA); b) Declaración hecha por la directora general Adjunta de la Unesco para las Ciencias Sociales y Humanas; y c) Actividades derivadas del Foro Global sobre la Ética de la Inteligencia Artificial, celebradas en el año 2023 y 2024.

a) Recomendación sobre la Ética de la Inteligencia Artificial (REIA)

La REIA, sin duda, se trata de la fuente más importante formulada y publicada por un organismo subsidiario perteneciente al Sistema ONU. En palabras de Reyes,

En su conjunto es un marco integrador de valores, principios y acciones que persigue guiar el desarrollo de las TIA's en una dirección enfocada en la dignidad humana, los derechos humanos, así como también en la igualdad de género, la justicia social y el desarrollo económico, el bienestar físico y mental de las personas, la diversidad cultural, la inclusión, la interconectividad y la protección del medio natural y los ecosistemas. (2023, p. 4)

Por la fecha de su adopción (23 de noviembre de 2021) constituye un instrumento internacional que aparece en un contexto en el que el concepto de disrupción tecnológica cobraba la mayor importancia, debido a las condiciones de emergencia sanitaria por el COVID-19, situación que aceleró el uso de estas herramientas, plataformas, aplicaciones, etcétera, todas ellas fundamentalmente operadas mediante IA.

En amplio sentido, se trata de un documento en el que se congregan, entre otras cosas, las mejores prácticas identificadas, se adopta una reflexión normativa siste-

mática y se tiene como punto de partida un marco integral, global multicultural y evolutivo de valores, principios y acciones interdependientes. En consideración de la propia Unesco, el marco anteriormente aludido puede servir como una guía fundamental para los miembros de la comunidad internacional a efecto de orientar la toma de decisiones, así como las acciones de cara a los efectos visibles, previsibles e, incluso, desconocidos de la IA en una multiplicidad de ámbitos: en la propia humanidad, en el medio ambiente, en los ecosistemas, en la dimensión laboral, etcétera. En la REIA se adopta una concepción de 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. (Unesco, 2021, p. 10)

Con la publicación de la REIA, la Unesco busca ofrecer un instrumento normativo moral que articule valores y principios con impacto en la comunidad internacional. Se aspira —con el conjunto de criterios éticos y metodológicos imbricados en el texto— a funcionar como una especie de guía que permita operar la IA siempre en beneficio de los seres humanos, de los propios Estados, del medio ambiente y de

los ecosistemas, entre otros. Ahora bien, vale la pena destacar que, tanto en la sede de la ONU como en el propio Sistema que la conforma, siempre ha estado presente la constante preocupación respecto de las tecnologías en general y la IA en específico y su uso bélico, preocupación que también se hace manifiesta en la fuente aquí reseñada. Como bien lo menciona Drnas de Clement en un artículo intitulado “Inteligencia artificial en el derecho internacional, Naciones Unidas y Unión Europea”, al señalar que: “Las implicaciones éticas de las tecnologías basadas en la IA desdibujan la frontera entre los sujetos humanos y los objetos tecnológicos, afectando las categorías centrales de la ética: la percepción del actor, la responsabilidad y los marcos de valores” (2022, p. 5).

Derivado de lo anterior, la Unesco apela a que, con la adopción de los principios éticos contemplados en la REIA, se fomente el uso pacífico de las tecnologías y, en particular, la IA. Se contemplan cuatro valores en el documento, los cuales están explicitados en rubros y se describen a continuación:

- a) Respeto, protección y promoción, tanto de los derechos humanos, las libertades fundamentales y la propia dignidad humana;
- b) Prosperidad en materia medioambiental/ecosistemas;

- c) Garantía de diversidad (cultural, de ideas, etcétera) e inclusión (social, principalmente); y
- d) Vida en comunidades pacíficas, justas y, además, interconectadas.

Mientras que los principios adoptados son los de:

- 1) Proporcionalidad e inocuidad; 2) Seguridad y protección; 3) Equidad y no discriminación; 4) Sostenibilidad; 5) Derecho a la intimidad y protección de datos; 6) Supervisión y decisión humanas; 7) Transparencia y explicabilidad; 8) Responsabilidad y rendición de cuentas; 9) Sensibilización y educación; y 10) Gobernanza y colaboración adaptativas y de múltiples partes interesadas.

En términos generales, la REIA constituye un parteaguas en el contexto de una sociedad que ha experimentado un tránsito acelerado hacia las tecnologías digitales. En esos escenarios de disrupción tecnológica aludidos en este ensayo, la IA ha sido uno de los temas centrales en la agenda de la ONU y sus organismos subsidiarios (Sistema ONU). Sin embargo, como todo desarrollo tecnológico jamás emerge exento de peligros y amenazas —más allá de sus bondades y promesas—, la Unesco se dio a la tarea de concentrar sus esfuerzos en la elaboración de un documento de largo aliento que permee en la toma de decisio-

nes entre quienes forman parte de la comunidad global, y qué mejor manera de comenzar por el diseño de un conjunto de criterios éticos que marquen la pauta en ello, como la REIA.

No obstante las respuestas y el consenso que la REIA ha podido generar, aún persisten voces que invitan a preguntarse, dentro de las coordenadas del Sistema Universal de Derechos Humanos, la cuestión de la suficiencia del enfoque de derechos humanos. Así, por ejemplo, para Asís (2022, pp. 35-36) no resulta suficiente adoptar como referencia los derechos humanos, sino que, en todo caso, se antoja necesario una teoría de los derechos, en la que los asuntos clásicos de esta (concepto y fundamento) adquieran una nueva relevancia. El autor se enfoca en el posicionamiento hecho por Michelle Bachelet, la Alta Comisionada de las Naciones Unidas para los Derechos Humanos, en su discurso pronunciado el 17 octubre de 2019 intitulado “Derechos humanos en la era digital. ¿Pueden marcar la diferencia?”. Ante los desafíos que derivan de la revolución digital, básicamente la Alta Comisionada comparte el dilema de si estos se deben enfrentar desde la ética o los derechos humanos. Para ella todos los segmentos de la denominada “revolución digital” pueden y deben analizarse desde el enfoque de los derechos humanos, puesto que —asegura— los códigos éticos y el cumplimiento voluntario no constituyen una garantía

suficiente y contundente de cara a los profundos problemas a los que nos enfrentamos, sobre todo en el terreno de la IA.

b) Posicionamiento de la Dirección General Adjunta de la Unesco para las Ciencias Sociales y Humanas

Como ya se señaló, el progreso tecnológico siempre viene acompañado de amenazas que denotan (y detonan) nuestras vulnerabilidades, resaltan las deficiencias de los sistemas jurídicos, visibilizan sesgos, perpetúan estereotipos o profundizan contrastes sociales, entre otros aspectos. Es inevitable que todos estos factores generen desconfianza y temores, pero sobre todo resulta insoslayable que los diversos actores políticos y sociales asuman posicionamientos con el propósito de mejorar la zozobra social y, con ello, dar señales sobre los aspectos que les resulta prioritarios. Este es el caso de la titular de la Dirección General Adjunta de la Unesco para las Ciencias Sociales y Humanas.

Desde la perspectiva de esta área de la Unesco, la situación de emergencia sanitaria que se vivió en todo el planeta por el COVID-19 fue un escenario fértil para la manifestación de una serie de actos que atentaron claramente contra la integridad de las personas y contra la dignidad. Además de aquellas expresiones, la Dirección General Adjunta advirtió sobre los múltiples actos y conductas discriminatorios que se observaron en el espectro digital, pero

también racalcó el fenómeno de la desinformación en virtud del cual el concepto de “infodemia” retomó una singular importancia: los lamentables —pero también controvertidos— discursos de odio, la violencia digital, entre muchos más. Las expresiones descritas afectaron —y continúan afectando— negativamente a grupos concretos plenamente identificables.

En razón de lo anterior, Gabriela Ramos, titular de la Dirección General Adjunta, señaló que si se sigue haciendo uso de las tecnologías sin una base ética que apunte a evitar daños, actos de discriminación, sesgos en los algoritmos y prejuicios, todas estas expresiones inevitablemente se magnificarán. De ahí el exhorto que ella hiciera en el sentido de que las autoridades de los Estados que conforman la comunidad internacional inviertan sus esfuerzos a efecto de que, en el desarrollo e implementación de estas tecnologías, se procure evitar aquellas manifestaciones de la vida cotidiana que no resultan deseables, o que simplemente se repudian porque constituyen, a todas luces, una amenaza y un desafío a la protección de los derechos humanos. El objetivo —para la titular de la Dirección General Adjunta— es invertir esfuerzos para construir sociedades más inclusivas y no discriminatorias. Parte del exhorto también se dirigió a diseñar una propuesta que contemple la creación de comités nacionales y regionales, cuya función principal sea monitorear y dar segui-

miento respecto de la manera en cómo los diversos países y sus gobiernos han estado implementando la REIA (Naciones Unidas México, 2022). Para el caso específico de México, la propia Unesco, en voz de Gabriela Ramos, señala que México:

enfrenta enormes desafíos para lograr el desarrollo ético de la IA. El gobierno de México empezó a participar activamente en los foros internacionales sobre estos temas desde 2018, cuando puso en marcha una iniciativa para elaborar una estrategia nacional digital que considerara explícitamente a la IA. Formó parte de los grupos de trabajo en la ONU y asumió el liderazgo de varias iniciativas en la Red de América Latina y el Caribe para el Desarrollo de Gobiernos Digitales. Sin embargo, desde entonces otros países con niveles similares de desarrollo han logrado avances importantes, mientras que México se ha alejado de sus posiciones de liderazgo en la materia. (2024, p. 8)

La anterior afirmación tiene su fundamento en la iniciativa de la Unesco formulada para comparar las condiciones de las naciones de la región de América Latina, particularmente el Índice Latinoamericano de IA, el cual, en su edición correspondiente al año 2023, informa que el nivel asignado a México es de “muy bajo” en lo que respecta a Visión e Institucionalidad: el puntaje es de 2.78 frente a la media LA-

TAM, que es de 33.68; mientras que en el rubro Estrategia obtuvo una puntuación de 8.33 frente a la media LATAM de 35.41, por lo que no titubea en concluir que México “... no cuenta con una estrategia de IA vigente y no ha puesto en marcha mecanismos con participación social para avanzar en esta materia” (p. 8).

c) Foro Global sobre la Ética de la Inteligencia Artificial

Dado que la Unesco asume la IA como una tecnología de vanguardia que advierte profundas implicaciones para los seres humanos, las culturas y el medio ambiente, diseñó este Foro como un espacio de discusión, diálogo, intercambio de experiencias, entre otros, en torno a los aspectos éticos que están en juego durante su diseño e implementación en todos los ámbitos del conocimiento: salud, derecho, medio ambiente, etcétera. La reconoce como una aliada en la búsqueda de un futuro más equitativo, justo y, por supuesto, sostenible. Todo ello en virtud de su capacidad analítica sin precedente, sobre todo en la toma de decisiones en los múltiples aspectos de la vida diaria. Como muestra de lo anterior, este organismo especializado de la ONU alude a la experiencia durante la pandemia por el COVID-19 y el impacto de la IA en el descubrimiento de las vacunas. No obstante, lo maravilloso y útil que pueda resultar esta herramienta en muchos aspectos, para esta entidad internacional su

uso siempre viene acompañado de enormes riesgos y desafíos de amplio impacto en los derechos humanos y las libertades fundamentales.

El Primer Foro Global sobre la Ética de la Inteligencia Artificial, se intituló “Garantizar la inclusión en el mundo de la IA” y se fraguó como un punto de encuentro para la construcción de una alianza internacional sólida enfocada en garantizar el desarrollo ético y su uso en el mundo. Se verificó el 13 de diciembre de 2022, con sede en Praga, Checoslovaquia. El programa se desarrolló en torno a tres ejes temáticos: a) Desarrollo ético y uso de la IA en la Unión Europea (UE); b) Colaboración global entre la UE y la Unesco para la implementación de la REIA; y c) Temas especiales, en los que se abordaron cuestiones sobre la exploración del estado del arte de la IA desde la perspectiva de la igualdad de género, protección ambiental, impacto ético, transparencia y el derecho humano a la no discriminación.

En términos generales, se hizo un balance de todos aquellos esfuerzos en cada una de las naciones en torno a la promoción del desarrollo ético y el uso de la IA; asimismo, se hizo un ejercicio de identificación de las mejores prácticas regulatorias y los entornos institucionales que apuntaran a garantizar el desarrollo ético de esta herramienta. Los participantes compartieron sus análisis sobre la manera en que la Recomendación de la Unesco podría

servir de apoyo a los esfuerzos de la UE para promover el desarrollo tecnológico basado en valores, de tal modo que estos esfuerzos regionales se hagan extensivos. De hecho, estos aspectos abordados en el Primer Foro, en cierta manera, son el eco de aquellas inquietudes que ya se habían venido socializando en el propio contexto del derecho internacional de los derechos humanos y, de acuerdo con López y García (2024, p. 118), una de ellas es, precisamente, los sesgos algorítmicos y los desafíos éticos y sociales que genera; de ahí la profunda responsabilidad que los autores le asignan a los desarrolladores de IA, a efecto de que consideren los aspectos éticos implícitos en sus tecnologías.

El Segundo Foro Global sobre la Ética de la Inteligencia Artificial, celebrado en Kranj, Eslovenia, los días 5 y 6 de febrero de 2024, se intituló “Cambiando el panorama de la gobernanza de la IA”. En la segunda edición del Foro compartieron sus conocimientos un grupo de expertos en exploración de oportunidades y desafíos generados por la IA: equidad, diversidad, derecho humano a la no discriminación, adopción de mejores prácticas emergentes de supervisión, asociaciones con el sector privado por medio de evaluaciones de impacto ético e impacto en lo concerniente a la igualdad de género.

Para este segundo encuentro, se diseñó un amplio programa de actividades, pero por cuestiones de espacio solo se descri-

ben las más destacadas, de acuerdo con los propósitos de este documento. Se insistió en el compromiso de la Unesco de promover la gobernanza ética de la IA más allá de la Recomendación, lo que implica traducir sus principios en políticas legislativas y públicas viables. En virtud de lo anterior, se presentaron ideas de aquellos países pioneros (más de cincuenta) en adoptar dichos principios en políticas viables como resultado de la aplicación de la Metodología de Evaluación de Preparación (RAM, por sus siglas en inglés). Para poner en contexto, la REIA, en su párrafo 49, reconoce los diversos estadios de preparación en que se encuentran los Estados Miembros, para aplicarla desde diferentes puntos de vista (científico, tecnológico, económico, entre otros).

En virtud de lo anterior y con el propósito de facilitar la aplicación efectiva de la REIA, la Unesco elaboró una metodología de evaluación acerca del estadio de preparación —de avance— en el que se ubican los Estados y que, además, permita determinar su situación en circunstancias concretas de toda su trayectoria de preparación por medio de un abanico de dimensiones (Unesco, 2021, p. 26). La metodología RAM es un instrumento macro mediante el cual se busca orientar a las naciones a diagnosticar en qué grado de preparación se encuentran estas en determinado momento para aplicar la IA de manera ética y responsable para la ciudadanía en su totalidad,

permitiendo con ello tener claro los cambios institucionales y legales que requerirían implementar (Unesco, 2023, p. 6).

El jueves 26 de octubre de 2023, el secretario general de la ONU, António Guterres, convocó y dio a conocer la conformación de un grupo consultivo de expertos sobre IA denominado Órgano Asesor de Alto Nivel sobre IA (HLAB, por sus siglas en inglés). El objetivo es que el grupo multidisciplinario de expertos preste asistencia a la comunidad internacional en la tarea que esta tiene de gestionar todo lo concerniente al desarrollo de este tipo de tecnología y su impacto en la gobernanza internacional. Los perfiles que lo integran provienen del sector público, privado, sociedad civil y académico. En el acto de presentación el secretario general afirmó que:

Para las economías en desarrollo, la inteligencia artificial abre la posibilidad de superar tecnologías obsoletas y servir directamente a las personas que más lo necesitan. El potencial de transformación positiva que ofrece la inteligencia artificial resulta todavía difícil de asimilar. Y sin entrar a considerar posibles escenarios catastróficos, ya ha quedado claro que el uso malintencionado de la inteligencia artificial puede socavar la confianza en las instituciones, debilitar la cohesión social y amenazar la democracia en sí misma.

Por estos motivos, he puesto en marcha una conversación mundial y multidisciplinar, en la que participen todas las partes interesadas, sobre la gobernanza de la inteligencia artificial, con el objetivo de maximizar los aspectos beneficiosos para la humanidad, para toda la humanidad, y contener y minimizar los riesgos. (ONU. Secretaría General, 2023)

A este órgano consultivo de alto nivel sobre IA, se le asignó la encomienda de elaborar y publicar un documento provisional que incluyera análisis, diagnósticos, es decir, lo relacionado con el estado del arte en torno a la IA y la gobernanza. Como resultado, el grupo publicó en diciembre de 2010 un informe provisional (ONU. *Advisory Body on Artificial Intelligence*, 2023). Dentro de las actividades del Segundo Foro, el lunes 5 de febrero de 2024 se llevó a cabo un encuentro entre la enviada del secretario general para Tecnología, Amandeep Gill, con los miembros del HLAB, a efecto de entablar un diálogo abierto entre ambas partes sobre cuestiones asociadas a la gobernanza de la IA.

Asimismo, se hizo público el lanzamiento del Observatorio Global de Ética de la Inteligencia Artificial, el cual se erige, en colaboración con el Alan Turing y la URT, como una plataforma que tiene como objetivo, entre otros, servir como una especie de epicentro de conocimientos sobre ética y gobernanza de la IA. Se le asigna la

responsabilidad de presentar los informes de la RAM, análisis de las mejores prácticas respecto de la gobernanza de la IA de una multiplicidad de países, además de análisis sobre las mejores prácticas sobre gobernanza de la IA en las diferentes redes, plataformas y asociaciones vinculadas a la Unesco. También, se hizo público el lanzamiento de la Red de Expertos en Ética en Inteligencia Artificial Sin Fronteras, perteneciente a la Unesco. Esta Red fue diseñada en colaboración con la Comisión Europea, con el propósito de apoyar a las naciones menos desarrolladas en la elaboración de sus agendas legislativas. Se le asignó un presupuesto de cinco millones de euros para su funcionamiento (Unesco, 2024, p. 19).

En términos generales, estos son los aspectos más destacables de la segunda versión del Foro, en cuya agenda se programaron algunos otros que también abordan aquellas cuestiones de la gobernanza de la IA.

RESULTADOS

Uno de los resultados de mayor trascendencia respecto de la IA y los derroteros éticos que están en juego son los retos a los que se enfrenta, tanto el propio órgano supranacional como el conjunto de organismos subsidiarios que conforman el Sistema ONU, de dar cauce a este tipo de herramientas digitales en todo lo que beneficie a la humanidad y al medio ambiente, en-

tre otras variables. Es innegable que todo el conjunto de acciones e instrumentos aquí compartidos constituyen un avance significativo; sin embargo, para Silva, Guerrero y González (2024, p. 91) aún falta mucho camino por recorrer en el ámbito específico de la agenda legislativa de los derechos humanos, pues, por paradójico que resulte, aún se siguen explorando terrenos desconocidos.

Resulta claro que los instrumentos (fuentes) adoptadas y las acciones previamente descritas son predominantemente de naturaleza *soft law*. De hecho, en términos generales, Rudas hace un repaso de aquellos instrumentos internacionales que han sido piedra de toque en la temática abordada en este artículo y señala que:

La Declaración Universal de los Derechos Humanos (DUDH) de 1948 reconoce el derecho a la privacidad y la no discriminación. Asimismo, conforme con el *soft law*, la OCDE estableció las Directrices sobre Inteligencia Artificial en el 2019 para promover la IA ética y segura, con énfasis en la transparencia, la responsabilidad y la gobernanza. Por su parte, la Asamblea General de la ONU, a través de sus resoluciones, aborda la IA y sus implicaciones, incluida la necesidad de proteger los derechos humanos en su desarrollo y uso. (2024, p. 125)

Lo anterior, para algunos podría implicar un alcance limitado. Esto no es del todo cierto por dos razones. La primera tiene que ver con el hecho de que, al hablar de ética e inteligencia artificial, la temática se coloca necesariamente en el mundo de la filosofía moral y no precisamente en una disertación jurídica de naturaleza aplicada, que tenga como propósito proponer reformas para, con ello, colmar las lagunas legales correspondientes. La segunda razón se dirige a señalar la importancia que, en el caso de México, han retomado en el ámbito jurisdiccional los instrumentos internacionales de naturaleza no vinculante, particularmente en la sede del Poder Judicial federal: Pleno, Primera Sala y Segunda Sala.

Merece subrayar que cuando nos referimos a instrumentos internacionales no vinculantes, la Suprema Corte de Justicia de la Nación en México ha dado cabida no solo a las declaraciones internacionales, sino a un conjunto de fuentes que no necesariamente revisten esa cualidad de declaración. Pero esto es tema para otro espacio, de tal modo que me limitaré a remarcar la magnitud de la importancia que reviste en los textos referidos en este ensayo, así como las acciones (los foros) verificadas en el seno del Sistema ONU, todos ellos encaminados, como se mencionó, a proporcionar un mínimo de coordenadas éticas que permitan guiar las acciones de

quienes se encuentren involucrados en todo lo concerniente a la IA.

CONCLUSIONES

Resulta procedente concluir sobre el papel determinante de la ONU y sus organismos subsidiarios que integran el Sistema ONU, frente a las amenazas a los derechos humanos que implica la IA, sobre todo en la formulación de criterios éticos que marquen el rumbo a seguir en el seno de la comunidad internacional.

En el caso de ciertas organizaciones internacionales, como la ONU, la Unesco, órganos de expertos, espacios de encuentro para la discusión e intercambio de experiencias en la asignatura de la IA, sus resoluciones, acuerdos, informes, etcétera, todo ello se traduce, indiscutiblemente, en fuentes de futuras normas jurídicas para los Estados. Incluso, aquellas resoluciones de todas las entidades descritas en este texto, aun cuando tengan solamente una naturaleza recomendatoria, ello no es óbice para que puedan constituir referentes en la conformación tanto de criterios éticos al interior de los países como del derecho transfronterizo. Es innegable que el conjunto de acciones, documentos y resoluciones emanados de las instituciones aquí abordadas contribuyen al nacimiento de normas, códigos deontológicos y, por tanto, se asumen como medios auxiliares para la determinación de la existencia y contenido de los derechos humanos.

REFERENCIAS

- Aneja, U. (2021). La gobernanza de la inteligencia artificial: de solucionar los problemas a diagnosticarlos. *Anuario Internacional CIDOB*, 1, 28-35. https://www.cidob.org/articulos/anuario_internacional_cidob/2021/la_gobernanza_de_la_inteligencia_artificial_de_solucionar_los_problemas_a_diagnosticarlos
- Asís, R. (2022). Ética, tecnología y derechos. En J. Garrido Martín y R. Valdivia Jiménez (Coords.), *Inteligencia artificial y filosofía del derecho* (pp. 25-44). Laborum Ediciones.
- Benavides Casals, M. A. (2021). El *soft law* en el contexto de las fuentes del derecho internacional: aportes para la discusión. *Revista de Derecho Público*, 94, 1-16. <https://doi.org/10.5354/0719-5249.2021.64239>
- Drnas de Clement, Z. (2022). Inteligencia artificial en el derecho internacional, Naciones Unidas y Unión Europea. *Revista Estudios Jurídicos*, Segunda Época, 22(22), 1-28. <https://doi.org/10.17561/rej.n22.7524>
- Escudero Alday, R. (2012). El concepto de *soft law*. En J. Moreso y J. L. Martí (Eds.), *Contribuciones a la filosofía del derecho* (pp. 127-148). Marcial Pons.
- González G., J. y Ballesteros, J. (2024). Ecos de sostenibilidad: reflexiones sobre la inteligencia artificial a la luz de los Objetivos de Desarrollo Sostenible. *EDETANIA*, 65, 27-45. https://doi.org/10.46583/edetania_2024.65.1142
- Laporta, F. (2014). Gobernanza y *soft-law*: nuevos perfiles jurídicos de la sociedad interna-

- cional. En A. Ruiz Miguel (Ed.), *Entre Estado y cosmópolis. Derecho y justicia en un mundo global* (pp. 41-81). Trotta.
- López, F. y García, J. H. (2024). IA y sesgos: una visión alternativa expresada desde la ética y el derecho. *Informática y Derecho. Revista Iberoamericana de Derecho Informático*, 15(1), 109-121. <https://revistas.fcu.com.uy/index.php/informaticayderecho/article/view/4738>
- Naciones Unidas México. (2022). Unesco pide detener la discriminación a través del mal uso de la inteligencia artificial. Recuperado el 23 de junio de 2022, de <https://mexico.un.org/es/187451-unesco-pide-detener-la-discriminacion-trav%C3%A9s-del-mal-uso-de-la-inteligencia-artificial>
- Organización de las Naciones Unidas (ONU). Advisory Body on Artificial Intelligence. (2023). Interim Report: Governing AI for Humanity. https://www.un.org/sites/un2.un.org/files/un_ai_advisory_body_governing_ai_for_humanity_interim_report.pdf
- Secretaría General. (2023). *El secretario general de las Naciones Unidas crea el "Órgano Asesor sobre Inteligencia Artificial" para tratar los riesgos, las oportunidades y la gobernanza internacional de estas tecnologías*. Comunicado de prensa. https://www.un.org/sites/un2.un.org/files/organo-asesor-ia_prensa_final.pdf
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (Unesco). (2021). Recomendación sobre la Ética de la Inteligencia Artificial. <https://www.unesco.org/es/articles/recomendacion-sobre-la-etica-de-la-inteligencia-artificial>
- (2023). Metodología de evaluación del estadio de preparación. Una herramienta de la Recomendación sobre la Ética de la Inteligencia Artificial. https://unesdoc.unesco.org/ark:/48223/pf0000385198_spa
- (2024). La Unesco en imágenes. https://unesdoc.unesco.org/ark:/48223/pf0000389050_spa
- (2024). México. Evaluación del estadio de preparación de la inteligencia artificial. <https://mexico.un.org/sites/default/files/202407/México.%20Evaluación%20del%20estadio%20de%20preparación%20de%20la%20inteligencia%20artificial.pdf>
- Pérez, A. (2018). Ciudadanía y gobernanza digital. Entre política, ética y derecho. En M. Barrio y J. Torregrosa (Coords.), *Sociedad Digital y Derecho* (pp. 137-155). Boletín Oficial del Estado; Ministerio de Industria, Comercio y Turismo; RED.ES.
- Restrepo Medina, M. A. (2018). Valor jurídico de la Declaración Universal de los Derechos Humanos en el derecho internacional de los derechos humanos. En C. Proner *et al.* (Coords.), *70.º aniversario de la Declaración Universal de los Derechos Humanos. La protección internacional de los derechos humanos en cuestión* (pp. 453-456). Tirant Lo Blanch.
- Reyes, P. (2023). Ética de la inteligencia artificial. Recomendación de la Unesco, noviembre 2021. *Compendium*, 26(50), 1-6. <https://doi.org/10.5281/zenodo.10271853>

- Rodrigo Hernández, A. J. (2022). El valor normativo de la Agenda 2030. En C. Fernández Liesa, E. López-Jacoiste y D. Oliva (Eds.), *El derecho internacional, los ODS y la comunidad internacional* (pp. 53-77). Dykinson.
- Rudas, C. (2024). La inteligencia artificial y los derechos humanos. *Informática y Derecho. Revista Iberoamericana de Derecho Informático*, 15(1), 123-135. <https://revistas.fcu.edu.uy/index.php/informaticayderecho/article/view/4739/4160>
- Silva, S., Guerrero, J. y González, J. M. (2024). Ética en la era digital. *Elementos*, 31(135), 89-94. <https://elementos.buap.mx/directus/storage/uploads/00000009972.pdf>
- (2024). Comisión de Ciencia y Tecnología para el Desarrollo. (2024). Resolución E/CN.16/2024/3: Cooperación mundial en ciencia, tecnología e innovación para el desarrollo, 1-18. https://unctad.org/system/files/official-document/ecn162024d3_es.pdf
- (2023b). Resolución E/RES/2023/3: Evaluación de los progresos realizados en la aplicación y el seguimiento de los resultados de la Cumbre Mundial sobre la Sociedad de la Información, 1-16. <https://documents.un.org/doc/undoc/gen/n23/171/28/pdf/n2317128.pdf?token=lrt2H2GEXICXT9sFaE&fe=true>

Resoluciones ONU

- Asamblea General de la Organización de las Naciones Unidas. (2024). Resolución A/74/821: Hoja de ruta para la cooperación digital: aplicación de las recomendaciones del Panel de Alto Nivel sobre la Cooperación Digital, 1-22. <https://documents.un.org/doc/undoc/gen/n20/102/54/pdf/n2010254.pdf?token=EiJQvGNpdUigJ83yEY&fe=true>
- Consejo Económico y Social. (2023a). Resolución E/2023/31-E/CN.16/2023/4: Comisión de Ciencia y Tecnología para el Desarrollo. Informe sobre el 26.º Periodo de Sesiones, del 27 al 31 de marzo de 2023. Documentos Oficiales, Suplemento núm. 11, 1-49. https://unctad.org/system/files/official-document/ecn162023d4_es.pdf
- (2023b). Resolución E/RES/2023/3: Evaluación de los progresos realizados en la aplica-