

La Inteligencia Artificial y la Integridad de la Información en el marco de las Organizaciones Internacionales

Artificial Intelligence and Information Integrity in the Context of International Organizations

*Dr. Melvin Uziel Porras Reynoso**

Resumen

En los últimos años, ha habido un auge en el área de la Inteligencia Artificial, ello ha conllevado una serie de beneficio en pro de la humanidad, pero al mismo tiempo ha traído consigo riesgos muy grandes como lo es: la manipulación y alteración de información por medio de la IA, que puede traer consecuencias catastróficas. Razón por la cual, el presente artículo hace un análisis de las implicaciones del derecho a la información y la integridad de la misma, además de identificar cuál es la postura y acciones emprendidas por los organismos internacionales para atender estos riesgos.

Palabras Clave: Inteligencia Artificial, Integridad de la Información, Organizaciones Internacionales.

Abstract

In recent years, the field of Artificial Intelligence has seen a surge in development, bringing numerous benefits to humanity. However, it has also introduced significant risks, such as the manipulation and alteration of information through AI, which can have catastrophic consequences. Therefore, this article analyzes the implications of AI for the right to information and its integrity, and identifies the positions and actions taken by international organizations to address these risks.

Keywords: Artificial Intelligence, Information Integrity, International Organizations.

*Profesor-investigador de la Benemérita Universidad Autónoma de Puebla (BUAP). Miembro del Sistema Nacional de Investigadoras e Investigadores. Secretaría de Ciencia Humanidades, Tecnología e Innovación. México. Doctor en Derecho por la BUAP. Estancia posdoctoral en la Universidad Autónoma de Nuevo León. Con formación como licenciado en Derecho. Secretario formado por el Instituto de la Judicatura Federal, hoy Escuela Federal de Formación Judicial. Áreas de investigación: Derecho Internacional y Derechos Humanos. Miembro de la Academia en Derecho Económico e Internacional de la Licenciatura en Derecho. Miembro de la Academia de la Maestría en Relaciones Internacionales y Derechos Humanos, ambos de la BUAP. Correo electrónico: melvin.porrasre@correo.buap.mx

1 Introducción

Sin lugar a duda la Inteligencia Artificial es una herramienta muy útil para eficientar procesos, análisis de datos e información, pero a la vez ha originado un gran número de riesgos entre los cuales puede destacar el uso de la IA para fines de desinformación, uso indebido de imágenes y voz de las personas, extorsiones, fraudes, etc.

En el momento histórico en el que nos encontramos, las redes sociales y la información que circula en ellas tiene un gran espectro de audiencia en el cual se comparten fake news (noticias faltas) y deepfakes (imágenes, videos o audios falsos o manipulados, creados con IA), lo cual es sumamente peligroso, ya que pueden atentar contra la reputación y dignidad de las personas, y sobre todo la incertidumbre de las personas en poder saber si el contenido de redes es verídico o si fue realizado con inteligencia artificial, lo que trasgrede el derecho humano a la información y de su integridad; siendo esto un problema que los Estados y la Comunidad Internacional deben de atender.

La justificación de la presente investigación deriva de los inminentes riesgos de la IA, por lo que, resulta imperativo hacer un estudio que los principales organismos internacionales como la ONU, la UNESCO, la Organización para la Cooperación y el Desarrollo Económicos (OCDE) y la Unión Europea (UE) entorno a la Inteligencia Artificial, para poder identificar los parámetros, directrices, acciones y recomendaciones que han adoptado para poder regular este fenómeno.

Finalmente, el estudio del Soft Law de los organismos internacionales es trascendental ya que son medidas que deben tomar los Estado en su territorio, y al conocer estos estándares internacionales pueden materializarlo en su marco normativo o en políticas estatales para poder contrarrestar los efectos negativos de la IA.

2 Integridad de la Información

La integridad de la información se refiere a la exactitud, coherencia y fiabilidad de la información. Esta se ve amenazada por la desinformación, la información errónea y el discurso de odio (Naciones Unidas, 2023: 5).

La desinformación y la información errónea son creadas por una amplia gama de actores, con distintas motivaciones, que en general pueden permanecer anónimos. Las campañas de desinformación coordinadas por actores estatales y no estatales han explotado sistemas digitales defectuosos para promover narrativas dañinas, con graves repercusiones (Naciones Unidas, 2023: 6).

En la era digital en el que la sociedad sigue tendencias sin detenerse a cuestionar si lo que se está compartiendo es verídico, pues el contenido realizado por IA puede ser difícil de identificar, y cuando esta contiene información falsa o manipulada, se trasgrede la integridad de la información. No cabe duda que los programas de inteligencia artificial deben ser regulados adecuadamente a través de un marco jurídico idóneo y políticas estatales que las respalden, con el objeto de garantizar la integridad de la información y reducir al mínimo los riesgos y consecuencias que trae consigo las noticias falsas o manipuladas por IA, para la sociedad.

El Derecho a la información lo tenemos consagrado en diversos tratados internacionales como podemos ver a continuación: La Declaración Universal de los Derechos Humanos en su artículo 19 reconoce el derecho a la información, así como el de libertad de expresión: “Todo individuo tiene derecho a la libertad de opinión y de expresión; este derecho incluye el de no ser molestado a causa de sus opiniones, el de investigar y recibir informaciones y opiniones, y el de difundirlas, sin limitación de fronteras, por cualquier medio de expresión”.

Por su parte, el Pacto Internacional de Derechos Civiles y Políticos reconoce el derecho a la libertad de expresión y a su vez buscar, recibir y difundir información: “*Artículo 19. 1. Nadie podrá ser molestado a causa de sus opiniones. 2. Toda persona tiene derecho a la libertad de expresión; este derecho comprende la libertad de buscar, recibir y difundir informaciones e ideas de toda índole, sin consideración de fronteras, ya sea oralmente, por escrito o en forma impresa o artística, o por cualquier otro procedimiento de su elección*”.

Es esa misma disposición se establecen restricciones a la libertad de expresión las cuales son: “*a) Asegurar el respeto a los derechos o a la reputación de los demás; b) La protección de la seguridad nacional, el orden público o la salud o la moral públicas*”.

También, el artículo 13 de la Convención Americana de Derecho Humanos reconoce ese derecho a la información, a la libertad de expresión y las restricciones de esta última: “*1. Toda persona tiene derecho a la libertad de pensamiento y de expresión. Este derecho comprende la libertad de buscar, recibir y difundir informaciones e ideas de toda índole, sin consideración de fronteras, ya sea oralmente, por escrito o en forma impresa o artística, o por cualquier otro procedimiento de su elección. 2. El ejercicio del derecho previsto en el inciso precedente no puede estar sujeto a previa censura sino a responsabilidades ulteriores, las que deben estar expresamente fijadas por la ley y ser necesarias para asegurar: a) El respeto a los derechos o a la reputación de los demás, o b) La protección de la seguridad nacional, el orden público o la salud o la moral públicas*”.

Estos tratados internacionales han establecido claramente el reconocimiento del derecho a la información, así como el de la libertad de expresión oral, escrita, impresa, artística o cualquier otro medio (a través de redes), pero también podemos establecer que la debida regulación de la

IA para controlar la integridad de la información no debe ser considerada como una violación a la libertad de expresión.

Al respecto la Corte Interamericana de Derechos Humanos en la Opinión Consultiva OC-5/85 en su párrafo 69 ha establecido: “*Considera la Corte, sin embargo, que el mismo concepto de orden público reclama que, dentro de una sociedad democrática, se garanticen las mayores posibilidades de circulación de noticias, ideas y opiniones, así como el más amplio acceso a la información por parte de la sociedad en su conjunto. También interesa al orden público democrático, tal como está concebido por la Convención Americana, que se respete escrupulosamente el derecho de cada ser humano de expresarse libremente y el de la sociedad en su conjunto de recibir información*”.

Para mantener el orden público en una sociedad democrático es necesario garantizar la libre circulación de noticias, ideas y opiniones, respetando el derecho humano a expresarse, pero a su vez, el amplio acceso de la sociedad a recibir información, misma que no debe dañar los derechos y la reputación de las personas.

Como se pudo observar el derecho humano a la libertad de opinión y de expresión no debe ser objeto de abuso ni de distorsión para perjudicar a otras personas o comunidades. Lamentablemente, el contenido generado por inteligencia artificial se ha utilizado cada vez más para difundir discursos de odio, xenofobia y retórica discriminatoria dirigidos contra poblaciones vulnerables, como los refugiados y las minorías étnicas (Naciones Unidas, 2025).

Un estudio de 2019 realizado por *Deeprtrace*, una empresa de ciberseguridad con sede en Ámsterdam, reveló que el 96 por ciento de todos los videos *deepfake* no eran consensuales y eran pornográficos. Además, el contenido generado por IA se puede utilizar en la trata de personas,

dirigido especialmente a mujeres y niños para atraerlos a la explotación (Naciones Unidas, 2025).

La Naciones Unidas establece principios Universales para establecer la integridad de la información (ONU, 2025):

- Credibilidad social y resiliencia: La credibilidad se refiere al grado de confianza que las personas tienen en la veracidad y exactitud de la información a la que acceden, y la resiliencia se refiere a la capacidad de las sociedades a gestionar las alteraciones o manipulaciones del ecosistema de la información.
- Buenos incentivos: Los anunciantes y las empresas de tecnología pueden adoptar modelos de negocio que defiendan los derechos humanos y refuercen la integridad de la información.
- Empoderamiento público: Los cambios en las políticas de las empresas de tecnología y una mayor alfabetización mediática pueden dotar a los usuarios de un mayor control sobre su experiencia en línea y el uso de sus propios datos.
- Medios de comunicación independientes, libres y pluralistas: La integridad de la información solo es posible con unos medios de comunicación independientes, libres y pluralistas.
- Transparencia e investigación: Es fundamental mejorar la transparencia y el acceso a los datos para conocer mejor los entornos de información de todo el mundo y ofrecer soluciones basadas en pruebas para promover la integridad de la información.

Queda por concluir que, para mantener el orden público en una sociedad democrático es necesario garantizar la libre circulación de noticias, ideas y opiniones, respetando el derecho

humano a expresarse, pero a su vez, el amplio acceso de la sociedad a recibir información, misma que no debe dañar los derechos y la reputación de las personas.

3 ONU

La Organización de las Naciones Unidas en el año 2024 para afrontar los desafíos tecnológicos adopto El Pacto Digital Global. El pacto es un marco integral diseñado para fomentar la cooperación global en la gestión de la tecnología digital y la IA. Adoptado en septiembre de 2024 junto con el Pacto para el Futuro, proporciona una hoja de ruta para mejorar la colaboración digital global, aprovechar el enorme potencial de las tecnologías digitales y reducir las brechas digitales existentes (Naciones Unidas).

Negociado por 193 Estados miembros y basado en consultas globales, este marco compromete a los gobiernos a respetar el derecho internacional y los derechos humanos en línea. Enfatiza la necesidad de acciones concretas para garantizar la seguridad del espacio digital (Naciones Unidas).

El pacto por el Futuro establece una serie de acciones a realizar por parte de los Estados Miembros de las naciones unidas sobre temas de desarrollo sostenible, seguridad internacional, juventud y generaciones futuras, gobernanza global y sobre ciencia, tecnología y cooperación digital.

El Pacto en la acción 27 denominada: “Aprovecharemos las oportunidades que ofrecen las tecnologías nuevas y emergentes y abordaremos los riesgos que puede plantear su uso indebido” señala: “Reconocemos que la rápida evolución tecnológica plantea oportunidades y riesgos para nuestros esfuerzos colectivos por mantener la paz y la seguridad internacionales. Nuestro enfoque para hacer frente a estos riesgos se guiará por el derecho internacional (ONU 2024: 22).

Además, establece que: “Las tecnologías digitales y emergentes, como la inteligencia artificial, son importantes para facilitar el desarrollo sostenible y están cambiando nuestro mundo de manera radical. Las posibilidades que ofrecen a las personas y el planeta de progresar, tanto en la actualidad como en el futuro, son enormes. Tenemos la determinación de materializar ese potencial y gestionar los riesgos aumentando la cooperación internacional, interactuando con las instancias pertinentes y promoviendo un futuro digital inclusivo, responsable y sostenible. En este sentido, adjuntamos el Pacto Digital Mundial como anexo del presente Pacto” (ONU 2024: 23).

El Pacto Digital Global tiene por objeto establecer un marco mundial inclusivo, fundamental para la acción de múltiples partes interesadas necesaria para superar las diferencias en materia digital, de datos y de innovación. El pacto establece los principios, los objetivos y las medidas para avanzar hacia un futuro digital abierto, libre, seguro y centrado en el ser humano para todos, basado en los derechos humanos universales (ONU, 2024).

El pacto digital destaca que los Estados reconocen la necesidad de determinar y mitigar los riesgos y garantizar la supervisión humana de la tecnología para promover el desarrollo sostenible y el pleno disfrute de los derechos humanos. Y aspiran a conseguir un futuro digital inclusivo, abierto, sostenible, justo y seguro para todos. En este Pacto Digital Global se enuncian los objetivos, los principios, los compromisos y las acciones que nos propondremos para lograrlo, fuera del ámbito militar (ONU 2024: 42). Además, el pacto digital tiene establecidos los siguientes objetivos (ONU 2024: 43):

1. Eliminar todas las brechas digitales y acelerar los progresos en todos los Objetivos de Desarrollo Sostenible;
2. Ampliar la inclusión en la economía digital y sus beneficios para todos;

3. Fomentar un espacio digital inclusivo, abierto y seguro que respete, proteja y promueva los derechos humanos;
4. Promover enfoques de la gobernanza de datos que sean responsables, equitativos e interoperables;
5. Mejorar la gobernanza internacional de la inteligencia artificial en beneficio de la humanidad.

Dentro del objetivo 3 que es fomentar un espacio digital inclusivo, abierto y seguro que respete, proteja y promueva los derechos humanos, en el ámbito de los derechos humanos los estados miembros se comprometen a (ONU 2024: 48):

- a) Velar por que, al elaborar y aplicar legislación nacional pertinente para las tecnologías digitales, se cumplan las obligaciones en virtud del derecho internacional, incluido el derecho internacional de los derechos humanos.
- b) Establecer salvaguardias adecuadas para prevenir y abordar cualquier efecto adverso en los derechos humanos derivado del uso de las tecnologías digitales y emergentes y proteger a las personas de las violaciones y los abusos de sus derechos humanos en el espacio digital, incluso ejerciendo la diligencia debida en materia de derechos humanos y creando mecanismos eficaces de supervisión y recurso.

Y exhorta a (ONU 2024: 48):

- a) Las empresas y los desarrolladores de tecnología digital a que respeten los derechos humanos y los principios internacionales, incluso ejerciendo la diligencia debida en materia de derechos humanos y evaluando el impacto durante todo el ciclo de vida de la tecnología.

- b) Las empresas y los desarrolladores de tecnología digital y las plataformas de medios sociales a que respeten los derechos humanos en el entorno digital, rindan cuentas por los abusos y tomen medidas para mitigarlos y prevenirlos, y a que den acceso a vías de recurso eficaces conforme a los Principios Rectores sobre las Empresas y los Derechos Humanos de las Naciones Unidas y otros marcos pertinentes.

En tema de confianza y seguridad digital el pacto busca contrarrestar y abordar urgentemente todas las formas de violencia, incluida la violencia sexual y de género, que se producen o amplifican por el uso de la tecnología, todas las formas de discurso de odio y discriminación, las informaciones erróneas y la desinformación, el ciberacoso y la explotación y los abusos sexuales de menores. Estableceremos y mantendremos estrictas medidas de mitigación de riesgos y vías de recurso que también protejan la privacidad y la libertad de expresión (ONU 2024: 50).

Por lo que se refiere a la integridad de la información el pacto señala lo siguiente (ONU 2024: 51, 52):

- a) El acceso a una información y unos conocimientos que sean pertinentes, fiables y exactos es esencial para lograr un espacio digital inclusivo, abierto y seguro. Reconocemos que las tecnologías digitales y emergentes pueden propiciar manipulaciones e interferencias en la información que son nocivas para las sociedades y las personas y perjudican el disfrute de los derechos humanos y las libertades fundamentales, así como la consecución de los Objetivos de Desarrollo Sostenible.
- b) Colaboraremos para promover la integridad de la información, la tolerancia y el respeto en el espacio digital, así como para proteger la integridad de los procesos democráticos. Reforzaremos la cooperación internacional para hacer frente al problema de las informaciones erróneas, la desinformación y el discurso de odio en el entorno digital y

mitigar los riesgos que plantea la manipulación de la información, de manera compatible con el derecho internacional.

- c) A las empresas y los desarrolladores de tecnología digital a que sigan desarrollando soluciones y comunicando públicamente las medidas que hayan tomado para contrarrestar posibles daños, como el discurso de odio y la discriminación, derivados de los contenidos de inteligencia artificial. Algunas de esas medidas son incorporar salvaguardias en los procesos de entrenamiento de los modelos de inteligencia artificial, identificar el material generado por la inteligencia artificial y certificar la autenticidad de los contenidos y su origen, así como etiquetas, marcas de agua y otras técnicas.

4 UNESCO

En noviembre de 2021 la UNESCO estableció una serie de principios y directrices con el fin de asegurar el uso responsable, ético y en beneficio de la sociedad: Recomendaciones sobre la ética de la inteligencia artificial, mismas que se centran en la protección de los derechos humanos, la transparencia, la seguridad y el bienestar general.

En el documento se recomienda que los Estados Miembros apliquen, de manera voluntaria, las disposiciones de la presente Recomendación mediante la adopción de las medidas adecuadas, en particular las medidas legislativas o de otra índole que puedan ser necesarias, de acuerdo con la práctica constitucional y las estructuras de gobierno de cada Estado, con el fin de dar efecto en sus respectivas jurisdicciones a los principios y normas enunciados en la Recomendación, de conformidad con el derecho internacional, incluido el derecho internacional de los derechos humanos (UNESCO, 2021: 7).

Unos de los fines de la presente recomendación son:

- Proporcionar un marco universal de valores, principios y acciones para orientar a los Estados en la formulación de sus leyes, políticas u otros instrumentos relativos a la IA, de conformidad con el derecho internacional;
- Aportar un instrumento normativo aceptado mundialmente que no solo se centre en la articulación de valores y principios, sino también en su aplicación práctica, mediante recomendaciones de política concretas (UNESCO, 2021: 14).

El respeto, la protección y la promoción de la dignidad humana y de los derechos establecidos por el derecho internacional, en particular el derecho internacional de los derechos humanos, son esenciales a lo largo del ciclo de vida de los sistemas de IA. La dignidad humana tiene que ver con el reconocimiento del valor intrínseco e igual de cada ser humano, con independencia de su raza, color, ascendencia, género, edad, idioma, religión, opiniones políticas, origen nacional, étnico o social, condición económica o social de nacimiento, discapacidad o cualquier otro motivo (UNESCO, 2021: 18).

Ningún ser humano ni comunidad humana debería sufrir daños o sometimiento, ya sean de carácter físico, económico, social, político, cultural o mental, durante ninguna etapa del ciclo de vida de los sistemas de IA. A lo largo de su ciclo de vida, los sistemas de IA deberían mejorar la calidad de vida de los seres humanos, dejando a las personas o los grupos la tarea de definir el concepto de “calidad de vida”, siempre que como resultado de esa definición no se produzca ninguna violación o abuso de los derechos humanos y las libertades fundamentales ni de la dignidad de las personas (UNESCO, 2021: 18).

En el ámbito de acción los Estados Miembros deberían utilizar los sistemas de IA para mejorar el acceso a la información y el conocimiento. Para ello, pueden, por ejemplo, apoyar a los investigadores, las universidades, los periodistas, el público en general y los desarrolladores, a

fin de mejorar la libertad de expresión, las libertades académicas y científicas y el acceso a la información, así como aumentar la divulgación proactiva de los datos y la información oficiales (UNESCO, 2021: 36).

Además de, garantizar que los actores de la IA respeten y promuevan la libertad de expresión y el acceso a la información en lo que respecta a la generación, moderación y conservación automáticas de contenidos. Marcos adecuados, incluso reglamentarios, deberían propiciar la transparencia de los operadores de comunicación e información en línea, velar por que los usuarios tengan acceso a puntos de vista diversos y prever procesos de notificación rápida a los usuarios sobre los motivos de la eliminación u otro tratamiento de los contenidos, así como mecanismos de recurso que permitan a los usuarios solicitar reparación (UNESCO, 2021: 36).

Invertir en competencias digitales y de alfabetización mediática e informacional y promoverlas, a fin de reforzar el pensamiento crítico y las competencias necesarias para comprender el uso y las implicaciones de los sistemas de IA, con miras a atenuar y contrarrestar la desinformación, la información errónea y el discurso de odio (UNESCO, 2021: 36).

Crear entornos propicios para que los medios de comunicación tengan los derechos y recursos necesarios para informar eficazmente sobre las ventajas y los inconvenientes de los sistemas de IA, y también alentar a los medios de comunicación a que hagan un uso ético de estos sistemas en su trabajo (UNESCO, 2021: 36).

5 OCDE

La OCDE (Organización para la Cooperación y el Desarrollo Económicos) estableció en mayo de 2019 la Recomendación del Consejo sobre Inteligencia Artificial los primeros estándares sobre IA y posteriormente en el año 2024 fue actualizada, se trata de directrices que intenta

promover la confiabilidad de la Inteligencia Artificial, teniendo como objetivos: fomentar la innovación y la confianza; Alinearse con los valores y derechos humanos; guiar a los actores principales que establecen políticas; y promover la cooperación internacional.

El texto establece que los actores de la IA deben respetar el Estado de derecho, los derechos humanos y los valores democráticos y centrados en el ser humano durante todo el ciclo de vida del sistema de IA. Estos incluyen la no discriminación y la igualdad, la libertad, la dignidad, la autonomía individual, la privacidad y la protección de datos, la diversidad, la equidad, la justicia social y los derechos laborales internacionalmente reconocidos. Esto también incluye abordar la desinformación y la información errónea que la IA difunde, respetando al mismo tiempo la libertad de expresión y otros derechos y libertades protegidos por el derecho internacional aplicable (OCDE, 2024).

Para tal fin, los actores de la IA deben implementar mecanismos y salvaguardas, como capacidad de agencia y supervisión humana, incluso para abordar los riesgos que surgen de usos fuera del propósito previsto, mal uso intencional o mal uso no intencional de una manera apropiada al contexto y consistente con el estado del arte (OCDE, 2024).

Los sistemas de IA deben ser robustos, seguros y protegidos durante todo su ciclo de vida para que, en condiciones de uso normal, uso previsible o mal uso, u otras condiciones adversas, funcionen adecuadamente y no planteen riesgos irrazonables para la seguridad (OCDE, 2024).

Se deben establecer mecanismos, según corresponda, para garantizar que, si los sistemas de IA corren el riesgo de causar daños indebidos o exhiben un comportamiento no deseado, puedan ser anulados, reparados y/o desmantelados de manera segura según sea necesario (OCDE, 2024).

También deberían establecerse, cuando sea técnicamente posible, mecanismos para reforzar la integridad de la información garantizando al mismo tiempo el respeto de la libertad de expresión (OCDE, 2024).

Las recomendaciones de la OCDE establecen siguientes Políticas nacionales y cooperación internacional para una IA confiable (OCDE, 2024).:

- Invertir en investigación y desarrollo de IA.
- Fomento de un ecosistema inclusivo que facilite la IA.
- Desarrollo de un entorno de gobernanza y políticas interoperable y propicio para la IA.
- Desarrollo de la capacidad humana y preparación para la transformación del mercado laboral.
- Cooperación internacional para una IA fiable.

6 UNIÓN EUROPEA

The La Unión Europea en 2024 estableció el primer ordenamiento jurídico global sobre Inteligencia Artificial denominado Ley de la IA de la Unión Europea. El objetivo del presente Reglamento es mejorar el funcionamiento del mercado interior y promover la adopción de la inteligencia artificial (IA) centrada en el ser humano y digna de confianza, garantizando al mismo tiempo un alto nivel de protección de la salud, la seguridad, los derechos fundamentales consagrados en la Carta, incluida la democracia, el Estado de Derecho y la protección del medio ambiente, contra los efectos nocivos de los sistemas de IA en la Unión y apoyando la innovación (Unión Europea, 2024: Art. 1).

El Reglamento establece:

- a) normas armonizadas para la comercialización, la puesta en servicio y el uso de los sistemas de IA en la Unión;
- b) prohibiciones de determinadas prácticas de IA;
- c) requisitos específicos para los sistemas de IA de alto riesgo y obligaciones para los operadores de dichos sistemas;
- d) normas de transparencia armonizadas para determinados sistemas de IA;
- e) normas armonizadas para la comercialización de modelos de IA de uso general;
- f) normas sobre supervisión, vigilancia, gobernanza y ejecución del mercado;
- g) medidas de apoyo a la innovación, con especial atención a las PYME, incluidas las empresas de nueva creación (Unión Europea, 2024).

El Reglamento se aplica a:

- a) los proveedores que comercialicen o pongan en servicio sistemas de IA o comercialicen modelos de IA de uso general en la Unión, con independencia de que dichos proveedores estén establecidos o ubicados en la Unión o en un tercer país;
- b) los implantadores de sistemas de IA que tengan su lugar de establecimiento o estén situados en la Unión;
- c) proveedores e implantadores de sistemas de IA que tengan su lugar de establecimiento o estén situados en un tercer país, cuando el producto generado por el sistema de IA se utilice en la Unión;
- d) importadores y distribuidores de sistemas de IA;
- e) fabricantes de productos que comercialicen o pongan en servicio un sistema de IA junto con su producto y bajo su propio nombre o marca;
- f) los representantes autorizados de los prestadores que no estén establecidos en la Unión;

g) personas afectadas que se encuentren en la Unión (Unión Europea, 2024).

El reglamento tiene prohibidas las siguientes prácticas de IA: La comercialización, la puesta en servicio o el uso de un sistema de IA que utilice técnicas subliminales que escapen a la conciencia de una persona o técnicas deliberadamente manipuladoras o engañosas, con el objetivo o el efecto de distorsionar materialmente el comportamiento de una persona o de un grupo de personas, mermando de forma apreciable su capacidad de tomar una decisión con conocimiento de causa, haciéndoles tomar una decisión que de otro modo no habrían tomado, de forma que causen o sea razonablemente probable que causen a esa persona, a otra persona o a un grupo de personas un daño significativo (Unión Europea, 2024).

Conclusiones

El derecho a la información se encuentra consagrada en diversos tratados internacionales, al ser un derecho humano le corresponde al Estado garantizar ese derecho, eso incluye los medios digitales y la información o contenido realizado a través de inteligencia artificial. Además, ante el uso de Inteligencias Artificiales que pretenden con técnicas deliberadamente manipuladoras o engañosas, con el objetivo o el efecto de distorsionar materialmente el comportamiento de una persona o de un grupo de personas, los Estados deben garantizar la integridad de la información, sin menoscabar el derecho a la libertad de expresión.

Es indispensable para toda sociedad democrática mantener el orden público, para ello es necesario garantizar la libre circulación de noticias, ideas y opiniones, respetando el derecho humano a expresarse, pero a su vez, el amplio acceso de la sociedad a recibir información, misma que no debe dañar los derechos y la reputación de las personas.

En ese sentido, los Estado para cumplir con su labor como garante de estos derechos es indispensable que adopte una normativa que regule y sancione aquellos que hagan mal uso de la IA en perjuicio del orden público y de la sociedad, y para ello, es imperativo que adopte las recomendaciones, parámetros y directrices que los organismos como: a) la ONU con su pacto por el futuro y pacto digital global, b) la UNESCO con su reglamento sobre el uso ético de la IA; c) La OCDE con su reglamento del Consejo sobre Inteligencia Artificial; y d) La Unión Europea con su Ley sobre la IA.

Estos organismos han establecido importantes estándares a nivel internacional para que los Estados puedan implementar en aras de regular a las inteligencias artificiales y a su vez garantizar la integridad de la información, mismos que se mencionaron en el cuerpo del trabajo, pero de los cuales quiero destacar los siguientes:

- Los Estados deben reconocer la necesidad de determinar y mitigar los riesgos y garantizar la supervisión humana de la tecnología para promover el desarrollo sostenible y el pleno disfrute de los derechos humanos.
- Los Estados deben velar porque, al elaborar y aplicar legislación nacional pertinente para las tecnologías digitales, se cumplan las obligaciones en virtud del derecho internacional, incluido el derecho internacional de los derechos humanos.
- Los Estados deben establecer salvaguardias adecuadas para prevenir y abordar cualquier efecto adverso en los derechos humanos derivado del uso de las tecnologías digitales y emergentes y proteger a las personas de las violaciones y los abusos de sus derechos humanos en el espacio digital, incluso ejerciendo la diligencia debida en materia de derechos humanos y creando mecanismos eficaces de supervisión y recurso.

- Las empresas y los desarrolladores de tecnología digital deben respetar los derechos humanos y los principios internacionales, incluso ejerciendo la diligencia debida en materia de derechos humanos y evaluando el impacto durante todo el ciclo de vida de la tecnología.
- Las empresas y los desarrolladores de tecnología digital y las plataformas de medios sociales deben respetar los derechos humanos en el entorno digital, que rindan cuentas por los abusos y tomen medidas para mitigarlos y prevenirlos.
- Los Estados deben contrarrestar y abordar urgentemente todas las formas de violencia, incluida la violencia sexual y de género, que se producen o amplifican por el uso de la tecnología, todas las formas de discurso de odio y discriminación, las informaciones erróneas y la desinformación, el ciberacoso y la explotación y los abusos sexuales de menores.
- Los Estados deben garantizar una información y unos conocimientos que sean pertinentes, fiables y exactos es esencial para lograr un espacio digital inclusivo, abierto y seguro.
- Los Estados deben promover la integridad de la información, la tolerancia y el respeto en el espacio digital, así como para proteger la integridad de los procesos democráticos.
- Los Estados deben mitigar los riesgos que plantea la manipulación de la información, de manera compatible con el derecho internacional.
- Las empresas y los desarrolladores de tecnología digital a que sigan desarrollando soluciones y comunicando públicamente las medidas que hayan tomado para contrarrestar posibles daños, como el discurso de odio y la discriminación, derivados de los contenidos de inteligencia artificial. Algunas de esas medidas son incorporar

salvaguardias en los procesos de entrenamiento de los modelos de inteligencia artificial, identificar el material generado por la inteligencia artificial y certificar la autenticidad de los contenidos y su origen, así como etiquetas, marcas de agua y otras técnicas.

- Los organismos internacionales deben orientar a los Estados en la formulación de sus leyes, políticas u otros instrumentos relativos a la IA, de conformidad con el derecho internacional.
- Los sistemas de IA deben mejorar la calidad de vida de los seres humanos.
- Los Estados deben lograr que los sistemas de IA para mejorar el acceso a la información y el conocimiento.
- Los Estados deben garantizar que los actores de la IA respeten y promuevan la libertad de expresión y el acceso a la información en lo que respecta a la generación, moderación y conservación automáticas de contenidos.
- Los Estados deben invertir en competencias digitales y de alfabetización mediática e informacional y promoverlas, con miras a atenuar y contrarrestar la desinformación, la información errónea y el discurso de odio.
- Los Estados deben crear entornos propicios para que los medios de comunicación tengan los derechos y recursos necesarios para informar eficazmente y también alentarlos a que hagan un uso ético de estos sistemas.
- Los Estados deben establecer mecanismos para reforzar la integridad de la información garantizando al mismo tiempo el respeto de la libertad de expresión
- Los Estados deben prohibir la comercialización, la puesta en servicio o el uso de un sistema de IA que utilice técnicas subliminales que escapen a la conciencia de una persona o técnicas deliberadamente manipuladoras o engañosas, con el objetivo o el

efecto de distorsionar materialmente el comportamiento de una persona o de un grupo de personas.

Finalmente, el soft law aporta a los Estados la experiencia de los especialistas a nivel internacional sobre el tema de Inteligencia Artificial, los cual a pesar de no ser vinculantes son documentos jurídicamente y técnicamente relevantes para asegurar el orden público, la protección de los derechos de sus nacionales y brindar certeza jurídica en el uso de la IA.

References

ASAMBLEA GENERAL DE LAS NACIONES UNIDAS. (2024). Pacto por el Futuro.

<https://docs.un.org/es/A/RES/79/1>

CIDH. (1985). Opinión Consultiva OC-5/85.

https://www.corteidh.or.cr/docs/opiniones/seriea_05_esp.pdf

CONVENCIÓN AMERICANA DE DERECHOS HUMANOS. (1978).

https://www.oas.org/dil/esp/1969_Convenci%C3%B3n_Americana_sobre_Derechos_Humanos.pdf

DECLARACIÓN UNIVERSAL DE LOS DERECHOS HUMANOS. (1948).

<https://www.un.org/es/about-us/universal-declaration-of-human-rights>

NACIONES UNIDAS. (2023). Our Common Agenda Policy Brief 8 Information Integrity on Digital

Platforms. <https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-information-integrity-en.pdf>

NACIONES UNIDAS. <https://www.un.org/en/global-issues/artificial-intelligence>

NACIONES UNIDAS. (2025). <https://www.un.org/es/cr%C3%B3nica-onu/la-protecci%C3%B3n-de-los-derechos-humanos-y-la-integridad-de-la-informaci%C3%B3n-en-la-era-de-la>

OCDE. (2024). <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

ONU. (2025). <https://www.un.org/es/information-integrity/global-principles>

ONU (2024). <https://www.un.org/es/summit-of-the-future/global-digital-compact>

PACTO INTERNACIONAL DE DERECHOS CIVILES Y POLÍTICOS. (1976).

<https://www.ohchr.org/es/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights>

UNESCO, Recomendaciones sobre la ética de la inteligencia artificial, 2021, p7

UNION EUROPEA (2024). Ley de la IA de la Unión Europea. <https://artificialintelligenceact.eu/es/ai-act-explorer/>