Rolando Medina Peña
Yisel Muñoz Alfonso
Yulier Campos Pérez
Daniela Lázara Hernández Domínguez

INTELIGENCIA ARTIFICIAL,

ética y derechos. Desafíos

Rolando Medina Peña
Yisel Muñoz Alfonso
Yulier Campos Pérez
Daniela Lázara Hernández Domínguez

INTELIGENCIA ARTIFICIAL,

ética y derechos. Desafíos

Diseño de carátula y edición: D.I. Yunisley Bruno Díaz

Dirección editorial: PhD. Jorge Luis León González

ISBN: 978-9942-560-11-7

Editorial EXCED
Dr. Kennedy Nueva. 2do Callejón 11
A. Manzana 42, Número 26.
Guayaquil, Ecuador.
E-mail: editorial@excedinter.com

Rolando Medina Peña
Yisel Muñoz Alfonso
Yulier Campos Pérez
Daniela Lázara Hernández Domínguez

INTELIGENCIA ARTIFICIAL,

ética y derechos. Desafíos

COMITÉ EDITORIAL

ÍNDICE

ÍNDICE

CAPÍTULO 3.

Marco jurídico y regulatorio de la inteligencia artificial. Desafíos legales

PRÓLOGO

Vivimos en una era de transformación sin precedentes, impulsada por una fuerza que redefine los límites de lo posible: la Inteligencia Artificial (IA). Esta tecnología, que alguna vez fue dominio exclusivo de la ciencia ficción y la academia, ha irrumpido con vigor en el tejido mismo de nuestra sociedad, reconfigurando industrias, economías y, de manera más profunda, las interacciones humanas y los marcos normativos. Nos encontramos en una encrucijada histórica donde el vertiginoso desarrollo tecnológico parece superar, con frecuencia, la capacidad humana para comprender sus implicaciones y, sobre todo, para regular sus efectos.

Este libro nace precisamente de la necesidad de tender un puente sólido entre estos tres ámbitos indisociables. No es un texto meramente descriptivo, sino una obra que busca diseccionar, analizar y proponer respuestas a las complejas interrogantes que la inteligencia artificial plantea a nuestra convivencia. A través de un análisis profundo que abarca desde sus fundamentos teóricos hasta los complejos retos éticos y jurídicos que plantea su desarrollo y aplicación, esta obra se erige como una contribución fundamental para investigadores, profesionales y legisladores interesados en comprender y regular esta tecnología de manera responsable y efectiva.

En el primer capítulo, se establece el marco conceptual y tecnológico necesario para entender la inteligencia artificial desde una perspectiva jurídica. Se revisa su evolución histórica y desarrollo tecnológico, se define su naturaleza y características principales, y se examinan los principios aplicables a su correcto uso. Esta sección cobra relevancia al mostrar cómo la inteligencia artificial no solo es un desafío técnico, sino un fenómeno de gran trascendencia jurídica que exige criterios claros y precisos para su regulación.

El segundo capítulo se adentra en los fundamentos éticos vinculados a la inteligencia artificial, una dimensión indispensable para orientar su

implementación hacia el respeto a los valores humanos. Se analizan los valores éticos centrales, los límites necesarios para su uso, y los desafíos específicos que surgen en áreas sensibles como las aplicaciones biomédicas, el reconocimiento facial y las interfaces cerebro-computadora. Además, se reflexiona sobre la ética en la administración pública y las iniciativas internacionales y regionales que buscan establecer un régimen ético coherente y efectivo en la gobernanza de la inteligencia artificial.

Finalmente, el tercer capítulo aborda el ámbito jurídico y regulatorio, presentando un minucioso análisis del marco normativo internacional y regional vigente, así como los principales desafíos legales derivados del uso de la inteligencia artificial. Se examinan temas cruciales como los derechos digitales, la propiedad intelectual, la responsabilidad ante daños, los sesgos algorítmicos, la privacidad, los ciberdelitos, y la necesidad de un esquema jurídico dinámico capaz de adaptarse a la rápida evolución tecnológica. También se incluye un estudio de las regulaciones y políticas sobre inteligencia artificial en varios contextos regionales,

Este libro se sitúa en la frontera del diálogo entre tecnología, ética y derecho, proponiendo una reflexión imprescindible para pensar cómo las sociedades contemporáneas pueden integrar la inteligencia artificial sin perder de vista los principios fundamentales de justicia, equidad y protección de los derechos humanos. Su enfoque interdisciplinario y plural enriquece la discusión y aporta herramientas conceptuales y prácticas para enfrentar los desafíos que plantea la inteligencia artificial en el mundo actual y futuro.

Este libro se sitúa en un contexto vital para el desarrollo tecnológico y social: la Recomendación de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura sobre Inteligencia Artificial, aprobada el 18 de febrero de 2021, que establece principios orientadores para el uso ético, inclusivo y responsable de la inteligencia artificial. A través de sus páginas, se profundiza en la importancia de estos principios como fundamento para diseñar y aplicar tecnologías que respeten los derechos humanos, promuevan la equidad y contribuyan al bienestar colectivo.

Su papel es clave para fomentar un debate informado y ofrecer herramientas que permitan a gobiernos, instituciones y ciudadanos entender y aplicar estos estándares internacionales, asegurando que la inteligencia artificial se utilice como una fuerza positiva para el desarrollo sostenible y la cohesión social. Este libro, por tanto, no solo aporta conocimiento técnico y conceptual, sino que también se erige como un llamado a la responsabilidad compartida frente a los desafíos y oportunidades que la inteligencia artificial presenta en nuestro tiempo.

Por último, el manual se presenta como una herramienta fundamental para comprender y aplicar los cuatro principios rectores que aseguran la alineación de los sistemas de inteligencia artificial con los valores y derechos humanos desde una perspectiva jurídica. La robustez se traduce en la capacidad de la inteligencia artificial para operar con seguridad y resiliencia ante escenarios adversos o ataques, garantizando la protección de los derechos fundamentales. La Interpretabilidad es clave para la transparencia, permitiendo que las decisiones algorítmicas sean comprensibles y auditables por operadores y reguladores. La controlabilidad asegura que el control y la supervisión permanezcan en manos humanas, permitiendo la intervención legal o correctiva cuando sea necesario. Finalmente, la ética reafirma el compromiso del sistema con estándares morales universales y la defensa de los valores humanos, que conforman el núcleo del Estado de Derecho.

Este enfoque pone al alcance de juristas, legisladores y operadores del derecho una guía para evaluar, regular y asegurar que la inteligencia artificial se desarrolle y utilice de manera responsable y conforme a los principios legales y éticos vigentes.

Dr.C María Matilde García Lorenzo

Profesora titular e investigadora del Centro de estudios de Informática de la Universidad Central “Marta Abreu” de las Villas”

INTRODUCCIÓN

Desde finales del pasado siglo la humanidad ha venido experimentado un desarrollo vertiginoso de las tecnologías de la informática y las telecomunicaciones. Como es de suponer ello ha impactado todas las actividades de los seres humanos, tal es el caso del Derecho. En consecuencia, cada día es más común encontrarse con normas jurídicas que se encarguen de regular algunos de los aspectos relacionados con esta área, ya sea para la estimulación y protección de los creadores y desarrolladores de sistemas de Inteligencia Artificial, así como para delimitar los efectos y cambios que la actividad genera en el resto del ordenamiento jurídico y el ámbito privado de los ciudadanos.

A nivel internacional varios son los ejemplos de disposiciones que se encargan de regular alguno de los aspectos relacionados con el uso de la tecnología en este campo, en especial lo relacionado con el uso ético de la inteligencia artificial y su debida articulación con los derechos de las personas. En este orden, destaca la Declaración de Principios de Ginebra de la Cumbre Mundial sobre la Sociedad de la Información que establece que “el uso de las TIC y la creación de contenido deben respetar los derechos humanos y las libertades fundamentales de otros, incluida la privacidad personal y el derecho a la libertad de pensamiento y de conciencia” (Organizaci0on de las Naciones Unidas, 2003). De igual manera, destacan los aportes realizados por la Comisión Mundial de Ética del Conocimiento Científico y la Tecnología de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura.

En el contexto internacional también se reconoce la importancia y la contribución de la informatización, y en particular las herramientas de inteligencia artificial, para el desarrollo sostenible como objetivo primordial

de la comunidad internacional tanto desde la dimensión ambiental, económica como la social, especialmente desde los derechos humanos y el derecho al desarrollo receptado en la "Agenda 2030 para el Desarrollo Sostenible de las Naciones Unidas" (Res. 70/1 de la Asamblea General de las Naciones Unidas). No obstante, aunque las tecnologías de inteligencia artificial pueden respaldar avances en el logro de los Objetivos de Desarrollo Sostenible (ODS), también pueden tener consecuencias negativas que incrementen las desigualdades y afecten negativamente a las personas, las sociedades, las economías y el entorno humano, de ahí la necesidad de sistematizar sus implicaciones ético- legales.

Aunque la doctrina jurídica internacional es prolífera en el análisis de esta temática aún quedan importantes puntos en lo que todavía no existe un criterio único que ofrezca una solución a las problemáticas planteadas. Tal es el caso de los riesgos en el desarrollo de las herramientas de inteligencia artificial y la forma de minimizarlos, a ello se une la previsión de mecanismos legales que permitan por una parte la estimulación de estas creaciones y por la otra la correcta armonía con el resto del orden legal y los derechos y facultades individuales de los ciudadanos.

En el caso de Latinoamérica, la realidad es aún incierta puesto que, aunque los Estados han propiciado la introducción y desarrollo de herramientas informáticas esto no ocurre con la misma rapidez que el mundo desarrollado.

A lo anterior se une que desde el Derecho no existen antecedentes consolidados que se encarguen de regular de forma directa y expresa las cuestiones referentes al tema. No obstante, en el presente se han aprobado diferentes normas que conforman un marco regulatorio internacional y regional insipiente vinculado a la informatización de la sociedad. Estas normas jurídicas aunque suponen un avance en la regulación no incluyen todos los aspectos necesarios que garanticen una articulación orgánica con el resto del ordenamiento legal y las facultades de los individuos.

Por tanto, el presente texto contribuiría al desarrollo de la doctrina jurídica vinculada al tema y a la postre al perfeccionamiento del marco regulatorio. La investigación que se publica contiene en

el primer capítulo valoraciones sobre los presupuestos teóricos de la inteligencia artificial y sus implicaciones jurídicas, haciendo un recorrido sobre los antecedentes históricos, la recepción de esta categoría tecnológica en el campo de las ciencias jurídicas, con la definición de conceptos y del sistema categorial relativo al tema desarrolla un recorrido por los principios y aquellos aspectos novedosos en que impacta como son la robótica, los vehículos autónomos, la medicina y el sistema legal y judicial.

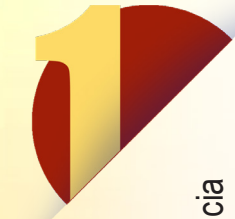
El segundo capítulo se encamina a establecer las implicaciones éticas de las IA, valorando las mutuas interrelaciones entre la ética y la inteligencia artificial, los límites para el campo de la tecnología, se profundiza en el régimen ético internacional y se realiza un estudio e derecho comparado que posibilita contrastar los avances en la regulación ética de los sistemas de IA.

El tercer capítulo contiene estudios sobre el marco jurídico y regulatorio de la inteligencia artificial y los desafíos legales que plantea en el contexto actual, con temas como los derechos digitales, la protección de en los regímenes de propiedad intelectual, los sistemas de responsabilidad e IA, el derecho penal y los ciberdelitos, la protección de datos, los asistentes legales y un análisis sobre su tratamiento en Ecuador.

En resumen, se alcanza a ofrecer una panorámica general sobre la Inteligencia artificial y el derecho con un enfoque de las principales tendencias y novedades sobre el tema, posibilitando un acercamiento al mismo a los estudiantes y profesionales del derecho y otras ramas interesados en el tema.

El texto contempla los resultados de investigaciones desarrolladas en la Universidad Metropolitana, desde la Facultad de derecho, planificadas en el proyecto de investigación: “Fundamentos epistemológicos del neoconstitucionalismo latinoamericano. Aciertos y desaciertos en su regulación jurídica y aplicación práctica en Ecuador” (Sede Machala) (Medina, et al., 2021) y asimismo comprende los resultados alcanzados en el proyecto Sectorial del MES en Cuba “Bases jurídicas para el uso ético y legal de herramientas informáticas y de inteligencia artificial en instituciones de la Educación Superior cubana”, con ellos

se alcanzan a sistematizar aspectos sustanciales del tema en Ecuador y Cuba desde la perspectiva regulatoria, constituyendo un referente investigativo desde la perspectiva del derecho comparado.



CAPÍTULO 1

Presupuestos teóricos de la inteligencia artificial. Implicaciones jurídicas

1.1. Evolución histórica. Desarrollo tecnológico e inteligencia artificial

Desde mediados del siglo XVIII, la humanidad ha experimentado profundas transformaciones tecnológicas que han marcado diferentes revoluciones industriales: la mecanización impulsada por la máquina de vapor y el ferrocarril; la electrificación que revolucionó la producción y las comunicaciones; y la era digital, caracterizada por la computación, la informática personal y la expansión del Internet. Más recientemente, con la llegada del siglo XXI (hace solo apenas dos décadas), el auge de los dispositivos móviles y la inteligencia artificial han impulsado una nueva revolución industrial (Schwab, 2017).

Desde inicios de la década de 1950 ya se hablaba de inteligencia artificial en las investigaciones de Alan Turing, influenciado por el trabajo previo de Marian Rejeswky, Jerzy Eozyeahi y Henryk Zygalski (Arana, 2023). Para

Turing si los humanos toman la información disponible y razonan con ella para tomar decisiones, una máquina bien programada con una tecnología suficiente podía hacer exactamente lo mismo. Ese pensamiento fue el marco teórico de un artículo que publicó en 1950, *Computing machinery and Intelligence* (Turing, 1950), en el que por primera vez se estaba planteado la idea de inteligencia en las máquinas creada por los humanos y programada en los dispositivos. En el trabajo se plantea un experimento conocido como Test de Turing, concretamente un "examen de la capacidad de una máquina para exhibir un comportamiento inteligente similar a la de un ser humano o distinguible de este" (Domingues Villarroel, 2021).

Cuando las ideas de Turing en este campo comenzaban a difuminarse, en 1956, McCarthy et al. (1955), también pioneros en el abordaje del tema llevaron a cabo el Proyecto de Investigación de verano de Dartmouth sobre Inteligencia Artificial, que inauguró formalmente el campo de estudio de la inteligencia artificial.

Las dos décadas posteriores, no obstante, no exhibieron un dinamismo significativo en ese ámbito aun cuando se obtuvieron algunos resultados. En 1966 llegó Eliza, un programa de procesamiento del lenguaje natural diseñado para simular una conversación terapéutica. Su algoritmo se basaba en determinar las palabras claves de las frases escritas por el usuario y responder con frases pre-programadas. Podía imitar lo suficientemente bien el lenguaje como para mantener una conversación y emular el comportamiento de un terapeuta, pero cuando los mensajes eran complejos, la conversación se tornaba incoherente (Haenlin y Kaplan, 2019).

Hacia los años '80, comienza a surgir la industria de los Sistemas Expertos (Waltz, 1997). Realizándose importantes inversiones en varios países de Europa, Asia y América, con el fin de lograr generar un sistema capaz de reproducir la actividad de un experto humano en tópicos específicos. Como en otras áreas de

la inteligencia artificial, los primeros resultados fueron atractivos y eso generó una expectativa desmesurada. Pero la comunidad halló severas dificultades en la manipulación de la gran cantidad de información necesaria para poder llevar a cabo una actividad realmente experta en el sentido humano.

Así, en los años ´90, con el crecimiento exponencial del Internet y el surgimiento de la *world wide web*, la empresa tecnológica multinacional estadounidense con sede en Nueva York, dedicada a la fabricación y comercialización de *hardware y software*, así como otros servicios relacionados con el sector de las tecnologías, International Business Machines Corporation (2024) se propuso el desarrollo de un ordenador cuyo propósito fue convertirse en un jugador de ajedrez al que ningún humano pudiera ganar, se trata pues de Deep Blue. Luego, el 3 de mayo de 1997, el mundo fue testigo del enfrentamiento en una partida de ajedrez entre la invencible máquina y el campeón del mundo Garri Kaspárov. Este constituyó un hecho mediático sin precedentes, llamado por muchos como la partida del siglo: el ser humano contra la máquina (Cárdenas Gallagher, 2020).

Con la entrada del siglo XXI se evidenció un aumento en el poder de procesamiento y en la disponibilidad del número de datos. Ya Internet había conectado al mundo y las máquinas tenían una cantidad de información tal que se abrió la puerta a una función revolucionaria, llamada: *machine learning*. Esta buscó conseguir que las máquinas aprendieran, acercándose a una verdadera IA. De esta forma se empiezan a cambiar los fundamentos de la inteligencia artificial, puesto que con los algoritmos convencionales se programaban a las máquinas para que realizaran una nueva acción, mientras que con el *machine learning* se programaban para que aprendieran a hacer una acción (Top de Impacto, 2023).

El *machine learning* o aprendizaje automático es “un conjunto de algoritmos que proporcionan a los sistemas la capacidad de aprender y mejorar automáticamente a partir de la experiencia

y acumulación de datos sin necesidad de estar programados previamente o de tener reglas previamente codificadas" (Kouvchinova y Voronine, 2024); así al introducir datos al software, este los procesa a partir de la identificación de patrones que le permiten construir modelos, hacer predicciones y tomar decisiones en función de esos datos, tal como lo haría una persona.

En su generalidad, cuenta con tres principales clasificaciones: el aprendizaje supervisado, el aprendizaje no supervisado y el aprendizaje de refuerzo. Las dos primeras se distinguen entre sí por la estructuración o no, mediando el factor humano, de los datos que se le facilitan: en el supervisado los datos son clasificados por una persona humana, mientras que en el no supervisado no existe intervención humana y los algoritmos aprenden de elementos no etiquetados. A su vez el aprendizaje de refuerzo utiliza el método de ensayo y error, dejando a la inteligencia artificial tomar decisiones independientemente y recompensándola mediante retroalimentación en los casos en que sean correctas, de modo que el sistema pueda evaluar las consecuencias de sus acciones y elegir una estrategia a largo plazo para maximizar la señal de recompensa que recibe.

Siguiendo la dinámica anterior, se hicieron posibles las recomendaciones de productos y servicios en Internet, creándose programas capaces de predecir los gustos y necesidades al analizar los historiales de compra, comportamientos en la red y preferencias. Por su parte los sitios *web* comenzaron a hacer publicidad personalizada, Google pudo desarrollar su sistema de búsquedas y recomendaciones, los sistemas de traducción llegaron al público en cualquier idioma, al tiempo que encontró su aplicación en la medicina como apoyo a los médicos en pos de interpretar imágenes y elaborar diagnósticos (Concha Flores, 2024).

Ante ese escenario, en 2010, un equipo de investigadores de la Universidad de Princeton creó una base de datos en formato de colección de imágenes: el programa ImageNet. Sin embargo, los

algoritmos de *machine learning* no respondían, por lo que parecía que se había llegado al límite, pero el 30 de septiembre de 2012 surgió un nuevo programa desarrollado en la Universidad de Toronto nombrado AlexNet, que mostró una capacidad de acierto de cerca del 85%. Los científicos detrás de este programa dieron un paso más allá en el desarrollo de la inteligencia artificial, algo que fue bautizado como aprendizaje profundo o *deep learning*.

Este resulta ser una especialización del *machine learning* con sus mismos principios de funcionamiento, pero que ofrece mayores capacidades de procesamiento al poseer una arquitectura de múltiples capas de un algoritmo denominado red neuronal, donde cada una de ellas proporciona una interpretación diferente de los datos en un intento de copiar la función de las redes neuronales del cerebro humano. De modo que, en un sentido tecnológico las redes neuronales artificiales no son más que ecuaciones matemáticas interrelacionadas que imitan la estructura de los circuitos cerebrales, con varias capas de nodos entre la entrada y la salida similares a neuronas que procesan información en pasos sucesivos, por lo que traen consigo un alto grado de especificidad en su resultado con menor orientación humana.

Con lo antedicho, se logró despojar a las máquinas del algoritmo, por lo que oficialmente la era de la inteligencia artificial había comenzado, creciendo sus usos y mejorando inmensamente. Ese verdadero aprendizaje de las máquinas permitió perfeccionar los anteriores sistemas y hacer viables los reconocimientos virtuales de imágenes y de voz con los asistentes virtuales que revolucionaron el mercado: Alexa y Siri en los teléfonos (The Path to Singularity, 2023), incluso en innovaciones que tiempo atrás parecía ciencia ficción, con un ejemplo claro en la conducción autónoma de vehículos. Sin embargo, lo que hace la inteligencia artificial es navegar a través de los datos y clasificar la información incorporada previamente, por lo cual por mucho que el desarrollo de la computación hubiera permitido el aprendizaje, las máquinas

solo jerarquizaban datos. El próximo paso sería potenciar aún más el *deep learning*, en lugar de clasificar datos, crearlos; insertar datos para que aprendiendo de ellos creara algo nuevo.

Fue entonces que en el año 2015 nació OpenAI y se creyó que la inteligencia artificial generativa era un reto viable. Los *chatbot* habían surgido mucho antes, el propio Eliza en los ´60 había demostrado ciertos avances, tantos años después, era evidente que se podría hacer algo mucho más potente: un programa que pudiera mantener conversaciones fluidas con los usuarios. Por ello, utilizando el modelo de lenguaje GPT-3 y usando 175 mil millones de parámetros fue alimentada con toda la información disponible en la red. Así, en noviembre de 2022, la aplicación conocida como ChatGPT fue lanzada al público (Top de Impacto, 2023).

Desde el 2022 a la actualidad puede decirse que el desarrollo de la inteligencia artificial ha ido *in crescendo* de forma tal que en su avance se emplea un enfoque interdisciplinario que abarca áreas como matemáticas, ciencias de la computación, lingüística, psicología, entre otras. En el presente está utilizando para analizar radiografías, resonancias magnéticas o tomografías, y para ayudar a médicos en el diagnóstico de enfermedades. Por ejemplo, proyectos como IDx-DR utilizan inteligencia artificial para detectar la retinopatía diabética en imágenes de retina (Healthvisors, 2020), o Watson for Oncology, un sistema de soporte de decisiones que ayuda a oncólogos a tomar decisiones informadas sobre el tratamiento del cáncer (Mellado et al., 2023).

Por otra parte, al ser capaces de analizar gran cantidad de datos, como imágenes satelitales, datos climáticos e históricos está siendo utilizada para predecir y gestionar mejor los desastres naturales. Por ejemplo, el modelo GraphCast y su proyecto “DeepMind for Flood Forecasting” para predecir inundaciones con mayor precisión (AI Policy Primer, 2024). Asimismo, se utilizan en sistemas de seguridad para analizar comportamientos

sospechosos en entornos públicos y privados. Por ejemplo, la tecnología de Pangiam junto a la Administración de Seguridad del Transporte de los Estados Unidos (TSA) se utilizó para identificar posibles artículos prohibidos en el equipaje de mano, como la incorporación del reconocimiento facial TSA PreCheck para la verificación de identidad de los pasajeros (Oremus, 2025).

En resumen, la evolución histórica y tecnológica de la inteligencia artificial revela un proceso de construcción científica marcado por desafíos epistemológicos, avances metodológicos y una creciente complementariedad interdisciplinaria. Lejos de constituir un desarrollo lineal o homogéneo, la inteligencia artificial ha atravesado periodos de auge y estancamiento, reflejando tanto las limitaciones tecnológicas de cada época como las expectativas sociales puestas en estas máquinas inteligentes. No obstante, el progreso sostenido en algoritmos, procesamiento de datos y capacidad computacional ha consolidado como una disciplina autónoma y transformadora, capaz de superar barreras iniciales y expandirse hacia ámbitos prácticos y teóricos con significativas repercusiones jurídicas. Este recorrido histórico subraya la necesidad de un análisis regulatorio y ético que reconozca la complejidad técnica y social, evitando tanto el entusiasmo acrítico como la desconfianza infundada, y fomentando así un marco de desarrollo tecnológico informado, responsable y sostenible.

1.2. Definición de inteligencia artificial desde un enfoque jurídico. Características y tipos

En la actualidad no se ha llegado a una definición acabada y completa de lo que realmente es la inteligencia artificial y este problema tiene su raíz inmediata en que aún no está claro el propio concepto de inteligencia humana (Sheikh et al., 2023). Abarcar hoy en un planteamiento toda la esencia es improbable, sobre todo teniendo en cuenta que se trata de un fenómeno que todavía no ha terminado de evolucionar, desde que comenzara a expandirse en el pasado siglo.

McCarthy et al. (1955), por su parte, tuvieron la visión de que "*every aspect of learning or any other future of intelligence can in principle be so precisely described that a machine can be made to simulated it*"; o sea, siguiendo este presupuesto, existe una correspondencia entre la inteligencia artificial y la capacidad de una máquina de imitar el comportamiento humano inteligente. Sin embargo, cuando se ofreció esta definición, ni siquiera existía un concepto incipiente de inteligencia artificial o al menos no como se conoce hoy. En adición a esto, cabe mencionar que los avances de la ciencia solo hacían soñar; por tanto, esta definición es abstracta, dada a luz en un momento histórico donde no existía la inteligencia artificial y esta solo era el sueño de unos pocos.

Por su parte, la Real Academia de la Lengua Española (RAE) define la inteligencia artificial como la "disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico" (Real Academia Española, 2024). Aunque aceptable, a los fines de la ciencia esta definición también es limitada, dado que no define exactamente la categoría en cuestión, pues, de conjunto con los programas informáticos, la inteligencia artificial incluye equipos y otras herramientas como el *hardware*; por ejemplo.

Así, como analiza Muñoz Vela (2021):

> La inteligencia artificial pretende descomponer la inteligencia humana en sus procesos más simples para poder describirlos y expresarlos lógicamente, es decir, hacerla computable, y transferirlos mediante algoritmos y programación para reproducir o emular la inteligencia humana en sistemas, programas informáticos y máquinas, es decir, en *software y hardware* (p. 63).

A su vez, se une a esto una consideración asociada a que en la actualidad en el concepto de inteligencia artificial intervienen ideas propias tanto de la filosofía y como de la ingeniería, que la

conciben como una máquina de cómputo capaz de pensar como humano. Para Serna et al. (2017), la inteligencia artificial es una rama de la ciencia de la computación, relacionada con la forma de dar a los computadores la sofisticación para desenvolverse inteligentemente, y hacerlo en ámbitos cada vez más amplios.

Se puede definir, además, como aquel campo que agrupa ciencias de la computación con robustos conjuntos de datos para facilitar la resolución de problemas (International Business Machines Corporation, 2024). Abarcando en la actualidad una gran diversidad de subcampos, que van desde áreas de propósito general, como el aprendizaje y la percepción, a otras más específicas como el ajedrez, la demostración de teoremas matemáticos, la escritura de poesía y el diagnóstico de enfermedades. Ello conduce a que, partiendo de la capacidad de la inteligencia artificial de sintetizar y automatizar tareas intelectuales pueda establecerse como un ente potencialmente relevante para cualquier ámbito de la actividad intelectual humana (Russell y Norving 2009).

Lo planteado por Russell y Norving (2009), constituye a juicio de los autores uno de los enfoques más completos en el abordaje que se ha hecho hasta el momento de la inteligencia artificial. Pues considera a la misma desde un punto de vista en el cual intervienen al mismo tiempo varias áreas del conocimiento como la filosofía, la matemática, la economía, la psicología y la informática. Al respecto ténganse en cuenta las palabras textuales de los referidos autores:

> Los filósofos (desde el año 400 a.C.) facilitaron el poder imaginar la IA, al concebir la idea de que la mente es de alguna manera como una máquina que funciona a partir del conocimiento codificado en un lenguaje interno, y al considerar que el pensamiento servía para seleccionar la acción a llevar a cabo. Las matemáticas proporcionaron las herramientas para manipular tanto las aseveraciones de certeza lógicas, como las inciertas de tipo probabilista. Asimismo, prepararon el terreno para un

> entendimiento de lo que es el cálculo y el razonamiento con algoritmos. Los economistas formalizaron el problema de la toma de decisiones para maximizar los resultados esperados. Los psicólogos adoptaron la idea de que los humanos y los animales podían considerarse máquinas de procesamiento de información. Los lingüistas demostraron que el uso del lenguaje se ajusta a ese modelo. Los informáticos proporcionaron los artefactos que hicieron posible la aplicación de la IA (Russell y Norving, 2009, pp. 33-34).

Por su parte, la definición legal que hace el Proyecto de Ley argentino "Marco legal para la regulación del desarrollo y uso de la Inteligencia Artificial" establece que esta "se refiere a sistemas computacionales diseñados para realizar tareas que requieren habilidades humanas, como el aprendizaje, la percepción, el razonamiento y la toma de decisiones". En igual sentido, el Proyecto de Ley del Senado colombiano 253/22, que supone un reglamento de inteligencia artificial, plantea que es una "disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico".

Aunque las definiciones legales contenidas en los anteriormente mencionados proyectos de ley argentino y colombiano reconocen acertadamente la capacidad de aprendizaje e inteligencia de estos sistemas, vale señalar que su conceptualización se limita solo a entender la inteligencia artificial exclusivamente como programas informáticos o sistemas computacionales. Resultando ser un enfoque restringido que deriva en insuficiente para abarcar la complejidad contemporánea de la inteligencia artificial.

Sobre todo, tomando en cuenta que en el marco del internet de las cosas (IoT) y la integración creciente de dispositivos físicos inteligentes; como vehículos autónomos, drones, robots y otros sistemas ciberfísicos, la inteligencia artificial trasciende el mero ámbito virtual y digital para arraigarse en contextos físicos y

operativos. Esta expansión implica que la inteligencia artificial no solo procesa datos y ejecuta algoritmos en entornos informáticos cerrados, sino que interactúa activamente con el mundo real, modificándolo y adaptándose a múltiples escenarios dinámicos y distribuidos.

Por tanto, una definición legal y regulatoria que aspire a ser plenamente operativa y contemporánea debe reconocer y contemplar esta dimensión híbrida y multisistémica de la inteligencia artificial, incorporando tanto su componente digital como su manifestación física integrada, con todas las implicancias técnicas, éticas y jurídicas que ello conlleva. De no considerarse esta amplitud, los marcos normativos quedarían constreñidos a una visión parcial que podría limitar la eficacia de la regulación, la protección de derechos y la gestión de riesgos derivada de tecnologías cada vez más autónomas, interconectadas y ubicuas.

Finalmente, en relación con lo abordado, la definición dada por el Grupo de Expertos de alto nivel en inteligencia artificial de la Comisión Europea proporciona una amplia visión del alcance que tiene hoy en la práctica la inteligencia artificial y del que tendrá en el futuro cercano, resultando ser, igualmente, una concepción considerablemente detallada de la categoría objeto de análisis:

> La inteligencia artificial (IA) consiste en sistemas de software (y posiblemente también de hardware) diseñados por humanos que … actúan en el mundo físico o digital al percibir su entorno mediante la adquisición de datos, interpretando los datos estructurados o no estructurados recopilados, razonando sobre el conocimiento o procesando la información derivada de estos y decidiendo la(s) mejor(es) medida(s) a tomar para lograr el objetivo dado … Como disciplina científica, la IA incluye varios enfoques y técnicas, como el aprendizaje automático … el razonamiento automático … y la robótica (High-Level Expert Group on Artificial Intelligence, 2019).

El Grupo de Expertos precitado establece una doble dimensión de la inteligencia artificial, por lo que es admisible concluir que esta es tanto un sistema como una disciplina científica. Como sistema se trata pues, en principio y sobre todo, de *softwares* que parten del manejo de datos y que, imitando las funciones cognitivas humanas, los procesan y generan decisiones que pudieran catalogarse como inteligentes en base a los datos analizados.

Visto lo anterior, y sin ánimo de establecer una definición propia, se considera oportuno delimitar los elementos que no deberían faltar en una definición de esta figura a partir del análisis de las definiciones ofrecidas. Ellos son:

- Inclusión de programas informáticos y dispositivos físicos.
- Uso de algoritmos y datos.
- Aprendizaje (datos, conductas, etc.).
- Toma de decisiones.
- Realización de tareas de forma autónoma.

Básicamente el funcionamiento de estos sistemas consiste en:

> Primero … identificar el problema -aprendizaje-, para luego analizar las situaciones del pasado y estudiar todas las posibles variables relacionadas con el problema, posteriormente a través de un sistema de estadísticas -entrenamiento- predice el resultado futuro del problema y por último una vez el sistema tiene todos los datos, proporciona la solución -resultados- más factible para el problema (Domingues Villarroel, 2023).

No obstante, es importante señalar que no todos los sistemas de inteligencia artificial operan con la misma eficacia ni presentan un funcionamiento homogéneo. La variabilidad en su desempeño está directamente vinculada al nivel de desarrollo tecnológico alcanzado por cada sistema, así como a la calidad y amplitud

de los datos con los que han sido entrenados. Factores como los algoritmos empleados, la capacidad computacional disponible, el diseño arquitectónico y la especialización para tareas concretas influyen considerablemente en su eficiencia y precisión. Por ende, la evaluación y selección de un sistema de inteligencia artificial deben considerar su adecuación específica al problema o ámbito de aplicación, reconociendo que sistemas más avanzados y con mejor desarrollo tendrán mayor capacidad para procesar información compleja, adaptarse a entornos dinámicos y ofrecer resultados confiables. En suma, la disparidad en el funcionamiento y efectividad de estos sistemas refleja la evolución tecnológica y la diversidad de aplicaciones que la inteligencia artificial abarca actualmente. De ahí que el profesor Tapia Hermida (2021), haya establecido dos tipos de inteligencia artificial de acuerdo al desarrollo de la misma:

- IA de bajo riesgo, sencilla o débil
- IA de alto riesgo, fuerte o compleja

Por un lado, la inteligencia artificial sencilla o débil está diseñada para cumplir tareas concretas y predefinidas, basándose en algoritmos y con una cierta intervención del ser humano. Siendo estas las inteligencia artificial más conocidas y accesibles por la mayoría, integrada sobre todo en herramientas de teléfonos inteligentes y ordenadores, por solo citar dos ejemplos. Son aquellos sistemas que únicamente parecen, conductualmente, tener un pensamiento inteligente similar al humano (simulan tener inteligencia), pero que en realidad no pasan de ser sistemas muy especializados que aplican técnicas más o menos complejas a la resolución de problemas muy concretos, hallándose lejos de mostrar cualquier síntoma revelador de estados cognitivos.

Por otra parte, la inteligencia artificial fuerte o compleja es aquella que se enfoca en la imitación de las habilidades cognitivas humanas, siendo este tipo de sistemas inteligentes los más avanzadas conceptualmente porque implican poder encontrar

soluciones a tareas desconocidas por estas máquinas, sin haberlas experimentado previamente (la variante fuerte, es la que por el momento no se ha desarrollado en gran medida), o sea, la Inteligencia Artificial compleja tienen capacidad de aprendizaje y desarrollo propio, o al menos no tan vinculado a la conducta dispositiva del ser humano.

En otro orden, puede clasificarse a la inteligencia artificial además, atendiendo a los componentes que la integren, como establece Muñoz Vela (2021):

- IA no integrada.
- IA integrada

Distinguiéndose entre ellas principalmente por su constitución, en el caso mínimo, únicamente por *software*, sin presencia física material, se entiende por inteligencia artificial no integrada a aquellos sistemas que funcionan exclusivamente como programas informáticos. Ejemplos claros de esta categoría incluyen asistentes virtuales, *software* de análisis de imágenes, motores de búsqueda o sistemas de reconocimiento de voz y rostro, los cuales operan en entornos digitales y no requieren una corporeidad tangible para realizar sus funciones.

En contraste, la inteligencia artificial integrada comprende aquellos sistemas en los que el *software* se encuentra incorporado en un soporte físico, es decir, en dispositivos o máquinas concretas. Esta integración física permite la interacción directa con el entorno, ofreciendo capacidades avanzadas de percepción, acción y autonomía. Ejemplos típicos de inteligencia artificial integrada son los robots, drones, vehículos autónomos y dispositivos conectados dentro del ecosistema del Internet de las Cosas (IoT). En estos casos, el *hardware* y el *software* funcionan de manera conjunta para ejecutar tareas complejas, donde la inteligencia artificial no solo procesa datos, sino que también responde y actúa en el mundo real en tiempo real.

Esta distinción es fundamental para comprender el alcance y las aplicaciones variadas de la inteligencia artificial, ya que mientras la inteligencia artificial no integrada se limita a procesos cognitivos y análisis dentro de entornos virtuales, integrada añade una dimensión física que amplía su impacto y funcionalidades hacia ámbitos cada vez más diversos e interactivos en la vida cotidiana y la industria.

Además de lo abordado hasta el momento, es significativo el criterio de Uribe (2024) al sintetizar a la inteligencia artificial de acuerdo a cuatro enfoques que permiten capturar la diversidad en su clasificación, destacando sus múltiples perspectivas y facilitando la integración de ese conocimiento. De ese modo, el autor plantea como criterios clasificadores los siguientes ámbitos: campo, funcionalidad, capacidad y tecnología.

El criterio de la funcionalidad parece estar en consonancia con otras clasificaciones abordadas por cuanto se equipara a la distinción entre inteligencia artificial estrecha y la general que ya se analizó, aunque sumándole un tercer grupo concebido como superinteligencia artificial, descrita como cualquier forma de inteligencia que excede considerablemente el desempeño cognitivo del ser humano en casi todos los ámbitos.

Se distingue igualmente entre inteligencia artificial experimentada y la teórica basándose en el hecho de que no todas las tipologías de inteligencia artificial de las que se habla han sido implementadas o probadas en la práctica, con lo cual una inteligencia artificial experimentada es aquella que ha sido desarrollada, implementada y sometida a pruebas en diversos contextos prácticos, mientras que una inteligencia artificial teórica engloba aquellos conceptos y prototipos que aún no han sido completamente materializados o aplicados en la realidad operativa.

En correspondencia con esa línea de razonamiento se escinde además la inteligencia artificial, por su funcionalidad, en máquinas reactivas, de memoria limitada, con teoría de la mente y

autoconsciente. Criterio que establece una diferenciación a partir del nivel de funciones que pueda realizar el sistema, partiendo de un sistema diseñado para ejecutar tareas específicas, a uno que puede recordar eventos y resultados anteriores para mejorar la toma de decisiones futuras, luego a uno; aún en fase teórica que pretende comprender y responder a las emociones y pensamientos de otras entidades, y por último a un concepto aún más avanzado y teórico que incluiría la capacidad de reconocer sus propios estados internos.

Finalmente, según el criterio clasificador que se guía por la tecnología, se establece una divergencia de sistemas de inteligencia artificial que incluye: tecnología de visión por computadora, el procesamiento de lenguaje natural, el aprendizaje automático y el aprendizaje profundo. No obstante, no puede estarse completamente de acuerdo con esta idea toda vez que varias de esas tecnologías pueden coexistir en un mismo sistema haciendo imposible su distinción en ese sentido, además de que el conjunto de tecnologías asociado a la inteligencia artificial es mucho más amplio (Banco Interamericano de Desarrollo, 2024) y se encuentra en constante evolución por lo que establecer un criterio de distinción de acuerdo a él resulta, cuanto menos, inviable.

La inteligencia artificial, desde una perspectiva jurídica, se configura pues, como un fenómeno complejo y en constante evolución que desafía las definiciones tradicionales debido a su naturaleza interdisciplinaria. Reconociéndose que esta trasciende la mera programación informática para incluir componentes de *hardware* y sistemas integrados que imitan funciones cognitivas humanas mediante el uso de algoritmos, datos y capacidad autónoma para tomar decisiones. Lo cual asociado a sus tipologías condiciona que exista una diversidad funcional que implica distintos niveles de intervención humana y consecuencias jurídicas diferenciadas. De modo que este marco conceptual, enriquecido por distintas

perspectivas filosóficas, técnicas y normativas, evidencia la necesidad de un abordaje regulatorio que contemple tanto su dimensión sistémica como disciplinaria, reconociendo la inteligencia artificial como herramienta transformadora con implicaciones jurídicas profundas que demandan una definición flexible y ajustada a los rápidos avances tecnológicos y sociales contemporáneos.

1.3. La inteligencia artificial como fenómeno tecnológico: su trascendencia jurídica

La Inteligencia Artificial (IA), es fruto del desarrollo tecnológico vertiginoso de los últimos 30 años, si bien sus orígenes se remontan a los años 50 del pasado siglo. Esta implica cambios susceptibles de transformar la realidad económica y social, además posee un elevado potencial para transversalizar la sociedad. La tecnología digital supone una revolución, según González Moreno (2023), se cataloga como la 4ta Revolución industrial; que se refiere a la transformación digital y tecnológica que está teniendo lugar en la sociedad y en la economía, caracterizada por la integración de tecnologías digitales en una amplia gama de industrias, desde la comunicación hasta la banca, la salud y la fabricación, y que desempeña un papel cada vez más importante en la sociedad y en la economía global, implica innovaciones tecnológicas que conducen progresivamente y a escala acelerada por su influencia en diversas esferas como la medicina, el derecho, la educación, las comunicaciones, la energía, la transportación, la informática, la robótica, entre otras y las mejoras continuas de las herramientas tecnológicas asociadas a la inteligencia artificial (Martínez, 2012). Al respecto Sein (2019) expone:

> Tales innovaciones tecnológicas tienen como base la obtención, gestión y procesamiento de la información. El elemento clave de esta revolución tecnológica es la información, la recopilación ingente de datos, el Big data, y el análisis y tratamiento

de los datos, basado en la inteligencia artificial. Y ello se proyecta en todos los órdenes de la vida social y económica, y, en particular, en lo que aquí más interesa destacar, sobre la actividad productiva, tanto en lo que se refiere a la forma de concebir el negocio, como a la gestión de los recursos humanos.

La inteligencia artificial se basa en sistemas computacionales que deben poseer la capacidad de generar rasgos particulares que se asocian al comportamiento humano y que logren un desempeño similar cuando se enfrentan a la solución de problemas, mediado por la tecnología y con un rendimiento superior.

Son varias las características de un sistema inteligente según refiere Martínez (2012):

a) Es un programa computacional que puede estar vinculado a otros elementos de transferencia y conversión de información;

b) Posee gran cantidad de conocimiento y datos sobre un problema y realiza un razonamiento similar al que haría un humano frente a un problema;

c) Puede operar con datos cuantitativos y con datos cualitativos;

d) Puede emitir conclusiones a partir de datos vagos o incompletos;

e) Puede interrumpir una línea de razonamiento para ocuparse de otra y ser capaz de volver a su línea anterior y,

f) Puede tener interfaces externas, o consultar una base de datos, esto es, ser capaz de comunicarse con otros y tener la posibilidad de operar en ambientes distintos.

La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura señala que la inteligencia artificial (IA) es una rama de la informática que desarrolla programas capaces de emular procesos propios de la inteligencia humana. Es decir, las máquinas pueden analizar el entorno y realizar determinadas acciones de manera más o menos autónoma con el fin de lograr objetivos concretos.

La inteligencia artificial (IA) es tecnología, se desarrolla mediante el procesamiento de grandes volúmenes de datos a través de algoritmos avanzados que imitan la forma en que los humanos aprenden, razonan y toman decisiones. Tales sistemas informáticos analizan patrones en la información, realizan el procesamiento y generan resultados basados en razonamientos lógicos y genera determinadas respuestas.

En esencia, la inteligencia artificial combina datos, modelos matemáticos y capacidades computacionales para llevar a cabo tareas que antes requerían la intervención humana, desde el reconocimiento de imágenes hasta la generación de texto o la toma de decisiones estratégicas (Martínez, 2024).

La Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno sobre la definición de inteligencia artificial de 2018, establece tres precisiones respecto a la definición de este concepto (Consejo de Europa, 2018) estas son.

- Inteligencia artificial no es un objeto único y homogéneo: en realidad es un conjunto de ciencias y técnicas (matemática, estadística e AI no es un objeto único y homogéneo: en realidad es un conjunto de ciencias y técnicas (matemática, estadística e informática) capaces de procesar datos para diseñar tareas de procesamiento informático muy complejas.

- Los motores de inteligencia artificial no producen inteligencia per se, sino que proceden utilizando un enfoque inductivo: la idea es asociar de manera casi automatizada un conjunto de observaciones (entradas) con un conjunto de resultados (salidas) posibles utilizando varias propiedades preconfiguradas.

- La fiabilidad del modelo (o función) construida depende en gran medida de la calidad de los datos utilizados y de la elección de la técnica de aprendizaje automático.

Resulta importante el tratamiento de la inteligencia artificial como herramienta tecnológica, en el Reglamento de la UE sobre IA, al considerarla como un conjunto de tecnologías en rápida evolución que contribuye a generar beneficios económicos, medioambientales y sociales muy diversos en todos los sectores económicos y las actividades sociales. Su uso puede proporcionar ventajas competitivas esenciales a las empresas y facilitar la obtención de resultados positivos desde el punto de vista social y medioambiental en diversos ámbitos (Parlamento Europeo, 2024).

En este sentido, el Reglamento reconoce como característica importante de la inteligencia artificial que es su capacidad de inferencia. Esta capacidad de inferencia se refiere al proceso de obtención de resultados de salida, como predicciones, contenidos, recomendaciones o decisiones, que puede influir en entornos físicos y virtuales, y a la capacidad de los sistemas de inteligencia artificial para deducir modelos o algoritmos, o ambos, a partir de información de entrada o datos.

Otra característica de la inteligencia artificial aportada por Flores Salgado & Martínez Xelhuantzi (2024), es incorporar a la informática el concepto de aprendizaje, es decir, la utilización de la experiencia y la deducción lógica para auto mejorarse, además utiliza símbolos no matemáticos.

Los basamentos tecnológicos y el rápido desarrollo de la inteligencia artificial con la diversidad de programas y aplicaciones, genera consecuencias positivas y negativas. Dentro de sus efectos favorables se encuentran la automatización de procesos, agiliza la toma de decisiones, fomenta e incentiva la creatividad, reduce las probabilidades de fallos o errores y por ende mejora la precisión y la calidad de los resultados, incide en el desarrollo de la medicina y favorece las conexiones humanas mediante redes cada vez más complejas.

Sin embargo, la inteligencia artificial posee impactos adversos, por la potencialidad que tiene de generar daños al medio ambiente,

afectaciones graves a los derechos humanos al incidir sobre la esfera de la intimidad, la privacidad, pueden generar sesgos de discriminación por diversas razones, otro impacto negativa es la posible exclusión laboral y la sustitución de los seres humanos por la tecnología; utiliza grandes volúmenes de datos, cuya manipulación y uso inapropiado puede generar otros impactos negativos en la seguridad y otras esferas de la vida humana.

El Libro Blanco sobre Inteligencia Artificial de la Comisión Europea (2020) define los problemas de la inteligencia artificial y estos son:

> Los daños pueden ser tanto materiales (para la seguridad y la salud de las personas, con consecuencias como la muerte, y menoscabos al patrimonio) como inmateriales (pérdida de privacidad, limitaciones del derecho de libertad de expresión, dignidad humana, discriminación en el acceso al empleo, etc.) y pueden estar vinculados a una gran variedad de riesgos. Los principales riesgos relacionados con el uso de la inteligencia artificial afectan a la aplicación de las normas diseñadas para proteger los derechos fundamentales (como la protección de los datos personales y la privacidad, o la no discriminación) y la seguridad, así como a las cuestiones relativas a la responsabilidad civil.

Se hace imprescindible la intervención del Derecho en los avances tecnológicos que acompañan los desarrollos de la inteligencia artificial, cabría preguntarse si es posible regular la inteligencia artificial, y esta es una vertiente de la relación bidireccional que se produce entre ambos elementos; existe otra, que se centra en la utilidad y la intervención profunda en la actuación de los operadores legales, dígase abogados, jueces, sistema penal, la actividad de asesoramiento empresarial, entre otros. La tercera vertiente está dada por la necesidad de protección a los productos de inteligencia artificial por medio de la propiedad intelectual u otros recursos legales per se.

En cuanto a la primera cuestión cabe establecer que debe procurarse que el desarrollo que supone la inteligencia artificial armonice con la tutela de los derechos humanos, que se garantice el respeto de principios éticos, la norma jurídica otorga legitimidad a la, pero permite establecer límites y presupuestos para el desarrollo ulterior de la tecnología y la actividad de quienes crean estos productos. La la inteligencia artificial debe poseer un marco ético y jurídico, expresado en instrumentos legales, políticas y principios, basado en el respeto absoluto de los derechos humanos, reconocidos en instrumentos internacionales y en los ordenamientos jurídicos internos.

La inteligencia artificial debe quedar sujeta al control de los humanos, al respecto Sánchez y Toro (2021) afirman:

> Asegurar la veeduría y la intervención humana supone … garantizar la presencia y el control de autoridades públicas y privadas, externas a los desarrolladores de sistemas de inteligencia artificial en sus procesos de diseño y desarrollo, con el ánimo de asegurar la defensa de intereses diferentes a los privados. Esto significa que la veeduría debe estar a cargo de actores estatales, la sociedad civil y otras empresas pares que evalúen el respeto a los derechos humanos en los sistemas de inteligencia artificial, para que estos prevalezcan sobre los intereses particulares y lucrativos de las empresas.

La inteligencia artificial trasciende a la actividad jurídica a través de programas y sistemas jurídicos de expertos en sus diversas variantes (modelo de procesamientos simbólicos, basado en el conocimiento y el modelo conexionista basado en redes neuronales), que acompañan y a veces sustituyen al operador legal, Martínez (2012) refiere que “un sistema de expertos jurídicos es un sistema computacional que puede plantear posibles soluciones a determinados asuntos jurídicos aplicando el conocimiento experto en la materia, así como explicar sus razonamientos”.

Otros autores abogan por sistemas avanzados de inteligencia artificial aplicados al derecho. Solar Cayón (2024), consideran que colaboran estrechamente en la construcción de soluciones adaptadas a las necesidades específicas de éste. Por ello, estos nuevos sistemas de "computación cognitiva" se configuran como plataformas interactivas que combinan diversas herramientas tecnológicas para posibilitar que el usuario customice su propia solución, de modo que el sistema seleccione, analice y sintetice la información jurídica de una manera totalmente ajustada al problema específico.

Esta es la principal aplicación de la inteligencia artificial al campo del derecho, constituyendo un sistema que contribuye a la adopción de la decisión judicial. También se encuentran en este campo los asistentes legales, que hacen revisión de extensos volúmenes de documentos, dictaminan, redactan instrumentos legales, sistematizan instrumentos jurídicos, sentencias, casos y ofrecen al abogado o asesor legal una herramienta útil para su desempeño.

La inteligencia artificial reviste gran utilidad e implica una transformación radical de la actividad jurídica, en la organización y procesamiento de textos jurídicos, la interactividad con los usuarios, la modelación de operaciones realizadas por los juristas, la resolución de casos, la redacción de instrumentos jurídicos, partiendo del razonamiento lógico y por casos, la modelación de proyectos de leyes con el fundamento jurídico correspondiente, la elaboración y dictamen de contratos, tasas e impuestos. Frente a esto cabe identificar como obstáculos la falta de aceptación de las herramientas de inteligencia artificial y de dominio de las mismas por los jurídicos, lo que plantea la exigencia de formación y capacitación en el campo de la informática y la inteligencia artificial desde el pregrado.

Sobre la influencia en la calidad y eficiencia de los servicios jurídicos se ha demostrado que son precisamente los clientes los

que buscan la flexibilidad y eficiencia que la inteligencia artificial brinda, según un reporte de Onecom, el 70% de los consumidores encuestados estarían dispuestos a que un lawbot los atendiera si el servicio era eficaz, accesible y relevante (Veracruz, 2024).

Otro análisis interesante es el de Campuzano Gallegos (2019), que afirma que el Derecho es un sistema de representación que varía de cultura en cultura y dentro de cada cultura hay diversas concepciones normativas, es decir, no hay consenso único sobre qué es el Derecho, cómo debe concebirse, así como una diversidad de Teorías Generales del Derecho; tampoco hay consensos explícitos sobre los hábitos de razonamiento de sus operadores, la incidencia del Derecho en los procesos de construcción social de la realidad, los insumos cognitivos jurídicos. A ello se agrega que el derecho regula relaciones sociales jurídica y socialmente relevantes, y estas indiscutiblemente varían de un país a otro, de ahí la dificultad de uniformar programas y productos de inteligencia artificial.

Los sistemas prevalentes en la actividad jurídica son los de *legal question answering*, o sea de búsqueda de respuestas jurídicas, capaces de a partir de un proceso de entrenamiento con sus algoritmos de aprendizaje profundo, sea capaz de responder automáticamente a preguntas legales formuladas en lenguaje natural alcanzando a generar en breve tiempo información procedente de diversas fuentes, extrayendo argumentos, razonamientos lógicos y elaborando una respuesta jurídicamente fundamentada (Cayon, 2020).

En otro orden, la cuestión relativa a cómo proteger los sistemas de inteligencia artificial. Es fundamental partir de que estos sistemas no solo se configuran con softwares, sino que pueden estar integrados por distintos elementos y componentes, variables en función de su tipología, como pueden ser además hardwares, algoritmos y bases de datos. Cada uno de ellos puede comportar independientemente una forma de tutela jurídica, de manera

que la defensa legal de todo el sistema de inteligencia artificial dependerá de la conjunción de formas de protección aplicables a los mismos, este es otro reto que queda sujeto al campo de la propiedad intelectual y los modelos de protección legal que se estructuren, lo que se abordara en el Capítulo III.

1.4. Principios del uso de la inteligencia artificial

La implementación de la inteligencia artificial y el desarrollo de nuevos sistemas entraña riesgos de diferentes niveles, siendo los más graves aquellos que trascienden a la esfera de los derechos humanos, afectando la dignidad de las personas y la vida, por ende, resulta necesario establecer principios y directrices que fomenten la transparencia de los procesos, el control y supervisión y la reducción de riesgos por discriminación y otros efectos negativos.

Los principios, recomendaciones y directrices, tiene como propósito que las personas que participan en el proceso de diseño, generación, implementación y despliegue de los sistemas inteligentes (ciclo de vida) cumplan con un grupo de valores, principios, y técnicas que guíen su comportamiento de forma que su utilización sea en pro del ser humano y del bien común, con el respeto de los derechos fundamentales. los principios expresan características a alcanzar en el desarrollo de la inteligencia artificial.

Son varios los instrumentos legales a nivel internacional, regional y de los países que contemplan los principios de inteligencia artificial fundamentalmente enfocados en el componente ético. Sirva de ejemplo la Recomendación de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura sobre ética de la inteligencia artificial, Principios rectores y funciones de la gobernanza internacional, Grupo Europeo de Ética de la Ciencia y de las Nuevas Tecnologías (grupo consultivo de la Comisión) publicó en marzo de 2018, declaración sobre

inteligencia artificial, robótica y sistemas autónomos, Directrices para su uso en los parlamentos europeos (2024) y el Libro blanco sobre la inteligencia artificial - un enfoque europeo orientado a la excelencia y la confianza de 2020, inteligencia artificial desde los cimientos: desafíos y oportunidades en el contexto de América Latina y el Caribe, Declaración sobre inteligencia artificial inclusiva y sostenible para las personas y el planeta, dentro de las iniciativas del sector privado se encuentra los 23 Principios de Asilomar, lineamientos globales (Future of Life Institute, 2017).

Para el desarrollo de este acápite se utilizó la herramienta de Atlasti (9), se examinaron 15 documentos, se establecieron 22 códigos que generaron 55 citas, con el propósito de establecer los principios concurrentes y efectuar un análisis de contenido en un mapa de palabras derivado de estos documentos identificando los términos más reiterados.

Los principios reconocidos de manera coincidente a partir de la bibliografía consultada son:

Derecho a la intimidad y la protección de datos: este es un derecho vinculado con la privacidad, es esencial la protección de la dignidad, la autonomía y la capacidad de actuar de los seres humanos, debe ser respetada, protegida y promovida a lo largo del ciclo de vida de los sistemas de inteligencia artificial. Es importante que los datos para estos sistemas se recopilen, utilicen, compartan, archiven y supriman de forma coherente con el derecho internacional y acorde con los valores y principios enunciados en la presente Recomendación, respetando al mismo tiempo los marcos jurídicos nacionales, regionales e internacionales pertinentes.

El principio de privacidad subraya la importancia de proteger los datos personales y no compartir información que no es de la incumbencia de empresas, particulares e incluso gobiernos. Esto significa que los diseñadores y constructores de sistemas de inteligencia artificial deben dar prioridad al diseño de sistemas

que sean transparentes y responsables con respecto a la gestión y el uso de los datos, y que permitan a las personas ejercer el control sobre su propia información (Comisión Europea 2020).

Supervisión humana: La supervisión humana se refiere, no solo a la supervisión humana individual, sino también a la supervisión pública inclusiva, según corresponda. Puede ocurrir que, en algunas ocasiones, los seres humanos decidan depender de los sistemas de inteligencia artificial por razones de eficacia, pero la decisión de ceder el control en contextos limitados seguirá recayendo en los seres humanos, ya que estos pueden recurrir a los sistemas de inteligencia artificial en la adopción de decisiones y en la ejecución de tareas, pero este sistema nunca podrá reemplazar la responsabilidad final de los seres humanos y su obligación de rendir cuentas. Por regla general, las decisiones de vida o muerte no deberían cederse a los sistemas de inteligencia artificial (Cruz et al., 2024; Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 2022).

Proporcionalidad e inocuidad: Las tecnologías de la inteligencia artificial no garantizan necesariamente, por sí mismas, la prosperidad de los seres humanos ni del medio ambiente y los ecosistemas. En caso de que pueda producirse cualquier daño para los seres humanos, los derechos humanos y las libertades fundamentales, las comunidades y la sociedad en general, o para el medio ambiente y los ecosistemas, debería garantizarse la aplicación de procedimientos de evaluación de riesgos y la adopción de medidas para impedir que ese daño se produzca (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 2022).

Seguridad y protección: Los daños no deseados (riesgos de seguridad) y las vulnerabilidades a los ataques (riesgos de protección) deberían ser evitados y deberían tenerse en cuenta, prevenirse y eliminarse a lo largo del ciclo de vida de los sistemas de inteligencia artificial para garantizar la seguridad y la protección

de los seres humanos, del medio ambiente y de los ecosistemas. La seguridad y la protección de la inteligencia artificial se propiciarán mediante el desarrollo de marcos de acceso a los datos que sean sostenibles, respeten la privacidad y fomenten un mejor entrenamiento y validación de los modelos de inteligencia artificial que utilicen datos de calidad (Comisión Europea, 2020). Bajo este contenido también es denominado: principio de fiabilidad, solidez y precisión. Solidez para evitar las vulneraciones de la seguridad, los ciberataques y los usos indebidos de los datos personales. Precisión que socave las posibles fugas de datos, las noticias falsas y la desinformación. Ambas respetuosas con la privacidad de los datos, dentro del acervo jurídico de la Unión en esta materia. Fiabilidad, basada en la calidad de los datos empleados, sin sesgos que dañen los datos acumulados y con unos algoritmos limpios, sin estrategias ocultas.

Principio de prevención de daños y responsabilidad: La necesidad de prevenir daños debe impulsar y desarrollar procesos de verificación, validación y control de los sistemas de inteligencia artificial que atribuyan certificados o etiquetados de buenas prácticas, crear códigos de conducta y valores dentro de los modos de actuación de los operadores con enfoque preventivo. Prever la exigencia de responsabilidad para indemnizar por daños y perjuicios, tomando en cuenta la responsabilidad del diseñador o desarrollador de la tecnología y quien la utiliza.

Transparencia y explicabilidad: Los sistemas de inteligencia artificial deberán usados con transparencia y deberán proporcionar información significativa para: Sensibilizar a las personas sobre sus interacciones con estos sistema, explicar a las personas cómo los sistemas de inteligencia artificial se relacionan con sus resultados y proporcionar a las personas que se ven afectadas negativamente por un sistema de inteligencia artificial, la oportunidad de desafiar sus resultados (Comisión Europea, 2020).

La inteligencia artificial se sirve de los algoritmos que son definidos como secuencia(s) finita(s) de operaciones e instrucciones lógicas que hacen posible obtener un resultado a partir de la entrada de información. Se necesita un código fuente o secuencia de instrucciones y un conjunto de datos que puedan ser explicables, que muestren la motivación de la estrategia que conduce a los resultados. Debe garantizarse el derecho a la información a los interesados, suministrada de manera comprensible, accesible cuando sea solicitada, normalizada, completa.

Totalmente ligado al respeto a la autonomía se encuentra el principio de explicabilidad o de trazabilidad, según el cual, los afectados tienen derecho a controlar el uso de nuestros datos y a conocer los algoritmos que los manejan.

Sostenibilidad: El desarrollo de las tecnologías de inteligencia artificial debe realizarse tomando en cuenta los Objetivos de Desarrollo Sostenible (ODS) de las naciones Unidas, la inteligencia artificial puede potenciar el cumplimiento de los objetivos e incidir en el desarrollo o retrasarlo tomando en cuenta las brechas digitales y las diferencias entre los países. La evaluación continua de los efectos humanos, sociales, culturales, económicos y ambientales de las tecnologías de la inteligencia artificial debería llevarse a cabo con pleno conocimiento de las repercusiones de dichas tecnologías en la sostenibilidad como un conjunto de metas en constante evolución en toda una serie de dimensiones (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 2022).

Equidad y no discriminación: Los actores de la inteligencia artificial deberían hacer todo lo razonablemente posible por reducir al mínimo y evitar reforzar o perpetuar aplicaciones y resultados discriminatorios o sesgados a lo largo del ciclo de vida de los sistemas de inteligencia artificial, a fin de garantizar la equidad de dichos sistemas (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 2022). Se debe

procurar eliminar los tratos discriminatorios y la introducción sesgada de datos que incidan en la generación de resultados portadores de inequidades, lo que implica además adoptar un enfoque inclusivo que posibilite la accesibilidad de los posibles beneficiarios en mujeres, afrodescendientes, grupos étnicos, personas en situación de discapacidad, personas en condiciones de vulnerabilidad y de pobreza, personas con diferencias culturales y educativas (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 2022; Randstad, 2019).

Los principios que expresamente se reconocen a la inteligencia artificial son los que estableció la Organización para la Cooperación y el Desarrollo Económico (2019), se firmaron por 44 países y se adopta por el G 20.

1. Centrada en el humano: la inteligencia artificial debería beneficiar a las personas y al planeta impulsando un crecimiento inclusivo, un desarrollo sostenible y bienestar. Por tanto, debe centrarse en el ser humano («human-centred»).
2. Justo: Los sistemas de inteligencia artificial deben diseñarse de manera que respeten el Estado de derecho, los derechos humanos, los valores democráticos y la diversidad, debiendo incluir las salvaguardias adecuadas; por ejemplo, permitiendo la intervención humana cuando sea necesario para asegurar una sociedad justa.
3. Transparente: Debe haber transparencia y divulgación responsable en torno a los sistemas de inteligencia artificial para asegurarse de que la gente entienda cuando se involucran con ellos y puede cuestionar e impugnar los resultados.
4. Seguridad: Los sistemas de inteligencia artificial deben funcionar de manera robusta, segura y sin riesgos a lo largo de su vida, y los riesgos potenciales deben evaluarse y controlarlos continuamente.

5. Responsabilidad: Las organizaciones e individuos que desarrollen, desplieguen u operen sistemas de inteligencia artificial deben responsabilizarse de su correcto funcionamiento de acuerdo con los principios anteriores.

1.5. Posturas teóricas respecto al uso de la inteligencia artificial

El desarrollo ingente de la inteligencia artificial transforma la manera en que se desarrollan y producen los bienes y servicios, su uso tiene un impacto importante en la economía y la sociedad en general.

Son diversos los impactos y usos de la inteligencia artificial que poseen trascendencia jurídica por su incidencia en los derechos fundamentales y por los retos que impone a las instituciones tradicionales del derecho, dígase personalidad, responsabilidad civil, protección de datos, contratación, sistema judicial, entre otras.

A continuación, se examinarán varios de estos usos y su relación con el entorno jurídico.

Robótica e IA, implicaciones jurídicas

Los robots constituyen una tecnología que unida a la inteligencia artificial plantea retos importantes por sus beneficios, pero también por las implicaciones en orden a la responsabilidad y la personalidad jurídica e-personalidad (personalidad electrónica), y los derechos legales de los robots, lo que determina un abordaje especial, sobre todo con el añadido de la complejización de estos con su aplicació. Los robots pueden desempeñar todo tipo de tareas, la definición de robótica proporcionada por la NASA, que concibe a los robots como máquinas capaces de realizar trabajos humanos con diversos grados de autonomía, habiendo algunos que pueden hacerlo por su propia cuenta y otros que

requieren de una persona que les indique previamente que hacer (Medina Rojas, 2025).

Resulta imprescindible mencionar las tres leyes robóticas, enunciadas por Isaac Asimov en 1950: 1. Un robot no debe dañar a un ser humano o, por su inacción, dejar que un ser humano sufra daño. 2. Un robot debe obedecer las órdenes que le son dadas por un ser humano, excepto cuando estas órdenes están en oposición con la primera Ley. 3. Un robot debe proteger su propia existencia, hasta donde esta protección no esté en conflicto con la primera o segunda Ley (Asimov, 2004).

Chávez Valdivia (2022), sintetiza tres componentes para determinar la esencia de un robot:

- Sensores que vigilan el entorno y detectan cambios en él.
- Procesadores o inteligencia artificial que deciden como responder.
- Actuadores que operan sobre el entorno de manera que refleje las decisiones anteriores, provocando algún tipo de cambio en el entorno mundo de un robot.

Son tres las características que Calo (2015) identifica en los robots: corporeidad a diferencia de los softwares tienen existencia material, esta es la característica más debatida por considerarse que el robot es un objeto corpóreo con la capacidad de actuación; impredecibilidad porque el robot decide con autonomía a través de sensores y el intercambio de datos y su capacidad de aprendizaje e impacto social con capacidad de adaptación al entorno (Parlamento Europeo, 2017).

Los robots representan tres posibilidades de acuerdo con la expresión anglosajona de the 3D, *dull* (tedioso), *dirty* (sucio) y *dangerous* (peligroso). estos pueden llevar a cabo tareas muy largas y tediosas, con altas exigencias físicas que pueden conllevar al cansancio, y las largas horas de trabajo que no puede asumir el

ser humano. Los Robots están especializados también en el trabajo en ambientes sucios, altamente contaminados (contaminación nuclear, química, biológica y radiactiva) extremadamente dañinos para el hombre. En cuanto a la peligrosidad reviste importancia que los robots puedan intervenir en ambientes peligrosos ante desastres naturales, lugares inaccesibles para el hombre, incendios y amenazas (Barrios, 2018).

La robótica cobra un matiz superior cuando se integra con la inteligencia artificial, esta última, representa un campo más reciente y dinámico, que se distingue por su capacidad de aprender de los datos, adaptarse a nuevos escenarios y ejecutar de manera más eficiente que el ser humano, por ello han significado una revolución en la creación de máquinas inteligentes.

Los robots se han integrado en la asistencia sanitaria, el transporte, la recopilación de información, la producción industrial, los cuidadores, los asistentes judiciales, la producción industrial y el entretenimiento, en cada vez más complejos sistemas de inteligencia artificial, que se engloban dentro del concepto de robot o sistema robótico, que llegarán a fundir a ambas en una sola categoría.

La Resolución del Parlamento Europeo que establece Normas de derecho civil sobre robótica, expone que entre 2010 y 2014, las ventas de robots aumentaron un 17 % de media cada año, que en 2014 las ventas registraron el mayor incremento anual observado hasta ahora, a saber, un 29 %, y que los principales motores de este crecimiento son los proveedores de componentes de automoción y la industria electrónica y eléctrica; a lo largo del último decenio se han triplicado las solicitudes anuales de patentes en el sector de la tecnología robótica (Parlamento Europeo, 2017).

En torno al reconocimiento de la personalidad jurídica de los robots, se parte de distinguir entre las personas y las cosas, considerando al robot en esta última categoría, por más que revista características antropomórficas.

Darling Kate, analiza si la forma en que las personas parecen reaccionar frente a las máquinas antropomorfas sugiere la necesidad de extender un conjunto limitado de derechos legales a los robots sociales o, al menos prohibiciones contra su abuso, aun cuando no sean reputados como seres vivos o sensibles a un nivel racional (Calo, 2015).

Otra postura considera reconocer una personalidad electrónica a los robots, que se coordina con la exigencia de responsabilidad y la posibilidad de responder en el ámbito civil, penal, comercial, sanitario, derecho tributario, libertad de expresión etc.

Los robots, especialmente los avanzados con capacidades autónomas y de autoaprendizaje, incluso a veces con apariencia humanoide, hace tiempo que son presentados al gran público utilizando una terminología que suele describir actividades o acciones de seres humanos y no de productos o, en su caso, de animales. los robots no son comparables con los animales, pero tampoco con otros productos del intelecto humano, son entes con características parecidas a las del hombre como autonomía y autoaprendizaje. A nivel jurídico, por lo tanto, la personalidad jurídica parece la categoría que mejor responde a las exigencias de atribución de personalidad, sobre todo si la personalidad jurídica se entiende en sentido técnico (Salardi, 2020).

El concepto de la personalidad electrónica debe verse en relación con el de su surgimiento o reconocimiento, para entender a los robots como sujetos de derecho, capaces de responder; a esto se une por ende el concepto de autonomía y capacidad para adoptar decisiones de manera responsable a partir de la inteligencia artificial generativa, que produce la capacidad de autoaprendizaje y por tanto, la capacidad para tomar decisiones no determinadas por instrucciones proporcionadas por su inventor. La capacidad de autoaprendizaje puede ser restringida o regulada por el creador, de ahí que se imita la autonomía del

equipo o tecnología robótica, son estas disquisiciones que cabe plantearse ante las particularidades de los robots.

La Resolución del Parlamento Europeo, Normas de derecho civil sobre robótica, de 16 de febrero de 2017, define la autonomía del robot como la capacidad de tomar decisiones y aplicarlas en el mundo exterior, con independencia de todo control o influencia externos; que esa autonomía es puramente tecnológica y que será mayor cuanto mayor sea el grado de sofisticación con que se haya diseñado el robot para interactuar con su entorno (Parlamento Europeo, 2017).

Respecto a la responsabilidad la postura de esta norma es definitoria al establecer que en el actual marco jurídico, los robots no pueden ser considerados responsables de los actos u omisiones que causan daños a terceros; que las normas vigentes en materia de responsabilidad contemplan los casos en los que es posible atribuir la acción u omisión del robot a un agente humano concreto, como el fabricante, el operador, el propietario o el usuario, y en los que dicho agente podía haber previsto y evitado el comportamiento del robot que ocasionó los daños; que, además, los fabricantes, los operadores, los propietarios o los usuarios podrían ser considerados objetivamente responsables de los actos u omisiones de un robot, inclinándose por la responsabilidad extracontractual objetiva.

Las normas tradicionales no bastan para exigir responsabilidad a un robot capaz de tomar decisiones autónomas, ni siquiera las relativas a los productos defectuosos o por mal funcionamiento al fabricante. En este sentido, la norma insta a tomar decisiones por la UE para establecer regulaciones en cuanto a la capacidad, la autonomía, la responsabilidad, la inexistencia de vida en sentido biológico y llama a establecer medidas de seguridad, control y normalización.

Aplicaciones a la medicina y el derecho sanitario

En el ámbito de la medicina existen aplicaciones de inteligencia artificial trascedentes para la salud humana y la atención a las personas. Se producen chatbots capaces de analizar síntomas y emitir un diagnóstico preliminar, mediante el análisis de determinados datos se puede establecer la propensión a desarrollar enfermedades, encontrar patrones en los pacientes o reducir el riesgo de procedimientos sanitarios (Repsol, 2025).

La inteligencia artificial incide ampliamente en la llamada salud digital, así la tecnología digital integrada por dispositivos portátiles, telemedicina, aplicaciones médicas móviles y softwares, ha producido una revolución en la atención médica. Las herramientas de la salud digital pueden mejorar la capacidad de realizar diagnósticos y tratar enfermedades con precisión, obteniendo mayores resultados médicos y posibilitando que los pacientes tengan un mayor control sobre su salud (Mitelman, 2025). Mediante la inteligencia artificial se pueden analizar grandes volúmenes de datos y mejorar los resultados de los pacientes, así como reducir el trabajo de los médicos.

Una noticia interesante en el campo de la salud es la inauguración del hospital de “Agent Hospital”, es la simulación de un Hospital en China, por la Universidad Tsinghua. En el hospital simulado, hay dos clases de participantes: por un lado, los médicos y las enfermeras y por el otro, los residentes que simulan ser pacientes. La idea es simular la atención de cada paciente desde el ingreso, diagnóstico, tratamiento y monitoreo y que el agente médico vaya evolucionado a medida que interactúe con los pacientes. El hospital tiene una biblioteca con los casos exitosos y una base de datos con los erróneos. Se experimentó en alrededor de 10.000 casos, alcanzando una precisión del 88% en revisación, 95,6% en diagnóstico y 77,6% en tratamiento (Mitelman, 2024).

La inteligencia artificial, se perfila, sobre todo, como una herramienta capaz de aprender y analizar con rapidez enormes cantidades de información de los historiales de pacientes, de las

pruebas de imagen y de los avances científicos para ayudar a los profesionales a ofrecer mejores diagnósticos y tratamientos. La inteligencia artificial es un asociado que les liberará también de algunas tareas monótonas, como el análisis de las imágenes médicas, estos sistemas permiten evaluar a los pacientes de manera más eficiente y menos costosa (Carcar Benito, 2019).

Otras aplicaciones, son los robots cirujanos, empleados en el presente en cirugías como el implante cloquear, con un nivel de acierto, importante. Sin embargo, el uso de los robots plantea la problemática de la responsabilidad que deberá quedar atribuida al desarrollador autor del sistema, al programador o la institución hospitalaria que utiliza el producto y en donde se practica la operación quirúrgica.

Una de las principales dificultades de la inteligencia artificial aplicada a la medicina es su uso inapropiado o con una intención marcadamente dañina, dígase por la programación de algoritmos con sesgos, la introducción de datos, el razonamiento lógico y la falta de transparencia de los algoritmos, a ello se agregan las particularidades del lenguaje médico, que a veces posee ambigüedad, vaguedad y significados meramente para el contexto médico y pueden ser tergiversados por la tecnología.

Un aspecto a tomar en cuenta es la disminución del flujo del empleo por la automatización de tareas, que conlleva a la optimización de procesos y la reducción de empleos menos complejos y que pueden ser asumidos por aplicaciones de inteligencia artificial, lo que tiene un impacto importante sobre los derechos laborales y la disminución de la fuerza laboral, así como la afectación al derecho al trabajo.

Son varias las posturas teóricas relativas a la relación entre derecho sanitario e inteligencia artificial relacionadas por Carcar Benito (2019):

a) casos normativamente difíciles de conflictos (ético-jurídicos) por la indeterminación semántica y vaguedad conceptual del propio ámbito jurídico-sanitario; b) casos epistémicamente y metodológicamente difíciles en los que el hallazgo de la respuesta precisa un notable esfuerzo profesional; c) casos pragmáticamente difíciles por causas ajenas al derecho sanitario, por razones de trascendencia y conflicto político y social; d) casos tácticamente difíciles o que plantean dudas sobre la calificación jurídica de los hechos, muy corriente en el derecho a la salud; e) casos moralmente difíciles o de justicia distributiva, en los que la respuesta jurídicamente correcta comporta resultados injustos.

Otro aspecto sustancial a tomar en cuenta es el uso de los datos personales, la inteligencia artificial se programa a partir de la introducción de datos, en el ámbito de la salud se recolectan y almacenan datos de los pacientes y las historias clínicas, relativos a la salud y estado de las personas, en este proceso prima el principio de confidencialidad, privacidad y además el secreto profesional, por lo cual trabajar con este tipo de datos médicos demanda que se adopten todas las medidas con rigurosidad, para evitar las filtraciones y manipulación que sería nefasta para quien los ostenta, de otro lado se debe velar por la calidad de los datos.

Un aspecto sustancial a tratar sobre la relación entre inteligencia artificial y medicina es la responsabilidad al respecto son 4 los aspectos que identifican Price et al. (2024).

En primer lugar, el ámbito de la responsabilidad extracontractual por inteligencia artificial sigue evolucionando, la responsabilidad por inteligencia artificial en el ámbito sanitario aún no se ha abordado directamente en los tribunales, principalmente porque la tecnología en sí es muy reciente y aún se encuentra en fase de implementación.

En segundo lugar, la causalidad suele ser difícil en contextos de responsabilidad civil extracontractual con inteligencia artificial. Demostrar la causa de una lesión es difícil en el ámbito médico, donde los resultados suelen ser probabilísticos en lugar de deterministas.

En tercer lugar, resulta inviable captar con precisión las diferencias internacionales en materia de responsabilidad.

En cuarto lugar, desde una perspectiva sistémica, la responsabilidad individual de los profesionales sanitarios es bien compleja. En el ámbito de la inteligencia artificial médica interactúan numerosos actores, incluyendo actores que podrían asumir responsabilidad y reguladores que podrían configurarla.

Sobre los robots asistenciales la Unión Europea valora su importancia y los riesgos que poseen, al subrayar que, con el tiempo, la investigación y el desarrollo de robots de asistencia geriátrica han pasado a ser más habituales y menos costosos, ofreciendo productos con mayor funcionalidad y mejor aceptación entre los consumidores; pone de relieve la amplia gama de usos de estas tecnologías para ejercer funciones de prevención, asistencia, seguimiento, estimulación y compañía de las personas de edad avanzada o que padecen demencia, trastornos cognitivos y pérdida de memoria; pero señala que el contacto humano es uno de los aspectos fundamentales de la atención a las personas; considera que sustituir el factor humano por robots podría deshumanizar la prestación de cuidados, pero, por otra parte, reconoce que los robots podrían realizar las tareas automatizadas de quienes prestan cuidados, aumentado la atención prestada por seres humanos y haciendo más selectivo el proceso de rehabilitación (Parlamento Europeo, 2017).

La introducción de la inteligencia artificial en el campo de la salud se justifica desde el bienestar humano, tal como ocurre con el desarrollo tecnológico, para garantizar este propósito es imprescindible establecer un marco regulatorio que prevea sus

impactos negativos y que contenga regulaciones éticas para su desarrollo y uso. Definir esta relación entre IA, derecho y medicina, y los aspectos precisos del marco regulatorio de estos tres campos, resulta una compleja tarea y una de las vertientes de uso que por su incidencia en el bienestar humano demanda una mirada especial.

Sistemas de transportación

En el ámbito de la transportación los sistemas de inteligencia artificial han cobrado presencia importante con usos en los vehículos autónomos, en los sistemas de vialidad, la transportación de pasajeros y la disciplina vial. Tales aplicaciones permiten disminuir los accidentes de carretera, predecir posibles problemas de accidentalidad, anticiparse a las necesidades de mantenimiento de los vehículos, planificar rutas de transportación y posibilitar la trasmisión de datos entre los dispositivos conectados.

En este entramado de servicios se comenzará el análisis por los vehículos autónomos, que ha supuesto novedosamente una revolución en el ámbito de la movilidady los impactos de esta tecnología desde la perspectiva jurídica.

El vehículo autónomo es un vehículo sin conductor haciendo referencia esencialmente a la tecnología empleada. Debemos tener en cuenta que un vehículo autónomo está altamente robotizado. Como tal, funciona a través de un lenguaje-máquina. En esencia, no deja de ser un ordenador, programado mediante líneas de código para cumplir un cometido específico, que es circular desde el punto A al punto B sin chocar con ningún obstáculo (Combarros Merino, 2023).

El transporte autónomo abarca todas las formas del transporte por carretera, ferroviario, por vías navegables y aéreas, pilotadas a distancia, automatizadas, conectadas y autónomas, están incluidos en estas formas, los vehículos, los trenes, los buques, los transbordadores, las aeronaves y los drones, así como todas

las futuras formas que resulten del desarrollo y la innovación en este sector.

La convergencia de la inteligencia artificial y la seguridad en medios de transporte tiene una importante expresión en los vehículos autónomos. Debido a su capacidad para reducir la accidentalidad asociada al error humano, así como para permitir una navegación sin dificultades, estos vehículos representan una innovación significativa en el ámbito de la movilidad.

Esto ha introducido un nuevo desafío para los vehículos autónomos, que ha sido denominado como seguridad de la funcionalidad prevista o SOTIF (safety of the intended functionality) el cual se refiere a los problemas que surgen de las ineficiencias funcionales con respecto a lo previsto en las funciones de un vehículo autónomo o su implementación, de acuerdo a la estandarización de la ISO 21448, que se centra en la seguridad de estos vehículos, de los riesgos por deficiencias o el mal uso previsible por parte de los humanos (Silva y Cuevas, 2024).

El núcleo de esta tecnología reside en sistemas avanzados de reconocimiento de imágenes, respaldados por el aprendizaje profundo dentro del campo de las redes neuronales, que les permiten comprender y moverse por su entorno.

El sitio web HP (HP Development Company, L.P., 2024), establece como funcionan estos vehículos, un coche autónomo realiza su operación cuando:

1. **Define una ruta:** el usuario establece un destino y el software del coche calcula el curso a seguir.
2. **Capta su entorno:** mediante sensores, como el LIDAR y cámaras, el automóvil crea un mapa 3D de todo a su alrededor.
3. **Se orienta en el espacio:** el automóvil utiliza sensores de movimiento para determinar su ubicación con respecto al mapa.

4. **Evalúa peligros cercanos:** los sistemas de radar detectan la distancia a obstáculos adelante y atrás.
5. **Procesa la información:** el software de inteligencia artificial integrado unifica los datos, incluyendo aquellos obtenidos de mapas virtuales (como Google Street View) y cámaras internas.
6. **Toma decisiones:** la inteligencia artificial produce simulaciones de procesos de percepción y toma de decisiones, controlando la dirección y los frenos.
7. **Consulta mapas avanzados:** busca datos anticipados de mapas sobre señales de tráfico y puntos de interés.

Existen niveles de autonomía para considerar que un vehículo autónomo posee esta cualidad según el grado de autonomía del vehículo, o, dicho de otro modo, dependiendo del grado de supervisión requerido por el conductor:

La autonomía vehicular se define a través de seis niveles diferentes que la Sociedad de Ingenieros Automotrices (SAE) y la Dirección general de Tráfico de España:

- **Nivel 0**: sin automatización. Tú, como conductor humano, posees el control total del vehículo.
- **Nivel 1**: asistencia al conductor. Algunos sistemas como el frenado o la dirección son automáticos, aunque tu atenta supervisión resulta indispensable.
- **Nivel 2**: automatización parcial. La mayoría de los aspectos de la conducción los gestiona el vehículo, donde se requiere tu intervención en situaciones específicas.
- **Nivel 3**: automatización condicional. El coche navega de forma autónoma en la mayoría de las circunstancias, pero tú debes estar dispuesto a tomar el control.

- **Nivel 4**: autonomía elevada. Los coches requieren mínima entrada humana al circular en múltiples escenarios, aunque en situaciones atípicas se puede necesitar tu intervención.
- **Nivel 5**: autonomía completa. El automóvil funciona de manera autónoma, sin necesitar intervenciones externas, lo que permite viajar en todas las condiciones sin la necesidad de que tú conduzcas.

Actualmente, los vehículos autónomos se encuentran en el nivel 4 en pruebas por compañías como Waymo y Cruise, por ello, aun la autonomía es relativa, porque el vehículo requiere en todo momento la presencia del conductor, si bien puede desactivar el sistema de conducción autónoma a voluntad del conductor del vehículo, en el caso de los vehículos de niveles 2 y 3, lo que implicaría la exigencia de responsabilidad civil por negligencia en caso de accidentalidad y la producción de daños y perjuicios, los niveles 4 y 5, admitirían la conducción del vehículo con desconexión total, incrementando los riesgos de accidentalidad (Combarros Merino, 2023).

Ante la progresión de la tecnología de los vehículos autónomos, el Reglamento del Parlamento europeo, establece el Código de conducta ética para los ingenieros en robótica que, en la fase de diseño y desarrollo de los vehículos autónomos, deben respetar los siguientes principios:

- **Beneficencia**: Los robots deben actuar en beneficio del hombre.
- **Principio de no perjuicio o maleficencia**: En virtud del cual los robots no deben perjudicar a las personas (primero, no hacer daño).
- **Autonomía**: La capacidad de tomar una decisión con conocimiento de causa e independiente sobre los términos de interacción con los robots.

- **Justicia**: La distribución justa de los beneficios asociados a la robótica.

Un sistema de vehículos autónomos trae consigo innumerables ventajas, pero también innegables riesgos que se acompañan de la desconfianza pública.

Un análisis de los beneficios es que puede reducirse la accidentalidad y los siniestros, y podrían implicar una mayor seguridad en la vía pública, mediante la inteligencia artificial, estos vehículos reaccionarían con precisión a múltiples escenarios de tránsito, lo que reduciría las colisiones.

La Administración Nacional de Seguridad del Tráfico en las Carreteras de Estados Unidos estima que hasta un 94 % de los siniestros viales corresponden a errores humanos. Con la conducción autónoma, se eliminan factores humanos de riesgo como distracciones o infracciones a las reglas de tránsito (HP Development Company, L.P., 2024).

Genera beneficios en el sistema vial, que contaría con una red optimizada de coches autónomos disminuirían la congestión vial y se coordinarían los movimientos, habría flujos constantes, mejorando la eficiencia de la circulación y la prevención de accidentes y además implicaría una reducción de costos.

Darían cauce a un enfoque inclusivo al representar el acceso para personas con restricciones, como las personas adultas mayores, las personas en situación de discapacidad, se eliminarán las barreras de movilización incluso cuando la persona no posea autonomía para su movilidad por determinadas limitaciones físicas (Parlamento Europeo, 2017).

Dentro de los obstáculos más significativos se encuentra la determinación de la responsabilidad en caso de accidente provocado por un vehículo autónomo, se debe establecer quien respondería: el propietario, el proveedor del servicio o el fabricante, siendo cuestionable un sistema de autonomía infalible.

la industria debe garantizar un vehículo autónomo fiable, seguro, creíble.

Los desarrolladores en la industria automotriz manejan este problema con cautela, y se dedican a monitorear atentamente la retroalimentación de las pruebas para perfeccionar continuamente los sistemas. La industria debe comprobar no solo la seguridad de la tecnología, sino también su confiabilidad en comparación con las decisiones en situaciones críticas que puede hacer un ser humano. Cada prueba y dato aporta a construir esta credibilidad tan necesaria (HP Development Company, L.P., 2024).

En otro orden, se presentan dilemas éticos con relación a los vehículos autónomos y la inteligencia artificial, estos autos deben adoptar decisiones en cuestión de segundos, poniendo en riesgo la vida humana frente a los daños materiales y el interés de conservar el vehículo. En otro orden, a pesar de los avances impresionantes, la inteligencia artificial en automóviles aún no iguala la complejidad del razonamiento humano. Los sistemas actuales sobresalen en tareas específicas, pero afrontan limitaciones al interpretar contextos ambiguos y al ejecutar juicios basados en valores éticos (HP Development Company, L.P., 2024).

Los automóviles dependen de los sistemas de radares, y las señales de internet, no existe cobertura suficiente en todos los sitios lo que podría inutilizar los dispositivos, en zonas de silencio o en condiciones climáticas adversas.

Combarros Merino (2023) expone otro riesgo como la observancia de la diligencia al conducir uno de estos vehículos se complica en gran medida porque resulta posible que por parte del fabricante del vehículo se incurra en un defecto de información sobre las capacidades técnicas del vehículo. Esto significa que el conductor puede llegar a tener un exceso de confianza en las capacidades del vehículo, de modo que este defecto de información imputable al fabricante es un factor que contribuye en cierta medida a que

el conductor relaje su nivel de atención y por consiguiente incurra en una conducta negligente por falta de la diligencia debida.

En las Normas de derecho civil y robótica del Parlamento europeo se considera que la transición a los vehículos autónomos repercutirá en los siguientes aspectos: la responsabilidad civil (responsabilidad y seguros), la seguridad vial, todas las cuestiones relativas al medio ambiente (por ejemplo, eficiencia energética, uso de tecnologías renovables y fuentes de energía), las cuestiones relativas a los datos (por ejemplo, acceso a los datos, protección de los datos personales y la intimidad, intercambio de datos), las cuestiones relativas a la infraestructura TIC (por ejemplo, alta densidad de comunicaciones eficientes y fiables) y el empleo (por ejemplo, creación y pérdida de puestos de trabajo, formación de los conductores de vehículos pesados para el uso de vehículos automatizados); subraya que se necesitarán inversiones considerables en las infraestructuras viarias, energéticas y de TIC; pide a la Comisión que examine los aspectos mencionados en sus trabajos sobre los vehículos autónomos (Parlamento Europeo, 2017).

Sistema judicial

El sistema judicial se ha visto impactado por la inteligencia artificial, en tanto, su utilización permite automatizar procesos, que se realizan por los operadores legales, puede mejorar la eficiencia y aminorar los errores. El uso de la inteligencia artificial en este contexto abarca una amplia gama de posibilidades desde la aplicación en la gestión de información estandarizada como los expedientes judiciales, hasta impactar en la adopción de decisiones judiciales y la resolución alternativa de conflictos, permite descongestionar el sistema, disminuir la carga laboral de los jueces y hacer más eficiente el sistema, al manejar grandes volúmenes de información y contribuir a la elaboración y fundamentación de resoluciones judiciales (González Moreno, 2023).

Sin embargo, ello plantea diversas observaciones en orden a la transparencia de los algoritmos, prever limitar los sesgos en los datos que generen inequidades y discriminación, estandarizar la justicia en manos de la tecnología hace correr el riesgo de no considerar determinados valores y sentimientos humanos que influyen en la toma de decisiones, que ponga en juego la adecuada impartición de justicia.

La inteligencia artificial aplicada al ámbito judicial tiene mucha relación con los derechos fundamentales como el derecho a la libertad, la igualdad, a la vida, la propiedad, el habeas corpus, el habeas data, la dignidad y la privacidad.

Sobre la adopción de decisiones un sector de la doctrina se inclina por no introducir en el análisis de los datos información como la edad, el sexo, la etnia, el color de la piel, el origen social entre otras porque implicaría afectación a los derechos de estas personas y exclusiones y otros se inclinan por su valoración por la incidencia que tendría en el funcionamiento óptimo de la herramienta (Sarmiento Arango, 2020).

Los sistemas de expertos jurídicos basados en inteligencia artificial son programas que simulan el razonamiento humanos y la capacidad de solucionar problemas jurídicos, estos son capaces de analizar amplios volúmenes de datos, pero se puede objetar a estos que no pueden reemplazar el razonamiento humano, la lógica jurídica, sino que lo complementan al carecer de la capacidad de interpretar contextos, valores éticos, puede valorar tendencias, coadyuvar a la interpretación pero los principios éticos es tarea del ser humano (Álvarez Villajuana, 2025).

En el ámbito europeo se distingue entre las decisiones judiciales guiadas por la equidad (justicia equitativa), y las decisiones automatizadas producidas por el uso de la inteligencia artificial (justicia codificada).

Puede observarse que el núcleo principal de diferenciación entre ambos modelos es la discusión tradicional sobre la naturaleza y propósitos del Derecho, en especial lo referido a la deseabilidad o no de la discreción jurídica. Es decir, mientras que la unificación de doctrina en la justicia equitativa requiere siempre discusión entre jueces y participantes en el proceso, en definitiva, ponderación, la justicia codificada, en cambio, unifica automáticamente siempre, lo que puede tener consecuencias graves una vez que en el Estado de Derecho las decisiones judiciales han de ser individuales atendiendo a todas las particularidades del caso (Galindo, 2020).

Las consecuencias adversas de las decisiones judiciales codificadas o automatizadas, están centradas en las denominadas: "incomprensión", "datificación", "desilusión" y "alienación". Cada una se fija en lo siguiente. 1) Las decisiones judiciales automatizadas producen "incomprensión", el mayor riesgo del uso de la justicia codificada "funcionar de maneras que son difíciles o imposibles de comprender para los seres humanos. La naturaleza de esta preocupación puede variar dependiendo del tipo de inteligencia artificial utilizada" (Galindo, 2020); 2) Las decisiones judiciales automatizadas también producen "desilusión". Ello es debido a que el… "desarrollo y uso de la Administración de Justicia realizada por medio de la inteligencia artificial están, ya, provocando una reconsideración escéptica de las prácticas existentes con la Administración de Justicia tradicional; 3) Producen "alienación".

"A medida que los juicios de inteligencia artificial desempeñen un papel más importante en el sistema jurídico, la participación humana cambiará y, en algunos aspectos, disminuirá. Esos acontecimientos plantean la posibilidad de alienación, o la tendencia de que algunas o todas las personas dejen de participar en el sistema jurídico e incluso pierdan interés en sus operaciones" (Galindo, 2020).

IA en el ámbito de la educación

La Inteligencia Artificial (IA) proporciona el potencial necesario para abordar algunos de los desafíos mayores de la educación actual, innovar las prácticas de enseñanza y aprendizaje y acelerar el progreso para la consecución del ODS 4. Sin embargo, los rápidos desarrollos tecnológicos conllevan inevitablemente múltiples riesgos y desafíos, que hasta ahora han superado los debates políticos y los marcos regulatorios (Galindo, 2020), el riesgo de datificacion" consiste en que se define como la capacidad de producir datos que puedan ser leídos por un ordenador. es que los datos tomados en consideración se generan por la maquina y no por textos legales

La inteligencia artificial se distingue de otras tecnologías digitales por su potencial para remodelar profundamente las sociedades, las economías y los sistemas educativos. A diferencia de las tecnologías de la información y la comunicación (TIC) convencionales; su uso plantea desafíos éticos y sociales únicos, como son las cuestiones relativas a la equidad, la transparencia, la privacidad y la responsabilidad. Además, posee capacidad única para imitar el comportamiento humano afecta directamente a la actividad humana. Estos retos requieren competencias específicas que van más allá de la alfabetización digital tradicional (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 2025).

El futuro de la Educación Superior está siendo remodelado por la inteligencia artificial, y las decisiones que se tomen hoy determinarán si esta transformación conduce a una mayor inclusión y oportunidades de aprendizaje para todos (Molina y Medina, 2025).

1.6. Marco internacional de la inteligencia artificial

La presencia creciente de la inteligencia artificial a nivel global ha llamado la atención de la comunidad y organismos internacionales

para la propuesta de iniciativas políticas y legales de regulación que tienen por objetivo garantizar un uso adecuado de la misma y por consiguiente, minimizar los riegos que supone el uso de estas tecnologías y maximizar sus beneficios en aras de una sociedad más justa e inclusiva. En este sentido, pueden listarse las siguientes normativas internacionales:

- "Recomendación sobre la ética de la inteligencia artificial aprobada por la Conferencia General de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) el 23 de noviembre del 2021.
- Resolución A/RES/78/265 "Aprovechar las oportunidades de sistemas seguros y fiables de inteligencia artificial para el desarrollo sostenible" aprobada el 21 de marzo del 2024 por la Asamblea General de las Naciones Unidas.
- Resolución A/RES/78/311 "Aumentar la cooperación internacional para la creación de capacidad en materia de inteligencia artificial" aprobada por la Asamblea General el 1 de julio de 2024 por la Asamblea General de las Naciones Unidas.

Unido a este naciente marco regulatorio se han aprobado por la Organización Internacional de Normalización (ISO) y la Comisión Electrotécnica Internacional (IEC) un conjunto de normas que establecen estándares internacionales en pos de alcanzar niveles de homogeneidad en relación con la gestión, prestación de servicios y desarrollo de productos en la industria asociada a la inteligencia artificial. Estas son:

- ISO/IEC 22989:2022 Tecnología de la información — Inteligencia artificial — Conceptos y terminología de inteligencia artificial.
- ISO/IEC 23894:2023 Tecnología de la información — Inteligencia artificial — Orientación sobre gestión de riesgos.

- ISO/IEC 38507:2023 Tecnologías de la información — Gobernanza de TI — Implicaciones de gobernanza del uso de inteligencia artificial por parte de las organizaciones.
- ISO/IEC 5338:2023 Tecnología de la información — Inteligencia artificial — Procesos del ciclo de vida del sistema de inteligencia artificial.
- ISO/IEC 42001:2025 Tecnología de la información — Inteligencia artificial —Sistema de gestión.
- ISO/IEC 42005:2025 Tecnología de la información — Inteligencia artificial (IA) — Evaluación del impacto del sistema de inteligencia artificial.
- ISO/IEC 42006:2025 Tecnología de la información — Inteligencia artificial — Requisitos para los organismos que proporcionan auditoría y certificación de sistemas de gestión de inteligencia artificial.

En especial las normas jurídicas internacionales establecen:

Recomendación sobre la ética de la inteligencia artificial aprobada por la Conferencia General de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) el 23 de noviembre del 2021.

La *Recommendation on the Ethics of Artificial Intelligence* considerada la primera norma mundial sobre la ética de la inteligencia artificial y cuyo marco fue adoptado por los 193 Estados miembros de la organización. Según la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura a través de su sitio el eje fundamental de dicha recomendación versa alrededor de los derechos humanos y la dignidad. Basándose en principios como la equidad y la transparencia y que los sistemas de inteligencia artificial se encuentren bajo la supervisión humana.

La Recomendación pretende servir de base para que los sistemas de inteligencia artificial queden al servicio de la sociedad en general, del medio ambiente y los ecosistemas, así como para prevenir daños. Ser un instrumento normativo de aceptación mundial, para lo cual articula una serie de principios y valores. En este sentido, resulta importante destacar que lo que hace verdaderamente aplicable a esta Recomendación es la posibilidad de traducir esos valores y principios en acciones respecto a la gobernanza de datos, el medio ambiente y los ecosistemas, el género, la educación, la investigación, la salud y el bienestar social, entre otros.

Los objetivos de la Recomendación estriban en lograr un marco universal de valores éticos y acciones para orientar a los Estados en la formulación de sus políticas y leyes relativas a la inteligencia artificial, de conformidad con el Derecho Internacional. De igual manera, orientar las acciones de las personas, las comunidades, las instituciones y el sector privado con el propósito de asegurar en todas las etapas del ciclo de vida de los sistemas de inteligencia artificial la incorporación de la ética (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 2021). Además, promover y respetar los derechos humanos y las libertades fundamentales, preservar el medio ambiente y respetar la diversidad cultural. También, garantizar el acceso equitativo a los avances y conocimientos que se generen en materia de inteligencia artificial.

Para la consecución de estos fines y objetivos, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura fija una serie de valores y principios éticos que están en correspondencia con el Derecho Internacional y los Objetivos de Desarrollo Sostenible (ODS). Estos valores y principios libran una importante labor como guías que motivan las medidas de política y las normas jurídicas. Los mismos deben ser respetados y promovidos mediante modificaciones de las

leyes, los reglamentos y las directrices empresariales existentes y la elaboración de nuevas normativas, cuando sea necesario y conveniente (Reyes Vásquez, 2023). Estos valores son: el respeto, protección y promoción de los derechos humanos, las libertades fundamentales y la dignidad humana; la prosperidad del medio ambiente y los ecosistemas; por otra parte garantizar la diversidad y la inclusión; y no por ser el último deja de tener trascendental importancia; vivir en sociedades pacíficas, justas e interconectadas.

En el marco de la Recomendación, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, incluye diez principios éticos, los que revelan de manera más concreta y exacta el contenido de los valores. De igual manera, hacen más sencilla su materialización en acciones concretas encaminadas a trazar políticas y normativas respecto a la ética de los sistemas de inteligencia artificial.

Dentro de estos principios se encuentran: proporcionalidad e inocuidad, principio que subraya la importancia de garantizar que la inteligencia artificial se emplee de manera que se ajuste a la finalidad para la cual fue diseñada, para evitar excesos o peligros innecesarios; por lo cual debe adaptarse bien a las circunstancias concretas que se presentan basándose en principios científicos fiables. Por su parte, el principio de seguridad y protección plantea abordar y mitigar los daños, riesgos de seguridad y las vulnerabilidades de los sistemas de inteligencia artificial; ello implica tomar medidas proactivas, como el establecimiento de marcos sostenibles para el acceso a los datos, con el objetivo de prevenir y eliminar tales riesgos.

El principio de equidad y no discriminación diseña una perspectiva integradora para garantizar la disponibilidad y accesibilidad equitativas (Morandín-Ahuerma, 2023) de los beneficios de la inteligencia artificial para toda la sociedad, tomando en cuenta las necesidades de los diferentes sectores sociales; con el fin de

promover un orden mundial más equitativo en lo que se refiere a la información, educación, cultura, investigación, las condiciones socioeconómicas e incluso la estabilidad política (Boddington, 2020).

El principio de sostenibilidad aborda la posibilidad de que los sistemas de inteligencia artificial se diseñen de manera que consideren y tengan en cuenta minimizar el impacto negativo en el medio ambiente, en especial el control de emisiones y uso racional del agua y la energía.

Otro de los principios es el derecho a la intimidad y protección de datos, en este sentido, la intimidad se corresponde con el hecho de que los datos personales no sean divulgados, en todo caso la privacidad debe ser promovida y respetada como derecho humano fundamental que es. Así, cualquier tratamiento de datos personales implica el consentimiento informado; que constituye una manifestación de voluntad libre, específica, informada e inequívoca por la que el usuario de sistemas de inteligencia artificial acepta, ya sea mediante una declaración o una clara acción afirmativa el tratamiento de los datos personales que le conciernen y entre sus características esenciales se encuentra la individualidad ya que solo es concedido por la persona a cuyos datos se acceden, y la voluntad libre, o sea, dicho consentimiento no puede ser bajo coacción o ningún tipo de violencia, de manera que la persona afectada emita una autorización para el manejo lícito de su información, de lo contrario se trataría de un abuso (Véliz, 2022). El principio de transparencia y explicabilidad por su parte, fomenta y estimula la confianza en los sistemas de inteligencia artificial a partir de que los usuarios sean capaces de comprender los procesos y etapas de dichos sistemas inteligentes; lo cual determina la capacidad de esos sistemas de hacer comprensibles y explicables sus resultados.

Entiéndase por transparencia la propiedad de un sistema de inteligencia artificial que aporta información sobre los

procesos internos propios y se pone a disposición de las partes interesadas. Mientras, que la explicabilidad viene a expresar factores importantes que influyen en los resultados del sistema de inteligencia artificial de una manera comprensible para los humanos. El grado de exigencia de ambos depende del impacto del algoritmo en la decisión, el resultado y el ciudadano, el grado de autonomía en la toma de decisiones (o sea, hasta qué punto se garantiza la participación humana), y del tipo y la complejidad del algoritmo (Cotino Hueso 2019).

La supervisión y decisión humanas como principio ético significa que, aunque los humanos se apoyen o incluso deleguen en los sistemas de inteligencia artificial la realización de tareas puntuales, ese sistema no podrá, en ningún caso asumir la responsabilidad, que en tales circunstancias recaerá o bien sobre el operador, o bien sobre aquel que haya dado la orden al sistema; sobre todo si se trata de decisiones que colocan en riesgo la vida o la salud de una persona (AlgorithmWatch, 2025).

Vinculado al anterior se incluye el principio de responsabilidad y rendición de cuentas, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura enfatiza que debe estar en conformidad con la legislación internacional y nacional de cada país, en particular en materia de derechos humanos (Morandín-Ahuerma, 2023).

En cuanto al principio de sensibilización y educación, debe fomentarse la alfabetización y capacitación dirigida conjuntamente por los gobiernos, las organizaciones intergubernamentales, la sociedad civil, las universidades, los medios de comunicación, los dirigentes comunitarios y el sector privado.

Por último, se encuentra el principio de gobernanza y colaboración adaptativas y de múltiples partes interesadas, y su importancia radica en que la colaboración entre las partes involucradas es necesaria para facilitar la adopción de normas abiertas e interoperables (Taddeo y Floridi, 2021).

Una de las fortalezas de esta Recomendación es ofrecer un ámbito común de acción política para la puesta en práctica de los valores y principios éticos que en ella se refrendan. Estos ámbitos de actuación incluyen la evaluación del impacto ético, gobernanza y administración éticas, política de datos, desarrollo y cooperación internacional, medio ambiente y ecosistemas, género, cultura, educación e investigación, comunicación e información, economía y trabajo, y salud y bienestar social. En este sentido, la principal acción que se propone consiste en que sus Estados miembros deberán llevar a cabo medidas eficaces, como el establecimiento de marcos normativos a los cuales se adhieran desde las empresas del sector privado hasta las instituciones universitarias, científicas e investigativas, así como la sociedad civil.

Por demás, estos sectores e instituciones deberán volcar su labor hacia la confección de herramientas que permitan y garanticen la evaluación del impacto de los sistemas de inteligencia artificial ya sea en los derechos humanos, el estado de derecho, la democracia, e incluso la evaluación del impacto ético, además de la elaboración de instrumentos de diligencia debida. Estos procesos de elaboración de mecanismos regulatorios y de herramientas normativas, necesariamente deben incluir a las múltiples partes interesadas y tener en cuenta las circunstancias concretas, así como las prioridades de cada Estado.

De manera general, la Recomendación de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, propone que se aborde siempre la inteligencia artificial desde un enfoque ético, que ha de ser tomado como un marco global, multicultural y evolutivo de valores y de principios que se encarguen de orientar a las sociedades cuando tengan que enfrentarse a todos los efectos de tecnologías. Proporcionando consecuentemente, orientación ética a todos sus actores, incluidos los sectores público y privado.

Si bien la Recomendación de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, constituye el primer acuerdo intergubernamental de su tipo, y reviste vital importancia para regulaciones futuras por parte de sectores y Estados, no escapa de la esfera *soft law* puesto que se trata de una cooperación que se basa en un instrumento que no es legalmente vinculante y que, sobre todo, otorga un papel preponderante a la buena fe de los Estados para que lleven a sus normativas nacionales acuerdos relativos a la ética de la inteligencia artificial.

Resolución A/RES/78/265 "Aprovechar las oportunidades de sistemas seguros y fiables de inteligencia artificial para el desarrollo sostenible" aprobada el 21 de marzo del 2024 por la Asamblea General de las Naciones Unidas.

El 15 de septiembre del 2021 la Alta Comisionada de las Naciones Unidas para los Derechos Humanos, Michelle Bachelet hacía un llamado ante la comunidad internacional en pos de la necesidad urgente de regular el uso de sistemas de inteligencia artificial que amenazan gravemente a los derechos humanos. En ese orden, Bachelet señalaba que "mientras mayor sea el riesgo para los derechos humanos, más estrictos deben ser los requisitos legales para el uso de la tecnología de IA" (Organización de las Naciones Unidas, 2021).

Asimismo, el informe recalca la importancia de establecer un sistema de responsabilidad en cuanto a la recopilación, almacenamiento y uso de datos como uno de los cometidos más urgentes en materia de derechos humanos. La Alta Comisionada terminó su informe exhortando a establecer límites en torno a la utilización de la inteligencia artificial alrededor de los derechos humanos agregando que:

> La capacidad de la IA para servir a la población es innegable, pero también lo es su capacidad de contribuir a violaciones de

> derechos humanos en gran escala, de manera casi indetectable. Es necesario adoptar urgentemente medidas para imponer límites basados en los derechos humanos a la utilización de la IA por el bien de todos (Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos, 2021).

Consecuente a ello, las Naciones Unidas no ha cesado en su labor de promover un uso "seguro y fiable" de las tecnologías de la información y de las comunicaciones, y de la inteligencia artificial. Así, el 21 de marzo de 2024 la Asamblea General aprobaba por unanimidad, la primera resolución respecto a la inteligencia artificial para promover su desarrollo de forma segura, responsable y equitativa con el fin de evitar las desigualdades entre países y su posible uso malintencionado. El proyecto fue liderado por Estados Unidos y copatrocinado por más de 120 Estados miembros. Con el objetivo de promover, en lugar de obstaculizar, la transformación digital y el acceso equitativo a los beneficios de estos sistemas.

La resolución A/RES/78/265 "Aprovechar las oportunidades de sistemas seguros y fiables de inteligencia artificial para el desarrollo sostenible" define como sistemas "seguros y fiables" de inteligencia artificial a aquellos que no se encuentran en el ámbito militar. Y que entre otras características se encuentran:

> Centrados en las personas, son fiables, se pueden explicar, son éticos e inclusivos, respetan plenamente la promoción y la protección de los derechos humanos y el derecho internacional, mantienen la privacidad, están orientados al desarrollo sostenible y son responsables (Organización de las Naciones Unidas, 2024).

La fiabilidad es la probabilidad, durante un tiempo determinado, de que el sistema se ejecutará correctamente y brindará los servicios que el usuario espera. Por otra parte, la seguridad es un atributo del sistema que refleja la capacidad de este para

protegerse a sí mismo de cualquier ataque externo, que pueden ser accidentales o deliberados.

La supramentada resolución hace un llamado a la gobernanza de los sistemas de inteligencia artificial, y en este sentido destaca que deben tratarse de sistemas que sean interoperables, ágiles, adaptables e inclusivos y respondan a las distintas necesidades y capacidades de los países desarrollados y los países en desarrollo por igual, en beneficio de todos. Con el fin de salvar las brechas digitales para lograr los 17 Objetivos de Desarrollo Sostenible en sus tres dimensiones económica, social y ambiental, con especial atención a los países en desarrollo.

En consonancia con lo anterior, exhorta tanto a los Estados, organizaciones internacionales y regionales, así como al sector privado a la cooperación con miras a lograr un acceso inclusivo y equitativo a los beneficios de los sistemas de inteligencia artificial, a partir de ampliar la participación de todos los países en la transformación digital mediante la promoción de actividades de intercambio y el aumento de la alfabetización digital. Fomentar alianzas efectivas para el mejoramiento de las infraestructuras digitales y el acceso a sistemas que incluyan a todos los sectores poblacionales y generen impactos positivos para los mismos.

De igual manera, la Asamblea reconoció los “diferentes niveles” de avance tecnológico entre los países y dentro de ellos, y que las naciones en desarrollo se enfrentan a retos únicos para seguir el rápido ritmo de la innovación (Organización de las Naciones Unidas, 2024b). En consonancia, destaca que durante todo el ciclo de vida los sistemas de inteligencia artificial deben respetar, proteger y promover los derechos humanos y las libertades fundamentales, así mismo exhorta al no uso de sistemas que no estén en concordancia con el Derecho Internacional o que supongan riesgos indebidos, sobre todo para las comunidades más vulnerables, reafirmando que los derechos de las personas

también deben estar protegidos en Internet, durante el ciclo de vida de estos sistemas.

La resolución también busca promover la elaboración y la aplicación de enfoques y marcos regulatorios y de gobernanza nacionales para apoyar la inversión responsable en inteligencia artificial para el desarrollo sostenible. Alentar medidas eficaces para la prevención y mitigación a nivel internacional de los riesgos durante el diseño, desarrollo y despliegue de sistemas de inteligencia artificial. Así como, el establecimiento de mecanismos de retroalimentación que permitan la notificación por parte de los usuarios y terceros de las vulnerabilidades y usos indebidos, incluyendo los incidentes relacionados con la inteligencia artificial.

Asimismo, fomentar la aplicación y divulgación de mecanismos de seguimiento y gestión de riesgos, protección de datos –incluidos los personales- y políticas de privacidad. Facilitar la elaboración de marcos y prácticas para el entrenamiento y la puesta a prueba de los sistemas de inteligencia artificial con el fin de proteger a las personas frente a cualquier forma de discriminación, sesgo o uso indebido, que contribuyan a perpetuar resultados discriminatorios y sesgados por parte de estos sistemas inteligentes. Promover la transparencia y fiabilidad de estos sistemas, así como la supervisión humana y la rendición de cuentas con el objetivo de garantizar la privacidad y la protección de los datos personales.

La Asamblea General persigue alentar la cooperación internacional para elaborar medidas encaminadas a evaluar el posible impacto de la implementación de sistemas de inteligencia artificial en los mercados laborales en aras de mitigar consecuencias negativas para la fuerza de trabajo (Comisión Económica para América Latina y el Caribe, 2022; Martínez, 2023). Reconoce además, el papel de los datos para el desarrollo y funcionamiento de estos sistemas, poniendo de relieve que la gobernanza justa y eficaz de los datos, así como la mejora en cuanto a la accesibilidad son elementos esenciales para aprovechar el potencial de sistemas seguros y

fiables, que generen beneficios para todos y en pos de acelerar la consecución de los 17 Objetivos de Desarrollo Sostenible.

De manera general la resolución A/RES/78/265 “Aprovechar las oportunidades de sistemas seguros y fiables de inteligencia artificial para el desarrollo sostenible” de la Asamblea General de las Naciones Unidas representa una contribución histórica, por aclamación y sin votación,122 basándose, sobre todo, en la cooperación internacional para facilitar la inclusión y representación de los países en desarrollo en el diseño de estos marcos regulatorios en materia de inteligencia artificial.

Sin embargo, no deja de ser un documento meramente declarativo, que no impone como obligación a los Estados signatarios establecer mecanismos regulatorios en materia de inteligencia artificial que incluyan aspectos éticos; y por tanto su cumplimiento no puede ser exigido. La redacción del texto evidencia el predominio del *soft law* como una característica intrínseca del Derecho Internacional Público, sin efectos jurídicamente vinculantes para las partes, y mucho menos una sanción por no incluir en las legislaciones internas aspectos referentes al marco ético de la inteligencia artificial.

Resolución A/RES/78/311 “Aumentar la cooperación internacional para la creación de capacidad en materia de inteligencia artificial” aprobada por la Asamblea General el 1 de julio de 2024 por la Asamblea General de las Naciones Unidas.

En la Resolución se reconoce que el rápido avance de la inteligencia artificial puede brindar nuevas oportunidades para el desarrollo socioeconómico y acelerar los progresos hacia los Objetivos de Desarrollo Sostenible y su consecución y el desarrollo sostenible en sus tres dimensiones: económica, social y ambiental, de forma equilibrada e integrada. De igual manera, incluye que el diseño, desarrollo, despliegue y uso incorrectos o maliciosos de los sistemas de inteligencia artificial; por ejemplo,

sin salvaguardias adecuadas o de manera contraria al derecho internacional, podrían plantear riesgos y desafíos potenciales.

También se conciben los propios principios de la Recomendación de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, en tanto se realiza sinergia con los mismos cuando establece que los sistemas de inteligencia artificial están centrados en las personas, son fiables, se pueden explicar, son éticos e inclusivos y respetan plenamente la promoción y la protección de los derechos humanos y el derecho internacional, deben ser seguros y fiables, en consonancia con el principio de una inteligencia artificial para el bien de todos, en el marco de una visión de la sociedad de la información centrada en las personas, inclusiva y orientada al desarrollo, teniendo en cuenta que la presente resolución se centra en la cooperación internacional para la creación de capacidad en materia de inteligencia artificial en el ámbito no militar y no afecta al desarrollo o uso de la inteligencia artificial con fines militares.

Por otra parte, se prevé el respeto de las normas jurídicas, incluidas la de Propiedad Intelectual. Es importante hacer notar que esto constituye uno de los principales desafíos que enfrenta el fenómeno en materia legal en el presente. Vinculado al acceso a la innovación promueve los softwares de código abierto, los modelos abiertos y los datos abiertos, entre otros métodos y modelos de negocio.

Resuelve reducir la brecha digital y otras brechas digitales entre los países y dentro de ellos, y mejorar la cooperación internacional en materia de creación de capacidad en los países en desarrollo, incluso mediante la cooperación Norte-Sur, Sur-Sur y triangular, teniendo plenamente en cuenta las necesidades, políticas y prioridades de los países en desarrollo, con el fin de aprovechar los beneficios de la inteligencia artificial, minimizar sus riesgos y acelerar la innovación y el progreso hacia la consecución de los 17 Objetivos de Desarrollo Sostenible.

CAPÍTULO 2

Fundamentos éticos vinculados a la inteligencia artificial

2.1. Ética- Derecho-Ciencia: breve aproximación

Habitualmente se habla de ética y/o de moral, no son términos nuevos ni desconocidos, aunque con la misma frecuencia se les suele confundir, incluso ignorar la verdadera connotación que tienen. Aquí resulta acertado hacer una distinción entre ambos, a saber, que no son lo mismo a pesar de que etimológicamente sus significados se conciben como sinónimos (Jalomo-Aguirre, 2012).

Así, al decir de Cortina (2013):

> La palabra ética, nacida del griego *ethos*, se refiere pues al carácter que forjamos en nuestro madurar, para cumplir con el fin mismo de la vida humana. Mientras que la moral, del latín *mos-moris*, se refiere al carácter, costumbres y usos, pero también a la morada en que habita el individuo. De este modo, la ética trata de la formación del carácter de las personas, de las instituciones y de los pueblos (p. 34).

En igual sentido, la Real Academia de la Lengua Española define la ética como ese conjunto de normas morales que rigen la conducta de la persona en cualquier ámbito de la vida. Otra de sus acepciones lo define como parte de la filosofía que trata del bien y del fundamento de sus valores.

Por su parte, el filósofo griego Aristóteles (384-322 a.n.e), en sus escritos a su hijo Nicómaco planteaba que la vida propiamente humana es la vida ética, que debe lograrse mediante el cultivo de virtudes éticas, y que en ello consiste además la felicidad y, por esto, la ética y la política son la realización del fin de la naturaleza humana (Toledo, 2013). Así mismo, distingue dos tipos de virtudes –una moral y otra intelectual- a saber, que las primeras nacen del hábito y de las costumbres (Aristóteles, 2016). Bajo esta visión aristotélica la ética se identifica con la moral.

Desde un punto de vista más moderno, la ética tiene un sentido puramente teórico, y se compone ya no de dirigir las conductas humanas sino del estudio de las conductas morales. Dejando a la moral entonces, esas costumbres referentes a regular las acciones humanas encaminadas al bien. La ética pasa a ser, efectivamente, una rama o disciplina filosófica con un enfoque en categorías y conceptos.

Si bien la ética no dicta normas bajo la estructura de reglas de conducta (Figueroa Jácome y Barrios Lira, 2014) no puede encerrársele en un plano que describe y analiza meramente ya que sus reflexiones fundamentan los principios que sirven de base a dichas reglas de conducta.

La ética es irreductible a una ciencia exacta y la causa fundamental de ello radica en que su objetivo esté aparejado al mejoramiento de la vida humana. Entendido esto, la vida humana no es más que la vida del hombre como ser que se relaciona y se desarrolla en sociedad; por tanto, la ética incluye necesariamente, decisiones que afectan a los demás.

Gabriel (2021) plantea que existen tres ámbitos morales para la actuación del hombre: lo bueno, lo malo y lo neutro; valores éticos de validez universal, que van más allá de diferencias culturales y temporales. Esos valores no solo indican lo que puede hacerse, sino aquello para lo que no hay margen de realización; la ética es la encargada de determinar qué acciones son esas que son de relevancia para la vida en sociedad. Visto así, no va de establecer un margen esquemático sobre lo que está bien o lo que está mal, sino que se enfoca en analizar los diferentes criterios que justifican por qué actuar de un modo o de otro.

La generalización o universalización de los preceptos éticos, va más allá de las distintos principios y valores morales que existen, incluso llegan a trascender fronteras espaciales y temporales. Esta es una característica particular de la ética, además de la eminente connotación histórica que representa ya que existe homogeneidad en cuanto todos los seres humanos tiene una competencia moral que les permite hacer juicios de valor (Castrillón et al., 2023).

Otra posición, para nada discordante con lo que ya se ha abordado, según de Zan (2024) la ética:

> Alude en este sentido a una concepción de la buena vida, a un modelo de la vida virtuosa y a los valores vividos de una persona o de una comunidad, encarnados en sus prácticas e instituciones. La "ética" así entendida se interesa ante todo por el sentido o la finalidad de la vida humana en su totalidad, se interesa por el bien o el ideal de la vida buena y de la felicidad (p. 22).

En este sentido, los principios confluyentes de la ética son reflejo de la evolución del hombre como ente individual y social (Palencia, 2020) modelos a los que se aspiran impregnados de justicia, igualdad, solidaridad por solo mencionar algunos. Sin embargo, dichos modelos no determinaran el actuar del hombre,

quien, en un acto subjetivo bajo la responsabilidad individual, deberá encontrar el justo medio para dirigir su actuar, sin obligar en ese acto a toda la sociedad en su conjunto.

Fue Hegel quien, en un plano personal otorgó autonomía normativa a la eticidad categorizándola constitutiva de un conjunto de sistemas normativos sociales o colectivos, del cual forma parte el Derecho (Friedrich Hegel, 2004).

Por tanto, la ética puede definirse como el conjunto de normas encaminadas a regir la conducta humana; y desde una visión más amplia como esa disciplina que estudia analiza y sistematiza los conceptos y valores morales de carácter universal, que permiten la evolución de la vida en sociedad.

Por otra parte, vinculado a lo legal la relación entre ética y Derecho constituye un tema polémico alrededor del cual se han generado múltiples estudios y discusiones por parte de filósofos y juristas a lo largo de la historia; entre ellos destacan: Enmanuel Kant, con su Introducción a la Teoría del Derecho, Carlos Cossío, con La teoría de la verdad jurídica, Luigi Ferrajoli, con Derecho y razón y Gregorio Peces-Barba, con Ética, poder y Derecho. Se trata de un debate que no tiene una única respuesta, sobre todo, si se tiene en cuenta la diversidad de criterios que existen entre sus estudiosos. La ética rige la conducta en aras de garantizar el desarrollo del hombre como ser social. Desde esta posición resulta fácil comprender que está influenciada entonces por factores históricos, sociales incluso culturales; mientras que el Derecho, como ente normativo y regulador de la conducta humana, igualmente ha estado condicionado por factores históricos, políticos y sociales (Palencia, 2021).

Los juristas en su misión de aplicar el Derecho Positivo deben remitirse a sus normas e interpretarlas. Sin embargo, como la realidad suele ser más amplia que el Derecho, lo que el ordenamiento jurídico fija como respuesta, no coincide con lo que verdaderamente resolvería el asunto en cuestión. En

esta situación se debe buscar el justo medio que posibilite un equilibrio entre la norma positiva y las cuestiones éticas, buscar una interpretación justa (Ramírez-García, 2008), de ese Derecho apelando a criterios morales.

Por tanto, no es suficiente entender el Derecho y las normas jurídicas desde un punto de vista meramente técnico, eso sería privarlo de esos principios éticos, capaces de matizarlo y reducirlo a norma y solo norma (Kelsen, 1982); es prescindir de valores éticos que garantizan un reconocimiento y una legitimación, suficiente dentro del entramado social. Aunque en este punto resulta necesario dejar claro, que la norma positiva de Derecho se impondrá por encima de criterios éticos, así está previsto y regulado por los ordenamientos jurídicos.

Esta relación entre ética y Derecho se traduce al aspecto científico y en especial al de la Inteligencia Artificial en tanto, hoy día no solo se apela a pronunciamientos legales para ordenar el tema sino también a cuestiones estrictamente éticas que deben respetar usuarios y desarrolladores de esas tecnologías.

Por otra parte, en este acápite es esencial analizar el vínculo ética-ciencia, en tanto dicha relación también trasciende al fenómeno que aquí se analiza.

La ciencia, al ser una actividad humana no puede ser vista como un fenómeno neutral. La interacción entre esta y la ética tiene un crecimiento exponencial en la época contemporánea. Sobre todo, si se toma en consideración que el auge de avances científicos y tecnológicos que se producen a diario, traen aparejados nuevos dilemas éticos, que requieren el diálogo constante entre investigadores, instituciones, la sociedad y la intervención del Derecho como mecanismo que garantice una estabilidad en el plano social.

En campos como la Ingeniería Genética, centrada en investigar técnicas para retirar, modificar o agregar genes a moléculas de

un organismo ya sea un humano, planta o animal. Tal es el caso de la clonación de seres humanos, que permite obtener seres con idéntico material genético que el de su donador, lo cual significa un atropello a la individualidad bilógica y desencadenas una serie de trastornos genéticos para los que la ciencia no tiene solución. De otro lado, se encuentra la producción de organismos transgénicos con fines comerciales, que puede generar efectos sobre la salud humana como el aumento de la toxicidad, resistencia a antibióticos, recombinación de virus y bacterias, etc. En otros campos como el de las llamadas tecnologías inteligentes que suscitan polémicas en torno a la responsabilidad cuando sistemas con estas tecnologías producen un daño, o cuando empleados de una industria son reemplazados de su puesto laboral por robots inteligentes, o cuando es vulnerada la privacidad de usuarios en el ámbito digital y revelados sus datos; incluso para la fabricación de material bélico lo que genera un desequilibrio en el orden público internacional, por solo hacer mención a algunos ejemplos.

Es cierto que la comunidad científica ha incorporado planteamientos éticos en torno a sus actividades, generando incluso códigos de buenas prácticas y que son aceptados como normas de obligatorio cumplimiento, orientándose en este sentido, la ciencia hacia el bienestar social. Sin embargo, en general se quedan en cuestiones de principios y actúan más como un decorado moral de autorregulación, sin lograr forzar un marco legal obligatorio para la actividad científica (Baeza, 2018).

Hoy día, la ciencia debe estar llamada a exaltar valores como la verdad, la libertad: debe ser útil. Asimismo, permite la autodeterminación y consecuentemente el autogobierno, tanto a escala individual como colectiva (Schulz, 2005).

En contraposición, la ciencia puesta en manos de la destrucción, de políticas represivas y de guerra, ya se ha demostrado a lo largo de la historia, lo catastrófico de sus consecuencias. Se habla de falta de ética en las aplicaciones de la ciencia, culpando

a científicos e investigadores de la mayoría de los males que acusan a la humanidad; desconociendo que a veces son terceros los que se encargan de administrar por medio de sus políticas esos conocimientos científicos.

Reforzando lo antedicho, desde octubre del 2017, la Unión Europea (2024), financia el Proyecto SIENNA *(Stakeholder-informed ethics for new technologies with high socio-economic and human rights impact)*. Su objetivo es realizar una propuesta de marco legal para la regulación de tres áreas tecnológicas de gran impacto, la genética humana y la genómica, la mejora del ser humano, y la inteligencia artificial y robótica.

El mundo contempla un avance vertiginoso y sin precedentes de la ciencia y la técnica, entre las que se inscribe IA, que se alza como cúspide del conocimiento. Con ella, se abren paso nuevos problemas emergentes asociados a la transferencia parcial o total de la toma de decisiones a sistemas técnicos (Delgado Díaz, 2023).

Problemas que versarán sobre todo en las consecuencias éticas que traerá aparejado el comportamiento y el uso de esta tecnología, creada por el propio hombre. Básicamente porque "se trata de sistemas que ni son morales ni producen reflexiones morales, pero toman decisiones, que sí tienen significado moral para los seres humanos y exigen replantearnos la pregunta por su moralidad" (Delgado Díaz, 2021, p. 41).

Por todo lo que actualmente la inteligencia artificial representa, se hace indispensable una mirada ética a esta tecnología, que analice los beneficios y/o los riesgos de su uso. Ética para marcar límites, establecer pautas y fomentar valores en las actuales generaciones, para las que el desarrollo tecnológico forma parte inescindible de su cotidianidad. Comprender que la ética no es un añadido, sino un componente intrínseco que guía y orienta la práctica científica. Su importancia para la ciencia está en que, al ser parte del proceso evolutivo de la humanidad, acompaña

y estimula los nuevos conocimientos, y la forma en que estos deben ser abordados. Una ciencia desde el respeto, la igualdad, la responsabilidad y que rompa los sesgos de la discriminación, es una ciencia comprometida con la ética.

2.2. Valores éticos y el uso de la inteligencia artificial

Los valores y principios éticos constituyen el primer límite a considerar respecto al uso de sistemas de inteligencia artificial. Estos valores éticos, que sirven como barrera al uso desmedido y desenfrenado de la inteligencia artificial, deben comparecer desde la primera fase del diseño de la tecnología.

En 2018, el proyecto AI4 People llegó a contabilizar 47 principios éticos proclamados internacionalmente. No obstante, la propia organización logra sintetizarlos en 5 principios. En este sentido, establece el principio de beneficencia (promover el bienestar, preservar la libertad y sostener el planeta); no maleficencia (privacidad, seguridad y "precaución de capacidad"); autonomía (el poder de decidir, o si decidir o no); justicia (promover la prosperidad y preservar la solidaridad), estos como principios básicos, y el último que se agregó la explicabilidad (permite la habilitación de otros principios a través de la inteligibilidad y la rendición de cuentas).

La Recomendación sobre la ética de la inteligencia artificial, emitida por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, proporciona una serie de valores éticos, que han de considerarse como límites éticos a nivel global para el empleo y aplicación de los sistemas de inteligencia artificial. En este sentido, se encuentran los siguientes valores:

El respeto, la protección y la promoción de los derechos humanos, la dignidad humana y las libertades fundamentales: límite que constituye la base para el desarrollo de la inteligencia artificial porque se trata de una condición que es inherente a cada ser humano, en tanto derecho inalienable de cada individuo. la

inteligencia artificial no puede ir contra la dignidad humana en ninguna etapa de su ciclo de vida, independientemente de la raza, el sexo, la religión, el origen o etnia, discapacidad del individuo. En el marco de interacción de la sociedad debe mejorar la calidad de vida social, sin menoscabar, abusar o violar las libertades particulares de cada cual, por el contrario, debe promover, impulsar y defender los derechos humanos (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 2024).

La prosperidad del medio ambiente y los ecosistemas: es un tema que involucra no sola a las generaciones presentes sino también a las futuras. Todos los agentes involucrados en la gestión y el despliegue de sistemas de inteligencia artificial deben enfocarse en garantizar que estos sean sostenibles y respetuosos con el medio ambiente (High-Level Expert Group on Artificial Intelligence, 2019). Debiendo reducir el impacto ambiental a partir de los grandes centros de procesamiento de datos que trabajan con fuentes de energías no renovables. Por otro lado, se encuentra la gran cantidad de desechos electrónicos generados a partir del sistema actual de producción y consumo (Organización de las Naciones Unidas, 2024ab) de inteligencia artificial. En este sentido, estos sistemas se encuentran en calidad de asegurar la mitigación del cambio climático.

Diversidad, no discriminación, inclusión y equidad: son puntos de referencia a considerar en el diseño y aprendizaje de los sistemas inteligentes que deben estar exentos de sesgos por motivos de color de la piel, nacionalidad, idioma, condición económica o cualquiera otra circunstancia. En este sentido The Guidelines de la Comisión Europea señalan que deben evitarse los prejuicios injustos y discriminatorios, ya que generan implicaciones negativas que van desde la marginación de grupos vulnerables hasta la exacerbación de los prejuicios. También, se trata de que se generen beneficios y prosperidad compartida, lo cual

incluye eliminar las discriminaciones históricas y prevenir nuevos estigmas.

Resistencia y seguridad: en este aspecto brindar un plan de contingencia en caso de producirse daños no deseados. La seguridad y la protección de la inteligencia artificial se materializa a través de marcos sostenibles de acceso a los datos, respeto a la privacidad y un entrenamiento de estos sistemas a partir de datos de calidad. Se trata de la exigencia de algoritmos seguros, fiables y sólidos para resolver errores o incoherencias durante todas las fases del ciclo de vida útil de los dispositivos inteligentes (Villalba, 2020) esto es integridad y seguridad tanto para los usuarios como para los sistemas inteligentes.

Derecho a la intimidad y a la protección de datos: límite que se encuentra estrechamente ligado a los derechos humanos y las libertades fundamentales. En este sentido, resulta importante que los datos utilizados por los sistemas de inteligencia artificial se recopilen, utilicen, compartan, archiven y supriman de forma coherente. Además, garantizar el pleno respeto de la privacidad y la protección de los datos personales, obtener el consentimiento informado de quienes brindan información es un requisito clave si lo que se busca es un uso ético de la inteligencia artificial. La intimidad forma parte de la autonomía y la capacidad que tiene todo ser humano de actuar conforme a sus intereses, por ello el uso que los sistemas de inteligencia artificial hagan de esta deben ser de acuerdo a marcos de protección de datos previamente establecidos para el tratamiento de datos personales.

Límite a la autonomía, responsabilidad y la rendición de cuentas: por parte de los sistemas de inteligencia artificial (Cortina Orts, 2019) la cuestión radica en que la autonomía es una condición específica del ser humano. Cuando se habla de la inteligencia artificial podemos estar frente a sistemas autómatas, o sea, sistemas que en su diseño están preparados para tomar ciertas decisiones, pero no para asumir sus consecuencias. Puede

suceder que, el ser humano por cuestiones de eficacia decida dejar determinada decisión ya que esta puede reemplazar al hombre para la ejecución de tareas puntuales, sin embargo, un sistema de inteligencia artificial nunca podrá reemplazar la responsabilidad final de los seres humanos y su obligación de rendir cuentas.

El caso de la responsabilidad y la rendición de cuentas deben designarse desde la fase del diseño de la inteligencia artificial no es excusa la autonomía de la máquina para diluir las responsabilidades. Por autónoma que sea el sistema, esa autonomía viene dada por una programación humana (Sebio Martín, 2020).

En cualquier caso, siempre debe ser posible atribuir la responsabilidad ética y jurídica donde intervengan sistemas de inteligencia artificial, a personas naturales o jurídicas, la supervisión humana es un requisito sine qua non para el uso de estos. Ademas la responsabilidad, La auditabilidad, puede desempeñar un papel de vanguardia en estos aspectos porque por medio de ella se realiza la evaluación de algoritmos y de datos, así como de los procesos de diseño.

Por otro lado, Floridi et al., también lo señala como un principio básico. En este sentido los humanos siempre deben conservar el poder de decidir qué decisiones tomar (Floridi et al., 2018). Las mejoras sociales y el impacto positivo de la inteligencia artificial no pueden ser a costa de reducir o de menoscabar el control humano sobre cualquier esfera, ni tampoco de limitar la prevención de daños y el establecimiento de un responsable ante la ocurrencia de un daño generado por estos sistemas.

En relación con esto, el reparto de responsabilidades entre los sujetos y los sistemas de inteligencia artificial debe seguir los principios éticos inherentes al ser humano (Guerrero Arévalo, 2024), dejando un amplio margen de decisión a las personas.

Para ello hay que garantizar la vigilancia y el control de estos sistemas por parte de los seres humanos.

Transparencia y explicabilidad: desde esta óptica las decisiones que tomen los mismos tienen que explicarse de una manera que involucre a las partes interesadas. En casos así, los seres humanos deben ser conscientes de que están interactuando con sistemas de inteligencia artificial y como consecuencia ser informados de las capacidades y limitaciones de estos sistemas.

Por tanto, estos requisitos tienen que formar parte de los modelos inteligentes puesto que son condiciones previas que posibilitan el respeto y la promoción de los derechos humanos, de esta forma es necesario que cualquier decisión tomada por una máquina inteligente pueda ser trazada, es decir, que se entienda el razonamiento seguido y que se puedan identificar los datos utilizados y los pasos seguidos.

Un tanto contrastante resulta entonces que, si bien, debe potenciarse la transparencia y explicabilidad de estos sistemas, ello debe hacerse en correspondencia con el contexto y los efectos que se genere, ya que debe respetarse, en todo caso, el límite de la privacidad, la protección y la seguridad de los datos.

La transparencia tiene como objetivo proporcionar información adecuada a los respectivos destinatarios para permitir su comprensión y fomentar la confianza. Para que los consumidores depositen su confianza en los sistemas de inteligencia artificial y la mantengan, resulta fundamental la explicabilidad. Es decir, que los procesos sean claros, que se informe públicamente de las posibilidades y la utilidad de los sistemas de AI y que las decisiones se expliquen, en la medida de lo posible, a las personas afectadas directa o indirectamente por ellas (Guerrero Arévalo, 2024).

Bajo estos enfoques la inteligencia artificial no puede ir más allá que desarrollarse para el bien común y el beneficio de la

humanidad, mejorar el bienestar individual y colectivo, generar prosperidad, valor y maximizar la riqueza y sostenibilidad. Asimismo, debe buscar una sociedad justa, inclusiva y pacífica, ayudando a aumentar la autonomía de los ciudadanos, con una distribución equitativa de oportunidades económicas, sociales y políticas.

Según plantea Cotino Hueso (2019), los sistemas de inteligencia artificial: también debe tener objetivos como la protección del proceso democrático y el Estado de derecho; la provisión de bienes y servicios comunes a bajo costo y de alta calidad; alfabetización y representatividad de los datos; mitigación de daños y optimización de la confianza hacia los usuarios.

Así, de manera general, los valores que sirven como límites éticos al uso de la inteligencia artificial son el respeto, la protección y la promoción de los derechos humanos, la dignidad humana y las libertades fundamentales; prosperidad del medio ambiente y los ecosistemas; diversidad, no discriminación, inclusión y equidad; resistencia y seguridad; derecho a la intimidad y a la protección de datos y la transparencia y explicabilidad.

En todo caso, lo que sí debe quedar claro es que los sistemas de inteligencia artificial son los que deben adaptarse al ser humano y a sus singularidades, nunca viceversa, ya que en definitiva es el propio ser humano quien crea y desarrolla dichas tecnologías. con el bien de la sociedad, que es a lo que definitiva debe aspirarse.

2.3. Desafíos éticos en tecnologías específicas (aplicaciones biomédicas, reconocimiento facial, interfaces cerebro-computadora)

La ética en la ciencia ha sido objeto de una larga y profunda reflexión, y el ámbito de la inteligencia artificial (IA) no constituye una excepción a esta tendencia. A medida que las tecnologías de inteligencia artificial se desarrollan y consolidan, emergen crecientes preocupaciones relativas a los riesgos inherentes a

su implementación, tales como la falta de transparencia en sus procesos, la perpetuación de sesgos y formas de discriminación, la complejidad en la asignación de responsabilidades frente a decisiones autónomas adoptadas por los sistemas, y la protección efectiva de los datos personales.

Estos desafíos éticos no solo reflejan problemáticas técnicas, sino que también impactan directamente en la dignidad y los derechos fundamentales de las personas. En consecuencia, sin una gobernanza adecuada y reflexiva, la adopción irresponsable de estas tecnologías puede poner en riesgo valores humanos esenciales (Breceda y Castillo, 2023).

El Libro Blanco sobre la Inteligencia Artificial (Comisión Europea, 2020) identifica y agrupa un conjunto de lo que considera como principales riesgos asociados a la misma en: riesgos para los derechos fundamentales, riesgos para la seguridad y riesgos para las cuestiones relativas a la responsabilidad civil.

En primer lugar, los riesgos para los derechos fundamentales se refieren a las posibles vulneraciones de principios esenciales como la privacidad, la no discriminación y la libertad individual, especialmente cuando la inteligencia artificial gestiona datos personales o toma decisiones automatizadas que afectan a los ciudadanos. En segundo lugar, los riesgos para la seguridad aluden a las amenazas que pueden surgir a nivel físico o digital, incluyendo fallos en sistemas críticos, ataques cibernéticos o un mal desempeño que comprometa la integridad de infraestructuras y la protección de las personas.

Finalmente, los riesgos relativos a la responsabilidad civil destacan la dificultad de atribuir y garantizar responsabilidades legales en escenarios donde las decisiones son tomadas autónomamente por sistemas de inteligencia artificial, planteando desafíos regulatorios sobre quién responde ante daños o perjuicios derivados de dichas tecnologías. Esta clasificación proporciona un marco conceptual esencial para entender y abordar las

implicaciones éticas y jurídicas inherentes al desarrollo y despliegue de la inteligencia artificial, con miras a salvaguardar valores democráticos y derechos fundamentales en el contexto tecnológico actual.

Dentro de esa lógica, una de las preocupaciones éticas fundamentales en el desarrollo de la inteligencia artificial se ha transformado en discusiones sobre transparencia y explicabilidad respecto a la opacidad y falta de conocimiento que existe sobre qué hace (Arriagada-Bruneau, 2024); así como acerca de sus datos de entrenamiento. De modo que la transparencia se refiere a la obligación de hacer visibles y comprensibles los procesos internos mediante los cuales los algoritmos de inteligencia artificial toman decisiones, ofreciendo claridad respecto a la lógica, los criterios y los datos que subyacen a sus resultados. En tanto, la explicabilidad busca que estas decisiones no solo sean accesibles en términos técnicos, sino que puedan ser interpretadas y justificadas de manera que los usuarios y las partes interesadas comprendan el porqué y el cómo de cada resultado.

Paralelamente, la selección y calidad de los datos utilizados para entrenar estos modelos revisten una importancia crítica, dado que pueden reproducir y amplificar sesgos sociales existentes, perpetuando dinámicas discriminatorias e injustas en la sociedad. De esta forma, la gestión responsable de la transparencia y la explicabilidad no solo contribuirá a la rendición de cuentas y a la detección de errores, sino también a la mitigación de impactos negativos en la equidad y los derechos humanos, preservando así valores fundamentales en contextos de creciente automatización y autonomía tecnológica.

Asimismo, se presentan un conjunto de riesgos a la seguridad, que abarcan tanto la seguridad física como la lógica (Muñoz, 2021), los cuales suelen tener su origen en deficiencias del diseño, en la calidad o disponibilidad inadecuada de los datos, así como en

un uso inapropiado o malintencionado de la tecnología. Dada la posibilidad de fallos que comprometan la seguridad de individuos o infraestructuras la integración de la inteligencia artificial en productos y servicios acarrea potenciales daños tanto a personas físicas como jurídicas. Asimismo, el uso malicioso o delictivo de estos sistemas puede amplificar esos riesgos, generando vulnerabilidades que ponen en peligro la integridad física, la privacidad y la continuidad operativa de diversas instalaciones y servicios.

Con respecto al acceso creciente a datos se están planteando cuestiones éticas y jurídicas claves: sobre todo en relación con las personas afectadas por el tratamiento de los mismos. La tecnología del *big data* ha propiciado un fenómeno muy preocupante denominado *neuromarketing* (BlogThinkBig, 2017), que no es más que la lectura de la mente de los consumidores para medir sus reacciones a los estímulos de marketing. Generando un conjunto de efectos que van más allá de la simple mejora en la segmentación y personalización de la publicidad, dado que puede influir en la toma de decisiones de los consumidores de manera subliminal, provocando un potencial riesgo de manipulación. Asimismo, el uso masivo y complejo de datos personales y biométricos mediante big data hace que el neuromarketing pueda consolidar perfiles detallados de comportamientos y preferencias, que, sin una regulación ética estricta, puede conducir a una erosión de la libertad individual y a la reproducción de sesgos sociales ya existentes.

Por otro lado, se halla el tema de la realidad virtual potenciada por IA, la cual presupone numerosos retos y riesgos, especialmente en materia de privacidad, debido al manejo y procesamiento de datos íntimamente ligados a la identidad personal, comportamientos y pensamientos más profundos. Esta tecnología implica un riesgo significativo frente al uso indebido, la falta de autorización o la sustracción de información sensible relacionada

con sentimientos, juicios y características físicas que pueden ser únicas entre individuos. La naturaleza altamente personal de los datos recopilados, como señales biométricas o patrones de comportamiento, expone, asimismo, a los usuarios a posibles vulnerabilidades, incluida la manipulación psicológica, robos de identidad digital y ataques cibernéticos que comprometen la integridad tanto física como digital.

Igualmente se presenta el riesgo de experiencias falsas difíciles de distinguir de la realidad, lo que puede llegar a influir de manera profunda en los comportamientos, opiniones y decisiones de los usuarios. Identificándose, además, riesgos importantes para la salud física y psicológica, derivados del posible distanciamiento del mundo real y de efectos mentales cuyo alcance y duración aún resultan desconocidos, dependiendo del uso y el contexto en que se empleen estas tecnologías. La conexión directa que estas tecnologías establecen con la mente y la percepción de la realidad plantea desafíos interpretativos y éticos, dada la limitada comprensión científica actual sobre sus implicaciones a largo plazo en la cognición y el bienestar humano.

En igual medida, se encuentran las interfaces neuronales, las cuales conectan el cerebro o el sistema nervioso con dispositivos digitales o sistemas informáticos. Algunas de estas interfaces se utilizan para registrar la actividad fisiológica, como las señales o movimientos cerebrales, otras se encargan de estimular estas funciones, y algunas realizan ambas actividades simultáneamente. Estas tecnologías representan riesgos significativos para empresas, gobiernos y ciudadanos, particularmente en materia de privacidad y otros derechos humanos, así como en términos de desigualdad social y seguridad. Maximizado por el hecho de que, en relación directa con los interfaces precitados, la biotecnología, neurotecnología o la nanotecnología van a tener un impacto cada vez mayor a partir de su aplicación en intervenciones en el cuerpo

o cerebro con el objetivo de entender, mitigar, predecir o curar enfermedades (Muñoz, 2021).

Del mismo modo se plantean los drones y los sistemas biométricos como tecnologías muy influyentes en la identificación de patrones delictivos y la asignación de recursos más eficientemente en la lucha contra el crimen organizado, el terrorismo y los delitos cibernéticos. No obstante, pese a los enormes beneficios que ello trae consigo, estas tecnologías también han generado preocupaciones sobre su impacto en los derechos individuales y la ética policial. Por un lado, la calidad de los datos de entrada puede perpetuar sesgos raciales y socioeconómicos, especialmente al identificar personas de grupos minoritarios, lo que puede llevar a discriminación en su aplicación. Esto puede resultar en una vigilancia excesiva y discriminatoria hacia comunidades vulnerables, socavando su derecho a la igualdad y dignidad (Chavarry et al., 2025).

La vigilancia mediante inteligencia artificial, por su parte, plantea considerables riesgos para los derechos fundamentales. Esta modalidad incorpora tecnologías como el reconocimiento facial, que capturan imágenes de forma anónima en múltiples lugares y a distancia, sin que las personas estén conscientes de dicha captura. Por demás, al hacerse depender su eficacia en gran medida de la calidad de las imágenes disponibles en las bases de datos, la incapacidad para reconocer un rostro en imágenes de baja resolución o las identificaciones erróneas pueden conllevar consecuencias graves, incluyendo fallos en la identificación o la detención injusta de personas inocentes.

Esta vigilancia biométrica, al operar sin un marco regulatorio sólido, no solo vulnera la privacidad, sino que fomenta un clima de desconfianza constante. Lo que unido a la ausencia de transparencia y control puede generar un efecto paralizante, donde los individuos, conscientes de estar potencialmente monitoreados, modifican su conducta por temor a represalias o

errores en la identificación, impactando de forma negativa en la libertad individual y la convivencia social. Además de que muchas veces operan sin el consentimiento informado de la ciudadanía lo cual representa un grave riesgo para la democracia.

Asimismo, la integración de los sistemas de inteligencia artificial dentro de los procedimientos judiciales, sin las salvaguardias necesarias, pone en riesgo el debido proceso, como el derecho a ser oído, la igualdad de las partes y la imparcialidad judicial. La opacidad de estos algoritmos puede limitar la posibilidad de revisar y cuestionar las decisiones basadas en IA, afectando la transparencia y la defensa adecuada. Además, la dependencia de datos con posibles sesgos puede generar discriminación y desigualdad en el tratamiento procesal. Por ello, es fundamental garantizar mecanismos que aseguren la explicabilidad, la supervisión humana.

En el espacio laboral, por su parte, ha comenzado a imperar un fenómeno asociado a la inteligencia artificial desde múltiples ángulos. Se observa una competencia creciente entre el ser humano y las tecnologías de inteligencia artificial debido a las amplias capacidades que estas últimas poseen. El mercado laboral dominado, por la tecnología, sobre todo en los países con altos niveles de desarrollo, plantea desafíos significativos, como el riesgo de un desempleo estructural, ya que predominará la necesidad de contar con profesionales con sólidas competencias digitales y experiencia en el manejo de estas tecnologías.

En este contexto, la creciente necesidad de que los trabajadores posean habilidades digitales para ser competitivos impulsará el acceso a niveles superiores de educación. Esta demanda conducirá a la sobre-cualificación profesional y académica en el mercado laboral, lo que ampliará la brecha social existente. La población con menos oportunidades enfrentará una realidad donde, además de competir entre sí, deberán confrontar a las máquinas, lo que se traduce en salarios más bajos y menores

posibilidades de empleo digno en comparación con quienes cuentan con mayores recursos y formación.

Por último, en relación con lo antedicho, se plantea lo que Byung-Chul Han (2012) define como sociedad del cansancio, caracteriza por un entorno en el que el uso constante de dispositivos electrónicos hace prácticamente imposible para los trabajadores desconectarse de sus obligaciones laborales. En este modelo, la distinción entre tiempo y espacio laboral se diluye, ya que las personas pueden desempeñar sus funciones desde la oficina, su hogar o incluso un lugar de ocio como un club deportivo, generando una presión continua que afecta su bienestar y contribuye al agotamiento.

El procesamiento de lenguaje natural (PLN) en la medicina se ha consolidado como una disciplina de creciente relevancia, especialmente en la mejora de la asistencia médica. Aprovechando estudios previos que destacan la eficacia del aprendizaje profundo, se han desarrollado modelos de inteligencia artificial capaces de asignar notas y códigos médicos para predecir diagnósticos finales a partir de datos no estructurados, como los registros de la enfermedad actual y los síntomas al ingreso del paciente. Estos modelos logran predecir los diagnósticos y procedimientos principales con una precisión cercana al 80%, lo que representa un avance significativo en la automatización y eficiencia del diagnóstico clínico. Sin embargo, factores como el error humano, el tiempo invertido, la sobrefacturación y los costos asociados a los reembolsos dificultan la rápida y flexible evolución de diagnósticos y procedimientos basados en el historial clínico a partir del uso de la inteligencia artificial (Varela-Tapia et al., 2024).

En un ámbito de acción similar actúa hoy en día la neurotecnología, que ha propiciado avances significativos mediante dispositivos capaces de leer, modificar e influir en la actividad cerebral, lo que abre la puerta a una transformación profunda de la experiencia humana. Estos avances permiten no solo potenciar las capacidades

cognitivas y tratar trastornos neurológicos complejos, sino también obtener un conocimiento más detallado sobre el funcionamiento de la mente humana. Además, se vislumbra la posibilidad de establecer conexiones directas y operativas entre el cerebro y las máquinas, lo que representa un salto cualitativo en la interacción entre seres humanos y tecnología.

No obstante, surge la preocupación de que un reducido número de empresas con acceso a esta tecnología pueda rastrear, analizar e incluso influir en los pensamientos, emociones y comportamientos humanos. Además, representa un riesgo considerable, ya que mediante ciertos dispositivos o flujos de información podrían generarse respuestas que no correspondan con nuestros valores o convicciones personales, comprometiendo así la autenticidad de nuestras decisiones y experiencias

Destacando entre los principales riesgos que se vislumbran la posible afectación de la identidad personal, el agravamiento de las desigualdades sociales, la vulneración de la privacidad mental y el impacto sobre el acceso equitativo a estas tecnologías y sus beneficios. Incluso, la capacidad de anticipar intenciones o comportamientos antes de que se manifiesten plantea serias preocupaciones éticas en torno al control, la libertad individual y la presunción de inocencia (Ienca y Malgieri, 2021).

De igual forma, los nuevos modelos de inteligencia artificial, debido a su capacidad para procesar, analizar y replicar datos personales, facilitan nuevas formas de suplantación de identidad. Pudiendo generar videos, fotografías o audios falsos que imitan a personas reales, lo que incrementa significativamente el riesgo de robo de identidad y viola un conjunto de derechos fundamentales como el derecho a la imagen y a la propia voz. Esta problemática, cada vez más prevalente, implica el uso indebido y malintencionado de información personal mediante tecnologías avanzadas, lo que representa un desafío creciente para la seguridad y privacidad digital en la era de la inteligencia artificial.

La concentración del poder en manos de un reducido grupo de grandes corporaciones tecnológicas y gobiernos, que controlan tanto los datos como las infraestructuras digitales, es otro de los riesgos inherentes al uso de la inteligencia artificial. Esta situación puede generar una marcada asimetría en el acceso a la información y en la toma de decisiones, provocando desigualdades significativas que dificultan la protección efectiva de los derechos humanos, especialmente en los sectores más vulnerables de la sociedad (Morozov, 2013).

En ese sentido, la brecha digital se agrava cuando el desarrollo y la implementación de tecnologías de inteligencia artificial están restringidos a contextos con altos niveles económicos, excluyendo así a comunidades y países en vías de desarrollo. Esta situación limita el acceso equitativo a las innovaciones tecnológicas y profundiza las desigualdades existentes, dificultando que amplios sectores puedan beneficiarse del avance tecnológico y quedando rezagados en la revolución digital.

En un intento de concluir lo analizado en este acápite toman fuerza las ideas de Munn (2023) al plantear que existe una brecha entre los principios éticos y la práctica de la inteligencia artificial, siendo esa disparidad especialmente pronunciada debido a la falta de mecanismos eficazmente vinculantes que obliguen a su cumplimiento. Pues, aunque múltiples marcos éticos a nivel internacional establecen directrices claras sobre transparencia, responsabilidad, justicia y privacidad, estos principios suelen carecer de fuerza legal o sancionatoria, lo que dificulta asegurar su observancia en el desarrollo y despliegue de sistemas de inteligencia artificial.

Por lo cual la ausencia de una regulación robusta y la limitada rendición de cuentas generan un vacío que permite la proliferación de prácticas que pueden vulnerar derechos fundamentales, reproducir sesgos o generar efectos discriminatorios. Por ello, el principal desafío ético en inteligencia artificial no reside solo

en definir qué debe hacerse, sino en asegurar los mecanismos que garanticen que efectivamente se cumpla, protegiendo así los valores y derechos en juego en cada aplicación concreta de la tecnología.

2. 4. La inteligencia artificial en la administración pública y la toma de decisiones, aspectos éticos

En el contexto actual, caracterizado por una creciente interconexión, se requiere una administración pública dinámica que promueva la transparencia y garantice un acceso sencillo y eficiente a los ciudadanos-usuarios. En este contexto, las tecnologías emergentes se presentan como herramientas esenciales para modernizar la gestión estatal y afrontar retos en diversas áreas.

Es por ello que a pesar de que la aplicación de la inteligencia artificial en las instituciones gubernamentales aún está poco desarrollada, actualmente se observa una tendencia hacia gobiernos que asumen la digitalización y emplean *chatbots* para la interacción con la ciudadanía, representando un avance significativo respecto a los primeros sistemas interactivos implementados en la atención al cliente.

Se coincide por esas razones con Henman (2020) al establecer que las nuevas tecnologías digitales están transformando aceleradamente el panorama de la prestación de servicios públicos, al integrar dispositivos móviles y aplicaciones que permiten llevar estos servicios en línea a cualquier lugar donde se encuentre el ciudadano, facilitando así un acceso flexible, inmediato y más inclusivo

El moderno enfoque de gobierno inteligente que se ha comenzado a manejar, basado en el uso intensivo de tecnologías emergentes, contribuye significativamente a fortalecer la transparencia, mejorar la toma de decisiones, y reducir los márgenes de discrecionalidad administrativa (Torreblanca Gómez, 2024) a

partir del efecto transformador de la inteligencia artificial en los métodos de almacenamiento y utilización de datos. Asimismo, promueve una mayor interoperabilidad entre entidades públicas, al facilitar el intercambio seguro y eficiente de información, y un mayor seguimiento de las acciones gubernamentales (Ospina Díaz et al., 2025).

Unido a lo anterior, Manet (2014) destaca que la tecnología representa una vía fundamental para optimizar la gestión de las ciudades, al proporcionar herramientas que facilitan la comunicación y el entendimiento entre los responsables de las entidades públicas y los diversos actores de la sociedad. Estas herramientas tecnológicas permiten una administración eficiente basada en la planificación, la colaboración y la participación ciudadana, con el objetivo de promover el bienestar de los habitantes.

De ahí que los instrumentos de inteligencia artificial se presenten como herramientas que ofrecen innumerables aplicaciones en la actividad administrativa, entre ellas: los sistemas de reconocimiento de voz permiten transcribir textos dictados y facilitar en términos de eficiencia la labor de empleados públicos; pueden además ofrecer información a la ciudadanía, realizar llamadas telefónicas para concertar una cita (para la vacunación, por ejemplo), recordar el cumplimiento de un plazo o comunicar la convocatoria de ayudas.

Otros permiten captar imágenes para determinar si un vehículo ha superado el límite de velocidad o ha transitado por una zona de acceso restringido, automatizando la denuncia que le llega al destinatario. Un sistema de lectura de expedientes (subvenciones, contratación, títulos habilitantes, solicitud de plazas universitarias, adjudicación de puestos de trabajo en procesos selectivos de empleo público) puede permitir la propuesta de una resolución administrativa, sobre la base de la configuración informática que se haya realizado del sistema.

Esta referencia somera a algunas de las aplicaciones y usos de la inteligencia artificial en el sector público remite al ejercicio de potestades conocidas: sancionadora, expropiatoria, tributaria, de inspección, así como a actividades vinculadas con el gasto público, como son la actividad subvencional o la contractual. Aun así, los instrumentos de inteligencia artificial son herramientas y su uso no tiene porqué comportar una delegación del ejercicio de la potestad administrativa en sentido estricto.

Entre las múltiples ventajas que comportan las mencionadas aplicaciones de la inteligencia artificial en el ámbito administrativo destacan la prestación de servicios efectivos, la racionalización y la agilidad de los procedimientos administrativos, la eficiencia de los objetivos fijados y la asignación y la utilización de los recursos públicos. Así, mediante la aplicación de algoritmos sencillos e inteligencia artificial, es factible automatizar diversos procesos administrativos. No obstante, su incorporación en la administración pública presenta desafíos particulares, especialmente cuando se intenta emplearlos en la gestión de servicios públicos y en la ejecución de actos administrativos, donde la complejidad y las demandas específicas del sector requieren un manejo cuidadoso y responsable.

De entre las utilidades mencionadas supra, la introducción de análisis predictivo y tecnologías de apoyo a la toma de decisiones es una de las principales formas en que la inteligencia artificial está comenzando a aplicarse en el gobierno local (Vogl et al., 2019). Permitiendo, entre otras cosas, que el gobierno trabaje de manera más eficiente, a la par que deja tiempo libre para que los empleados construyan mejores relaciones con los ciudadanos, fortaleciendo así la participación ciudadana y mejorando la percepción y satisfacción con los servicios públicos.

Ya ha sido documentado que en países desarrollados como España se implementan en su administración pública un conjunto de sistemas de inteligencia artificial amplio que contemplan por

ejemplo a VioGén, RisCanvi, Estrada, Bosco, Max, Tensor, Issa, Hermes, escanea.ods.c~, Send@, VeriPol, a las que se suma un sin número de *chatbots* y asistentes virtuales generalistas y temáticos como Clara, Gon, Rosi, Llum, Guillermina y Llaume, Gestri (Valcárcel y Hernández, 2025). En ese marco, partiendo de la premisa de que los sistemas de inteligencia artificial no poseen una autonomía absoluta; aunque es posible imaginar escenarios con sistemas completamente autónomos se observa que las decisiones automatizadas, aunque sean efectivamente ejecutadas por máquinas, quedarán formalmente atribuidas a una *fictio iuris*: un órgano administrativo y, en última instancia, a un responsable humano que ostenta la titularidad de dicho órgano.

Esto genera diversas cuestiones, por un lado, el debate sobre la personalidad o naturaleza jurídica de los sistemas de inteligencia artificial. Por otro lado, surgen inquietudes respecto a los derechos de los ciudadanos en cuanto al uso de sus datos, la explicabilidad de las decisiones, el derecho a ser informados sobre la utilización de un sistema de inteligencia artificial en un procedimiento específico, y sobre el funcionamiento y alcance de dicho sistema. Asimismo, se plantean aspectos relacionados con la motivación clara y comprensible en lenguaje natural de las decisiones administrativas basadas en IA, la supervisión y posible intervención humana en estas decisiones, la regulación de los recursos administrativos frente a decisiones automatizadas y la trazabilidad de la información de entrada y salida en los expedientes administrativos gestionados mediante IA.

En ese contexto, válido es apuntar que no posee la misma carga de responsabilidad la automatización de actos de mero trámite o de impulso, como pueden ser la expedición del recibo del registro electrónico, la foliación del expediente administrativo, o la remisión de las comunicaciones electrónicas al ciudadano por defecto que un acto administrativo que entre a valorar la aplicación de la norma. Es por ello que alrededor del uso de

sistemas de inteligencia artificial en el segundo grupo de tareas se han generado una serie de dilemas éticos.

La opacidad de los sistemas algorítmicos se manifiesta en tres formas: jurídica, relacionada con la protección de otros derechos e intereses, como la propiedad intelectual; sociológica, que surge de la limitada capacidad para entender su funcionamiento desde un punto de vista tecnológico; y una opacidad intrínseca, inherente al propio modo en que operan estos sistemas. No obstante, junto a estos tipos, la opacidad que primordialmente debe ser combatida desde el ámbito jurídico es la denominada opacidad intencionada, la cual se produce cuando la Administración oculta el uso de estos sistemas algorítmicos y dificulta el acceso a información detallada sobre su funcionamiento. La misma carece de justificación y representa un gran obstáculo para la transparencia y el control público.

Como parte del ámbito administrativo destaca el espacio de la seguridad vial, en el cual ha influido también la inteligencia artificial. Entre sus principales aplicaciones para el control del tráfico se destaca el uso en los sistemas de videovigilancia instalados en las carreteras, conocidos como sistemas de videovigilancia inteligentes. Estos sistemas son capaces de detectar diversas infracciones, no solo cuando un conductor supera la velocidad límite, sino también cuando utiliza el teléfono móvil mientras conduce, no lleva el cinturón de seguridad o si el vehículo transporta a más personas de las permitidas.

Además, está vinculado el manejo de datos personales recogidos por las cámaras con fines de control del tráfico, ya que estas cámaras tienen la capacidad de leer la matrícula de los vehículos e identificar a sus propietarios. Cabe señalar que la regulación actual respecto al uso de cámaras de videovigilancia implica únicamente la obligación de señalizar su presencia, pero no exige informar si estas cámaras utilizan inteligencia artificial en algún grado.

Por este motivo, surge la inquietud acerca de si los avances tecnológicos impulsados por la inteligencia artificial en los sistemas de videovigilancia de tráfico podrían afectar la proporcionalidad de las medidas que limitan derechos fundamentales, dada la amplia gama de posibilidades que ofrece la inteligencia artificial.

Al respecto España (2025) analiza:

> Las mismas ejercen un control general e indeterminado sobre el tráfico como actividad pública que es. En principio, no persiguen la vigilancia de sujetos concretos. Además, en la captación de imágenes en el tráfico se da la particularidad que existe la figura mediata del vehículo y esto puede ejercer un plus de protección al ciudadano, que le brinda cierta indeterminación. Cuestión distinta es la aplicación de técnicas de reconocimiento facial y/o un tratamiento, posterior, de esas imágenes que se repute abusivo o excesivo (p. 378).

La creciente expansión de los sistemas de videovigilancia en el espacio público representa un avance significativo hacia la conformación de una sociedad hipervigilada, promovida y potenciada por el poder público. Este fenómeno se agrava a medida que los sistemas de inteligencia artificial se perfeccionan y se vuelven omnipresentes en todos los aspectos de la vida pública y privada, generando preocupaciones sobre la pérdida de privacidad y el control excesivo. Este contexto plantea serios dilemas éticos y sociales, ya que la presencia constante y avanzada tecnología de vigilancia puede limitar las libertades individuales y amplificar el poder del Estado o de entidades privadas sobre los ciudadanos, haciendo indispensable reflexionar sobre los límites, la regulación y la transparencia en el uso de estas tecnologías.

Por otra parte, con respecto a todo lo anterior surge la siguiente interrogante: ¿cuál es la disposición de la población en aceptar que sus gobiernos utilicen la inteligencia artificial como parte de la estructura administrativa que se encargará de la toma de

decisiones? Aborda esta cuestión Benitez (2025) al concluir que la aceptación de la inteligencia artificial en las instituciones gubernamentales por parte de la población estará condicionada principalmente por tres factores: la confianza que los ciudadanos depositen en el uso de esta tecnología dentro de las estructuras gubernamentales; el grado de acceso que puedan tener a estas herramientas tecnológicas; y la manera en que se realice la integración entre humanos e inteligencia artificial en los procesos administrativos y de toma de decisiones. Estos elementos serán determinantes para lograr una adopción efectiva y positiva de la inteligencia artificial en el ámbito público.

Ahora bien, si la inteligencia artificial permite una mayor eficiencia en la gestión de las administraciones públicas, también es cierto que el elemento negativo consustancial al desarrollo algorítmico es su opacidad. De ahí que se subraye que la lealtad y transparencia de los procesos decisionales basados en la inteligencia artificial planteen la "necesidad de desarrollar una estrategia de documentación del proyecto, que facilite lo que se ha venido a denominar 'la transparencia del algoritmo" (Martínez, 2019, p. 77). La transparencia deviene así en un principio básico sobre el que cimentar todo el desarrollo jurídico de la cuestión algorítmica (Arellano, 2019).

Por otra parte, las consecuencias adversas de una gobernanza inteligente sobrepasan los efectos directos que puedan experimentar los ciudadanos. De modo que la opacidad inherente a la inteligencia artificial genera impactos negativos para la propia administración pública, dado que esta puede desconocer el funcionamiento interno de los algoritmos que emplea. De este desconocimiento se desprende que la administración podría ni siquiera comprender los mecanismos mediante los cuales se toman las decisiones delegadas a estas tecnologías. Situación que provoca un daño evidente a derechos fundamentales de los ciudadanos, especialmente en cuanto a la incapacidad del

Estado para explicar su gestión, lo cual compromete gravemente la rendición de cuentas y, en consecuencia, resulta en la ausencia de control sobre la actuación administrativa (Cerrillo, 2019).

Es más, llegados a este punto es importante hacer notar el hecho trascendental en esta cuestión de que generalmente los algoritmos utilizados por la Administración pública provienen de empresas privadas, lo que implica que desde el inicio del proceso es el sector privado quien desarrolla el algoritmo que posteriormente tomará decisiones en el ámbito público. De este hecho surge un problema fundamental: las empresas que programan dichos algoritmos podrían tener acceso a datos que resulten esenciales para la toma de decisiones por parte de las administraciones públicas, generando así posibles riesgos en cuanto a la gestión y protección de dicha información.

En esa línea se pronuncia Valero al alertar que "la falta de transparencia puede ser un problema de singular importancia debido al relevante liderazgo del sector privado en el desarrollo de aplicaciones basadas en el uso de algoritmos" (Valero 2019, p. 89).

Así, en cuanto a la implementación de la inteligencia artificial en la administración pública, surgen otras problemáticas adicionales, entre las cuales destaca la limitada infraestructura tecnológica existente en numerosos territorios locales subdesarrollados. Esta insuficiencia dificulta la adopción de sistemas avanzados y contribuye a profundizar las brechas tecnológicas y digitales entre las grandes potencias y los países en vías de desarrollo, generando desigualdades significativas en el acceso y aprovechamiento de estas tecnologías (Campoverde Ortiz y Lujan, 2025); mientras que por otro lado, se presenta la cuestión de la resistencia institucional al cambio, que se manifiesta en la rigidez de estructuras burocráticas que dificultan la adopción de innovaciones digitales, manteniendo así modelos de gestión enfocados en procedimientos tradicionales y limitando la

capacidad de adaptación a nuevas formas de trabajo (Puyol-Cortez, 2023).

Además de ello, persiste el riesgo de ampliar las brechas digitales y sociales existentes. Siendo un fenómeno especialmente crítico en contextos con bajos niveles de conectividad, donde la falta de infraestructura tecnológica y recursos limita el acceso equitativo a los beneficios derivados de la digitalización pública. Las poblaciones rurales, los sectores con ingresos reducidos y quienes poseen menor alfabetización digital enfrentan barreras significativas que dificultan su integración en la administración digitalizada (Tejada y Balbin, 2025).

De manera concluyente afirma Dwivedi et al. (2019) que los impactos de la inteligencia artificial, en la gestión pública, contemplan tres aspectos. Primero afecta a la fuerza laboral del sector público, al delegar la toma de decisiones en sistemas de inteligencia artificial, lo que representa una amenaza clásica de sustitución laboral; segundo, impulsa una dinámica incrementada en la toma de decisiones públicas apoyadas en IA, aunque introduce elementos opacos que dificultan la auditoría por parte de no expertos; y tercero, reduce la transparencia, trazabilidad y explicabilidad de la inteligencia artificial en relación con su desempeño y accesibilidad para la población, ya que estas cualidades tienden a ser inversamente proporcionales a la complejidad de los algoritmos.

A partir del análisis realizado, se puede sostener que la inteligencia artificial y las tecnologías emergentes actúan como catalizadores esenciales del valor público, integrando la eficiencia administrativa, la transparencia en la gestión y la legitimidad democrática en la toma de decisiones estratégicas. No obstante, para que sus beneficios se consoliden, es fundamental cerrar las brechas digitales, mejorar la formación en competencias tecnológicas del personal, establecer marcos regulatorios éticos y fomentar una cultura de innovación que cuente con la

participación ciudadana. Solo bajo estas condiciones, dichas herramientas podrán impulsar efectivamente la modernización del Estado y fortalecer genuinamente la democracia.

2.5. Regulación ética de la inteligencia artificial en perspectiva comparada: Unión Europea, China y América Latina, con especial atención al caso ecuatoriano

La búsqueda de un marco jurídico coherente y funcional que regule los usos y aplicaciones de la inteligencia artificial, desde un punto de vista ético, es una preocupación latente en los contextos regionales y nacionales. Por ello la Unión Europea (2024) se propuso la elaboración de la primera ley de inteligencia artificial denominada Reglamento 2024/1689 por el que se establecen normas armonizadas en materia de inteligencia artificial, y consecuentemente ha sido adoptado en fechas recientes el Convenio Marco del Consejo de Europa sobre la Inteligencia Artificial y los Derechos Humanos, la Democracia y el Estado de Derecho de septiembre del 2024 (Framework Convention, 2024).

Por su parte, el gobierno de la República Popular China se ha involucrado activamente en la promoción y desarrollo de la inteligencia artificial, tanto en el sector público como el sector privado, y cuenta con una estrategia basada en la búsqueda de la "prosperidad común" y un Código de Ética de la Inteligencia Artificial de Nueva Generación (Ministerio de Ciencia y Tecnología de la República Popular China, 2024).

Otra visión muestra América Latina, quien aún no cuenta con un mecanismo normativo regional sobre la inteligencia artificial que englobe a todos los países de la región; a pesar de ello, países como la República Argentina, la República de Colombia y los Estados Unidos Mexicanos cuentan con estrategias nacionales en materia de inteligencia artificial, lo cual demuestra el ánimo por parte de especialistas y gobiernos sobre el tema en cuestión.

Específicamente en la República de Ecuador se han presentado algunas iniciativas en torno a la regulación de la inteligencia artificial. Sin embargo, existe consenso que hasta la fecha, la iniciativa legislativa más completa en materia de inteligencia artificial es el proyecto de "Ley Orgánica de Regulación y Promoción de la Inteligencia Artificial en Ecuador", presentado en junio de 2024. Este sigue el modelo regulatorio de la UE, basado en niveles de riesgos, y que será analizado posteriormente en este propio acápite.

También se inscriben en este país, el proyecto de "Ley de Fomento y desarrollo de la inteligencia artificial" que tiene finalidades más específicas, orientadas a apoyar el desarrollo de software de inteligencia artificial y promover la educación y formación de capital humano en esta materia. El proyecto adopta medidas concretas, como incentivos financieros y deberes para las universidades.

Por último, también esta, el proyecto de "Ley Orgánica de Aprovechamiento Digital e Inteligencia Artificial para Niñas, Niños y Adolescentes" es todavía más específico, y adopta un enfoque protector basado en los derechos fundamentales de un grupo etario específico de la población, pero que representa el futuro.

El enfoque europeo de regulación de la Inteligencia Artificial

El Reglamento de la UE, propone que los sistemas de inteligencia artificial sean seguros y respeten la legislación en materia de derechos fundamentales, deben ofrecer un marco de seguridad jurídica para potenciar la inversión y la innovación, mejorar la gobernanza y los requisitos de seguridad aplicables a estos sistemas, así como facilitar un mercado y desarrollo único, legal, fiable y seguro. Su articulado se compone, además, por una serie de objetivos y prohibiciones específicas en torno al uso y la aplicación de la inteligencia artificial. En este sentido, la

estructura de la regulación tiene una perspectiva basada en el riesgo entre los usos que crean un riesgo inaceptable, un riesgo alto y un riesgo bajo o mínimo, en lo que se conoce como sistema de semáforo (Gamero Casado, 2021).

Siguiendo la lógica de la normativa, se consideran sistemas de alto riesgo aquellos que puedan afectar negativamente a la seguridad o a los derechos fundamentales (Parlamento Europeo, 2024). Estos sistemas se clasifican en dos categorías principales. La primera incluye aquellos que se utilizan en productos regulados por la legislación de la UE sobre seguridad de los productos, como juguetes, aviación, automóviles, entre otros.

La segunda categoría abarca sistemas aplicados en ámbitos específicos que requieren especial atención debido a su impacto potencial en la sociedad. Entre ellos se encuentran: identificación biométrica y categorización de personas físicas; gestión y explotación de infraestructuras críticas; educación y formación profesional; empleo, gestión de trabajadores y acceso al autoempleo; gestión de la migración, el asilo y el control de fronteras; así como asistencia en la interpretación jurídica y en la aplicación de la ley, entre otros.

En cualquier caso, los sistemas de inteligencia artificial de alto riesgo están sujetos a obligaciones estrictas a lo largo de su ciclo de vida. Entre las que destacan la implementación de un sistema de gestión de riesgos, el sometimiento de los datos de entrenamiento de los sistemas a prácticas de gobernanza y gestión de datos, elaboración de documentación técnica actualizada antes de la introducción de estos sistemas al mercado, contra con un registro automático de acontecimientos para la detección de situaciones y la vigilancia del funcionamiento, garantizar un nivel de transparencia, supervisión humana, así como precisión, solidez y ciberseguridad durante todo el ciclo de vida de los sistemas.

Por su parte, en cuanto a los sistemas de inteligencia artificial de riesgo limitado o de uso general (Unión Europea, 2024) dentro de los que se incluyen los *chatbots* como Chat-GPT, están sujetos a requisitos de transparencia que son obligatorios, y donde los seres humanos sean informados que se encuentran interactuando con dichos sistemas, también todo contenido generado o modificado por estos sistemas debe ser etiquetado como tal. De esta manera se regula, proporcionalmente la intensidad de los límites y de los requisitos que deben cumplir los sistemas, atendiendo a los riesgos que comportan.

Otra cuestión de interés en materia de gobernanza es la estructuración de un Consejo Europeo de inteligencia artificial que cuenta con un representante de cada Estado miembro de la Unión, con lo cual se pretende facilitar una aplicación armonizada del Reglamento. Por otro lado, también se establece un sistema de vigilancia pos-comercialización consecuente a la naturaleza de los sistemas de inteligencia artificial y a los posibles daños que pueden causar los sistemas de alto riesgo, así como la información e investigación en caso de la ocurrencia de incidentes relacionados con un mal uso o funcionamiento de sistemas inteligentes.

Se trata de un aporte significativo a la estrategia de regular el uso de sistemas de inteligencia artificial a pesar de su extensión y de la complejidad técnica de su contenido (Casado, 2021). las instituciones europeas han sido capaces de elaborar un marco robusto y novedoso, que permitirá la adecuación de los sistemas a los presupuestos del Reglamento 2024/1689.

No obstante, la contribución europea a la normativa del uso de sistemas de inteligencia artificial va más allá del citado Reglamento, en tanto, el 5 de septiembre de 2024 a través de su portal oficial el Consejo Europeo hizo público la firma del Convenio Marco del Consejo de Europa sobre Inteligencia Artificial y derechos humanos, democracia y Estado de derecho (Consejo

de Europa, 2024). Las negociaciones sobre la celebración de la Convención iniciaron en septiembre del 2022, bajo el vaticinio del Comité de inteligencia artificial establecido por el Consejo de Europa en Estrasburgo, e incluyen a la Comisión Europea como representante de la UE que, además, contó con el aporte de 68 representantes internacionales de la sociedad civil, el mundo académico, la industria y otras organizaciones internacionales garantizó un enfoque integral e inclusivo.

Se trata pues, del primer instrumento internacional- estrictamente regional- jurídicamente vinculante en materia de inteligencia artificial. Compatible, de manera general, con el Derecho Europeo, y sobre todo con la Ley Europea de inteligencia artificial, que es el primer Reglamento global sobre inteligencia artificial a escala mundial (Reglamento 2024/1689). Lo que se procura es un marco efectivo a nivel internacional para abordar los riesgos que plantea la inteligencia artificial para los derechos humanos, la democracia y el Estado de derecho, a la par que se apuesta por la innovación y la confianza en estos sistemas.

Así mismo, se incluyen una serie de preceptos relacionados con un enfoque de la inteligencia artificial basado en el ser humano, que respete los derechos humanos, la democracia y el Estado de derecho y que a la vez se base en el riesgo que puede generar su uso. Encierra un conjunto de principios para garantizar una inteligencia artificial fiable, a partir del establecimiento de obligaciones en cuanto a la rendición de cuentas y la documentación, en la gestión de riesgo y mecanismos que permitan la supervisión humana en las diferentes actividades.

El propósito de esta Convención es asegurar que las actividades contenidas dentro del ciclo de vida de los sistemas de inteligencia artificial sean compatibles con los derechos humanos, la democracia y el Estado de derecho. Para ello se adoptarán desde el plano legislativo, administrativo medidas encaminadas al cumplimento de ese objetivo. Dichas medidas, se diferenciarán,

siempre que sea necesario, de acuerdo a las probabilidades de que ocurran impactos negativos en cada una de estos tres ámbitos por parte de los sistemas de inteligencia artificial.

Dentro de los principios fundamentales relacionados con las actividades realizadas por los sistemas de inteligencia artificial se encuentran la dignidad humana y la autonomía individual, la inspección y la transparencia, responsabilidad, igualdad y no discriminación, incluida la igualdad de género, privacidad y protección de datos personales, confianza en los sistemas de inteligencia artificial y la innovación segura.

La Convención establece también remedios para aquellas personas que puedan ver perjudicados sus derechos humanos a partir de informaciones relevantes que son captadas o generadas por medio de sistemas inteligentes, en estos casos se les brinda la posibilidad de reclamar ante dichas decisiones tomadas por estos sistemas, incluso de formular una queja ante las autoridades competentes. Concede salvaguardas procedimentales que garantizan los derechos de quienes han visto perjudicado el disfrute de sus libertades fundamentales por medio de sistemas de inteligencia artificial, y que en todo caso la persona debe ser notificada que se encuentra interactuando con esta tecnología y no con otro ser humano.

El manejo de los riesgos que supone el uso de sistemas de inteligencia artificial, en el marco de la Convención, parte de la prevención y mitigación, para tomar medidas diferenciales y graduales en torno al contexto, la severidad del impacto, el monitoreo de los riesgos, informar y documentar sobre los impactos negativos que pueden tener los sistemas de inteligencia artificial en los derechos humanos, la democracia y el Estado de derecho. Así como la posibilidad por parte de las autoridades para poner en moratoria o proscribir usos de sistemas contarios a los derechos humanos y al derecho internacional.

La Convención 225 del Consejo de Europa (2024) se implementa en un marco de no discriminación, respeto de los derechos de las personas en situación de discapacidad y de los niños, consultas públicas tomando en consideración las implicaciones de la inteligencia artificial en ámbitos sociales, económicos, ambientales, legales y éticos, promover la alfabetización digital en los distintos sectores poblacionales para así ampliar el margen de identificación de los riesgos en cuanto a los usos de la inteligencia artificial, salvaguardas para la existencia de los derechos humanos a través de una amplia protección de los mismos, a lo largo de todo su articulado. Por otra parte, cubre las aplicaciones de inteligencia artificial que se realizan tanto por las autoridades públicas como por el sector privado. En el caso de estos últimos pueden directamente someterse al cumplimento de las obligaciones y principios del Convenio o bien tomar un de conjunto medidas encaminadas a asegurar el respeto a los derechos humanos, la democracia y el Estado de derecho.

En este orden, la secretaria general del Consejo de Europa, Marija Pejčinović, ha declarado: "debemos garantizar que el auge de la inteligencia artificial sostenga nuestras normas, en lugar de socavarlas. El Convenio Marco está diseñado para garantizar justo eso. Es un texto potente y equilibrado, resultado del enfoque abierto e inclusivo con el que ha sido redactado, lo que ha garantizado que se beneficie de perspectivas múltiples y expertas … es un tratado abierto con un potencial alcance mundial".

Por tanto, el Convenio abarca los sistemas de inteligencia artificial que puedan interferir en los derechos humanos, la democracia y el Estado de Derecho antes mencionados, siguiendo un enfoque diferenciado y basado en el riesgo. Y una característica peculiar del mismo es ser tecnológicamente neutral, o sea que no hace mención a una tecnología específica de la inteligencia artificial. Incluye, además, exenciones para la investigación y el desarrollo, así como para la seguridad nacional. Se trata pues, de una

norma equilibrada e inclusiva, en cuya redacción se tuvieron en cuenta la diversidad de criterios que existen alrededor del uso de sistemas de inteligencia artificial. Además de ser el primero de su tipo jurídicamente vinculante.

La República Popular China y su estrategia nacional de Inteligencia Artificial

El primer antecedente de China en la incursión de regular la inteligencia artificial fue el Plan de desarrollo de inteligencia artificial de nueva generación (Graham Webster et al., 2017). Esta fue la primera estrategia sobre inteligencia artificial a nivel nacional y reconocía su importancia clave para la ciencia teórica y aplicada, así como los avances y las limitaciones del país en la materia.

La potencia asiática cuenta con una serie de objetivos estratégicos en IA, que buscan establecer a China como una nación innovadora en inteligencia artificial y como una potencia mundial en ciencia y tecnología, gozando de las ventajas de ser el primero en establecerse y consolidarse en esta arena. Dentro de estos objetivos se encuentran el alcance del nivel global de competitividad en IA; progreso en el desarrollo de la tecnología base; fomento de aplicaciones innovadoras en el sector y la formación de profesionales especializados; establecimiento de empresas en inteligencia artificial referentes a nivel global.

Así mismo, en este país, para el período comprendido entre 2025-2030 una nueva generación de sistemas tecnológicos y teorías en el sector; nuevos hitos en la aplicación de la inteligencia artificial en sectores como: fabricación, medicina, *smart cities*, agricultura y defensa; y la entrada en vigor de leyes, políticas, regulaciones y normas éticas relacionadas con la inteligencia artificial. Además, posicionarse como líder mundial en la aplicación y el desarrollo de la industria de inteligencia artificial; en la generación de un mercado maduro en el sector con soluciones para todas las

industrias y el perfeccionamiento de las regulaciones derivadas de su aplicación en la industria (Salado García, 2018).

Para la consecución de estos objetivos el Comité Profesional Nacional de Gobernanza de la Inteligencia Artificial publicó en septiembre de 2021 el "Código Ético de la Inteligencia Artificial de Nueva Generación" (Ministerio de Ciencia y Tecnología de la República Popular China). El Código regula las actividades de las personas físicas y jurídicas dedicadas a la gestión, desarrollo y uso de la inteligencia artificial, para la cual propone una serie de normas éticas, de obligatorio cumplimiento.

Dentro de estas normas destacan: mejorar el bienestar humano, velar por los derechos humanos y los intereses fundamentales de la humanidad, así como el respeto por la ética nacional y regional. Promover la armonía entre humanos-máquinas, y a su vez el desarrollo económico, social y ecológico sostenible. Garantizar, en todo momento, que los sistemas de inteligencia artificial actúen bajo el control humano, y que sean equitativos y justos. Además, proteger la privacidad y la seguridad de las personas durante todo el ciclo de vida de los sistemas inteligentes.

Con esta estrategia nacional de inteligencia artificial el gobierno chino, por tanto, propone una política basada en parámetros éticos que son claros, dotándolos de carácter obligatorio, donde la población y su bienestar común se ubican como eje central de la misma.

Estrategias nacionales en América Latina para la regulación del uso de sistemas de inteligencia artificial: los casos de Argentina, Colombia, México y Ecuador

América Latina y el Caribe es una de las regiones con desigualdades económicas y sociales más marcadas en el mundo. Ello se evidencia, además, en la dependencia tecnológica, debido a la carencia de capacidades propias para el impulso de esta rama;

contradicción muy paradójica sobre todo si se tiene en cuenta que la región es una de las mayores exportadoras de materias primas (Banco Interamericano de Desarrollo. 2024) cruciales empleadas en la industria. Pese a que el desarrollo de políticas y marcos regulatorios ha sido discreto en materia de inteligencia artificial, en los últimos años la región ha mostrado avances en la generación de estrategias y políticas nacionales de la materia, que tomaron como base los principios emitidos por organismos internacionales como la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura y la Organización para la Cooperación y el Desarrollo Económico (OCDE).

Actualmente países como Argentina, Brasil, Chile, Colombia, México, Perú y Uruguay muestran un avance considerable en la regulación de la inteligencia artificial, con respecto a otros países de la región (Comisión Económica para América Latina y el Caribe, 2024), puesto que cuentan con estrategias nacionales propias, que partiendo de un conjunto de acciones y ejes estratégicos pretenden transformar el sector público, desde una perspectiva ética y de derechos humanos.

En el caso de Argentina, una de las primeras iniciativas fue el Plan Nacional de inteligencia artificial (ArgenIA),162 elaborado entre 2018-2019. Se trata de un informe descriptivo; y se centró entre otros puntos en la necesidad de formar recursos humanos, la importancia del uso de los datos, la infraestructura computacional y la ética y las regulaciones. En este sentido, el informe no logró distinguir entre las regulaciones sobre su uso y los problemas éticos que puede generar.

En junio de 2023 la Subsecretaría de Tecnologías de la Información aprobó las “Recomendaciones para una Inteligencia Artificial Fiable” (RIAF), que a través de un conjunto de principios éticos buscan el respeto, en todo momento, a los datos a los que se acceden durante todo el ciclo de vida de los sistemas de inteligencia artificial. Para ello se tomaron en consideración

entre otros documentos, la Recomendación sobre ética de la inteligencia artificial de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. Reúne principios como la supervisión y la responsabilidad humana, dejando de un lado problemas complejos relacionados sobre todo con la autonomización de estos sistemas, y va solo orientada a este tipo de procesos en el Estado.

El gobierno argentino también cuenta con el "Programa de Transparencia y Protección de Datos Personales en el uso de la Inteligencia Artificial" cuyo objetivo es fortalecer las capacidades estatales para el desarrollo y uso de la inteligencia artificial. Hasta la fecha el Programa solo actúa para la observación y el diagnóstico, sin aportar soluciones para el uso ilegal de datos personales.

En otro orden, Colombia cuenta con una "Hoja de Ruta para el Desarrollo y Aplicación de la Inteligencia Artificial" como muestra del compromiso de su gobierno de alinear los principios éticos y la sostenibilidad con el desarrollo y aplicación de la inteligencia artificial; y que incluye al sector estatal, el académico, el privado y el financiero, así como a la sociedad civil.

La Hoja de Ruta busca establecer un enfoque multifacético de la inteligencia artificial y para ello propone cinco entornos estratégicos de trabajo. El primero de ellos, ética y gobernanza, promueve la transparencia en los algoritmos, fomentando la participación ciudadana en la toma de decisiones y el establecimiento de principios éticos y marcos normativos para guiar el desarrollo y la implementación de la inteligencia artificial en el país.

En este sentido, es importante destacar que el país sudamericano cuenta con un "Marco Ético para la Inteligencia Artificial", el cual proporciona una serie recomendaciones específicas destinadas al sector público dentro de las que se incluyen: la transparencia y la explicabilidad, no discriminación, seguridad, responsabilidad, beneficio social y prevalencia de los derechos de niños, niñas y

adolescentes. Cada una de estas recomendaciones se abordan desde tres aspectos, los datos, los algoritmos y la práctica.

La Hoja de Ruta de la inteligencia artificial colombiana no solo promueve la integridad científica en el desarrollo de soluciones tecnológicas, sino que también busca establecer estructuras de gobernanza efectivas que aseguren su aplicación ética de la inteligencia artificial en la resolución de desafíos sociales, económicos y ambientales a nivel nacional. Ello evidencia la necesidad de que el Estado reconozca los desafíos éticos que representa su uso, lo cual implica diseñar un marco regulador que guíe el proceso de diseño, desarrollo e implementación de los sistemas inteligentes, desde un enfoque ético.

Estas iniciativas se alinean con los objetivos y metas del Plan Nacional de Desarrollo 2022-2026: "Colombia Potencia Mundial de la Vida" y la Política de Investigación orientada por Misiones del Ministerio de Ciencia, Tecnología e Innovación.

Por su parte México, no cuenta con una regulación en materia de inteligencia artificial, sin embargo, desde el 2020 ha habido 58 iniciativas legislativas (Meza Ruiz, 2024) que hacen alusión a la misma. La más recientes de estas propuestas fue presentada por la Alianza Nacional de Inteligencia Artificial (ANIA), bajo el nombre de Agenda Nacional de Inteligencia Artificial para México 2024-2030. La Agenda tiene por objeto establecer un marco referencial que suscite la integración de la inteligencia artificial para fortalecer el desarrollo económico, social, educativo y tecnológico mexicano; bajo una visión ética, responsable y equitativa en aras de proteger los derechos humanos y ambientales.

Desde este enfoque establece una serie de recomendaciones multidisciplinarias que incluyen políticas públicas, derechos, educación, mercados laborales, ciberseguridad y gestión de riesgos, género e inclusión, infraestructura y datos, e innovación. En el ámbito regulatorio, se proponen medidas para proteger la privacidad, garantizar la equidad y la transparencia en el uso

de inteligencia artificial, y asegurar una gobernanza adecuada. Además, se proponen mecanismos de gobernanza prácticos y democráticos, encadenando el sector privado, la academia y la sociedad civil.

La Agenda muestra la necesidad de una regulación basada en riesgos, neutralidad tecnológica, competencia económica, privacidad, seguridad, transparencia y libertad de expresión, para lo cual propone una serie de mecanismos y regulaciones en materia de inteligencia artificial tales como la Creación de la Ley de Datos Abiertos Públicos y de una Ley de Ciberseguridad, así como promover una gestión ética y responsable en el desarrollo, uso, aplicación e implementación de las tecnologías basadas en IA.

De manera general se han establecido una serie de mecanismos de regulación que consideran las implicaciones éticas que comprende este desarrollo tecnológico. Las políticas estatales incluyen un enfoque ético, a partir, de ejes temáticos y acciones estratégicas con la finalidad de garantizar que la implementación de la inteligencia artificial sea de acuerdo a principios éticos, en todo caso.

Por su parte, en la Republica de Ecuador la normativa sobre tecnologías digitales está contenida, principalmente en la iniciativa, que confiere al Ministerio de Telecomunicaciones y de la Sociedad de la Información (MINTEL) la rectoría del sector de telecomunicaciones y de la sociedad de la información. Complementariamente, la Ley Orgánica para la Transformación Digital y Audiovisual (LOTDA) y su reglamento, aprobada en 2023, incluye en su articulado a las tecnologías emergentes, entre las que se encuentra la inteligencia artificial.

El uso de la inteligencia artificial en Ecuador plantea desafíos en materia de derechos humanos tale como la privacidad, la igualdad y la seguridad jurídica. Si bien el país no cuenta aún con una legislación específica sobre IA, su marco normativo vigente

establece principios y protecciones aplicables a los posibles riesgos que estas tecnologías pueden generar.

Tal es así que hasta día de hoy Ecuador no cuenta, con una política nacional específica sobre inteligencia artificial. Sin embargo, se han realizado esfuerzos para su construcción. El MINTEL conformó en 2024 el Comité de Inteligencia Artificial, creado bajo la figura de Comités Consultivos establecida en la LOTDA y regulada por el reglamento.

Por su parte, la inteligencia artificial se reconoce como parte de la Política Pública de Transformación Digital del Ecuador 2025–2030, que la identifica como una tecnología emergente clave para el desarrollo sostenible.

En dicha política, uno de los ejes se enfoca en las "tecnologías emergentes para el desarrollo sostenible", proponiendo como estrategia el desarrollo y aprobación de un marco normativo flexible y eficiente, que permita la experimentación segura, reduzca la incertidumbre y facilite el crecimiento tecnológico. Esta política se construyó a partir de 576 propuestas, de las cuales se priorizaron 160 alternativas de solución (Ecuador. Ministerio de Telecomunicaciones y de la Sociedad de la Información, 2025), lo que llama la atención en establecer un marco regulatorio para la inteligencia artificial.

Actualmente, Ecuador cuenta con una propuesta Proyecto de Política para el Desarrollo y Fomento del Uso Ético de la Inteligencia Artificial, elaborado por dicho Comité. Otro aspecto a considerar es que la regulación de la inteligencia artificial no está vinculada únicamente por la normativa ligada a telecomunicaciones, sino también por normativas de carácter general como la Ley Orgánica de Protección de Datos Personas (LOPDP), el Código de la Economía Social de los Conocimientos, Creatividad e Innovación (COECSCI), el Código Orgánico Integral Penal (COIP) y otras.

También es clave declara que hasta lo aquí descrito se suma la ausencia de una política pública vigente y consolidada desde el gobierno, debido tanto a la inestabilidad política de los últimos años como a la rotación de funcionarios en entidades clave. Por otro lado, el poder legislativo ha impulsado diversas propuestas normativas pero que aún no se han aprobado. Asimismo, se evidencia la falta de mecanismos participativos amplios, lo cual restringe la inclusión de la academia, la sociedad civil, cuya experiencia y conocimiento resultan cruciales para la aprobación de un marco legal sistémico, orgánico y moderno.

CAPÍTULO 3

Marco jurídico y regulatorio de la inteligencia artificial. Desafíos legales

3.1. Análisis del marco jurídico internacional y regional vigente sobre inteligencia artificial

Las tecnologías de inteligencia artificial, con su amplia capacidad e impacto disruptivo, han permeado la sociedad de hoy de manera generalizada. Ello ha impuesto lidiar con numerosos desafíos derivados de su uso que están obligando ya a los Estados a crear un marco legal unificado para su regulación. El fenómeno de la inteligencia artificial requiere actuar en tiempo real (Nodals, 2021) es por ello que las iniciativas políticas y legales de regulación que se han aprobado tengan por objetivo minimizar los riegos que supone el uso de esta tecnología y maximizar sus beneficios en aras de una sociedad más justa e inclusiva.

En este sentido, las Naciones Unidas destacan en su abordaje institucional a partir del desarrollo de un conjunto de acciones (Drnas de Clément, 2022) que han sentado las bases del tratamiento regulatorio internacional que se exhibe hoy. La Declaración

de Principios de la Cumbre Mundial sobre la Sociedad de la Información celebrada entre 2003 y 2005 y el Mecanismo de Facilitación de Tecnología de las Naciones Unidas creado en 2015 son precursores de la postura seguida en la actualidad por la Organización de las Naciones Unidas en materia de inteligencia artificial pues en ambas herramientas se promueve el uso responsable y seguro de las nuevas tecnologías.

En junio de 2019 el Panel de Alto Nivel sobre Cooperación Digital de la Secretaría General de la ONU presentó su informe titulado "La era de la interdependencia digital" marcando los primeros pasos de avance en cuanto a políticas y normativización de la inteligencia artificial a nivel internacional. En estrecha relación con este informe de la Secretaría de la ONU, el Consejo de Derechos Humanos ha promulgado varias Resoluciones sobre "Las tecnologías digitales nuevas y emergentes y los derechos humanos" con el propósito de que los Estados sitúen los derechos humanos en el centro de los marcos reguladores relativos al desarrollo y el uso de las tecnologías digitales.

En una dirección analítica similar con respecto al fenómeno de la inteligencia artificial se encamina la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) centrando su programa en los asuntos relacionados con la ética, es decir, los aspectos morales y los impactos que estas tecnologías tienen en las personas, las políticas y la sociedad en su conjunto. El resultado fue la publicación en noviembre de 2021 de su Recomendación sobre la Ética de la Inteligencia Artificial cuyo objeto se centra en afrontar responsablemente los efectos de la inteligencia artificial en los seres humanos, las sociedades y el medio ambiente.

La Recomendación pretende ser un instrumento normativo de aceptación mundial, para lo cual articula una serie de diez principios éticos y valores. En este sentido, resulta importante destacar que lo que hace verdaderamente aplicable a esta

Recomendación es la posibilidad de traducir esos valores y principios en acciones concretas encaminadas a trazar políticas y normativas respecto a los sistemas de inteligencia artificial. A pesar de ello, no escapa de la esfera soft law puesto que se trata de un instrumento que no es legalmente vinculante y que, sobre todo, otorga un papel preponderante a la buena fe de los Estados para que lleven a sus normativas nacionales acuerdos relativos a la ética de la inteligencia artificial.

Las Naciones Unidas en su labor de promover un uso "seguro y fiable" de la inteligencia artificial, aprobaron el 21 de marzo de 2024 la resolución A/RES/78/265 "Aprovechar las oportunidades de sistemas seguros y fiables de inteligencia artificial para el desarrollo sostenible" con el objetivo de "promover la transformación digital y el acceso equitativo a los beneficios de estos sistemas". Por lo cual hace un llamado a su gobernanza, destacando que deben tratarse de sistemas que sean interoperables, ágiles, adaptables e inclusivos y respondan a las distintas necesidades y capacidades de los países desarrollados y los países en desarrollo por igual.

Enunciando algunos de sus principales iniciativas en estas esferas se encuentran resoluciones 77/320, de 25 de julio de 2023, relativa al impacto del cambio tecnológico rápido en la consecución de los Objetivos de Desarrollo Sostenible y sus metas, 78/132, de 19 de diciembre de 2023, relativa a las tecnologías de la información y las comunicaciones para el desarrollo sostenible, 78/160, de 19 de diciembre de 2023, relativa a la ciencia, la tecnología y la innovación para el desarrollo sostenible, 78/213, de 19 de diciembre de 2023, relativa a la promoción y la protección de los derechos humanos en el contexto de las tecnologías digitales, 77/211, de 15 de diciembre de 2022, relativa al derecho a la privacidad en la era digital; así como el establecimiento de la Oficina del Enviado del Secretario General para la Tecnología para coordinar su aplicación y del establecimiento por el Secretario

General de un Órgano Asesor de Alto Nivel sobre Inteligencia Artificial integrado por múltiples interesados y de su informe provisional publicado el 21 de diciembre de 2023.

La resolución también busca promover la elaboración y la aplicación de enfoques y marcos regulatorios y de gobernanza nacionales para apoyar la inversión responsable en inteligencia artificial para el desarrollo sostenible. Alentar medidas eficaces para la prevención y mitigación a nivel internacional de los riesgos durante el diseño, desarrollo y despliegue de sistemas de inteligencia artificial. Así como, el establecimiento de mecanismos de retroalimentación que permitan la notificación por parte de los usuarios y terceros de las vulnerabilidades y usos indebidos, incluyendo los incidentes relacionados con la inteligencia artificial.

Mientras tanto, en el espacio regional se destaca la Unión Europea (2024) al lanzarse a la elaboración de la primera ley de inteligencia artificial denominada Reglamento (UE) 2024/1689 por el que se establecen normas armonizadas en materia de inteligencia artificial, y consecuentemente ha sido adoptado en fechas recientes el Convenio Marco del Consejo de Europa sobre la Inteligencia Artificial y los Derechos Humanos, la Democracia y el Estado de Derecho de septiembre del 2024.

Por otro lado en la región asiática, el gobierno de la República Popular China se ha involucrado activamente en la promoción y desarrollo de la inteligencia artificial, tanto en el sector público como el sector privado, y cuenta con una estrategia basada en la búsqueda de la “prosperidad común” y un Código de Ética de la Inteligencia Artificial de Nueva Generación (Ministerio de Ciencia y Tecnología de la República Popular China 2021), que regula las actividades de las personas físicas y jurídicas dedicadas a la gestión, desarrollo y uso de la inteligencia artificial, para la cual propone una serie de normas éticas, de obligatorio cumplimiento.

Aunque en menor medida que el continente euroasiático, en Latinoamérica también se está apostando por la regulación

especial en temas de inteligencia artificial. En ese sentido países como Argentina, Brasil, Chile, Colombia, México, Perú y Uruguay (Agencia para el Gobierno Electrónico y la Sociedad de la Información y el Conocimiento de Uruguay, 2024; Subsecretaría de Tecnologías de Información, 2023; Ministério da Ciência, Tecnologia e Inovações Secretaria de Empreendedorismo e Inovação, 2021; Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, 2024) muestran un avance considerable en la regulación de la inteligencia artificial, con respecto a otros países de la región (Comisión Económica para América Latina y el Caribe, 2024), puesto que cuentan con estrategias nacionales propias.

En el caso de Argentina, una de las primeras iniciativas fue el Plan Nacional de de inteligencia artificial (ArgenIA). Un informe descriptivo que se centró en la necesidad de formar recursos humanos, la importancia del uso de los datos, la infraestructura computacional y la ética y las regulaciones. Además, en junio de 2023 la Subsecretaría de Tecnologías de la Información aprobó las "Recomendaciones para una Inteligencia Artificial Fiable" (RIAF) (Subsecretaría de Tecnologías de Información, 2023). Y cuenta también con el "Programa de Transparencia y Protección de Datos Personales en el uso de la Inteligencia Artificial".

En otro orden, Colombia cuenta con una "Hoja de Ruta para el Desarrollo y Aplicación de la Inteligencia Artificial" (Colombia. Ministerio de Ciencia, Tecnología e Innovación, 2024) como muestra del compromiso del gobierno de alinear los principios éticos y la sostenibilidad con el desarrollo y aplicación de la inteligencia artificial. También se propone desde Colombia la creación de un Registro Nacional de Inteligencia Artificial, "previo cumplimiento de los requisitos técnicos y jurídicos".

Por su parte México, no cuenta con una regulación en materia de inteligencia artificial. Sin embargo, desde el 2020 ha habido 58 iniciativas legislativas (Meza, 2024) que hacen alusión a la

misma. La más reciente de estas propuestas fue presentada por la Alianza Nacional de Inteligencia Artificial (ANIA) bajo el nombre de Agenda Nacional de Inteligencia Artificial para México 2024-2030, que en el ámbito regulatorio propone medidas para proteger la privacidad, garantizar la equidad y la transparencia en el uso de inteligencia artificial, y asegurar una gobernanza adecuada.

Cuba, por su parte, implementa desde mayo de 2024 una política para la transformación digital que concibe la transformación digital como una etapa superior a la informatización, centrada en el bienestar de las personas y orientada a la prosperidad y sostenibilidad del país. Para lo cual se estructura en ocho ejes estratégicos, que incluyen el marco normativo, la infraestructura tecnológica, la economía digital, la educación y cultura digital, el gobierno digital, la innovación, la ciberseguridad y los contenidos digitales.

El diseño de esta política tiene como objetivo ordenar y garantizar el derecho al acceso y participación de las personas naturales y jurídicas en el proceso de informatización, permitiendo al país asegurar la modernización, la ciberseguridad frente a las amenazas y riesgos, la sostenibilidad y soberanía tecnológica y la innovación con la elaboración y comercialización de productos y servicios asociados.

Unida a ella ordenó el país además la Estrategia para el Desarrollo de Software que viene incluida como parte del Decreto-Ley 370 de 2018, el cual plantea su Título II dedicado en específico a reglamentar aspectos propios de la Industria Cubana de Programas y Aplicaciones Informáticas. Es este Decreto-Ley precursor de una elaboración normativa subsiguiente, cuyo objeto es la reglamentación y garantía del desarrollo de dicha industria, a saber, el Decreto 359 de 2019 Sobre el Desarrollo de la Industria Cubana de Programas y Aplicaciones Informáticas, y el Decreto 45 del 2021 del Consejo de Ministros “Sobre el desarrollo integral de la automatización en Cuba”.

En ese marco se aprobó la Estrategia cubana para el Desarrollo de la inteligencia artificial (Consejo de Ministros de la República de Cuba, 2024) que contempla seis ejes centrales referidos a ética y marco normativo, capital humano, aplicaciones y servicios, administración pública, ciencia e innovación y comunicación social.

En tanto Ecuador, por otro lado, aunque todavía no registra una estrategia de inteligencia artificial, sí exhibe una voluntad legislativa enfocada hacia el logro de un marco normativo robusto orientado a la regulación de la inteligencia artificial. En ese sentido destacan su "Proyecto de Ley Orgánica de Regulación y Promoción de la inteligencia artificial en Ecuador" de junio de 2024, su "Proyecto de Ley de Fomento y desarrollo de la inteligencia artificial" de julio de 2024 y su "Proyecto de Ley Orgánica de Aprovechamiento Digital e Inteligencia para niñas, niños y adolescentes" de septiembre de 2024; así como un proyecto de política para el desarrollo y fomento de su uso ético y responsable en el Ecuador; propuesto en noviembre de 2023 (Zambrano et al., 2025).

La primera ley en proyecto tiene como alcance regular todas aquellas actividades de investigación, desarrollo, despliegue y comercialización y uso de sistemas de inteligencia artificial por parte de entidades públicas o privadas, nacionales o extranjeras, cuyo tratamiento se efectúe en territorio ecuatoriano o que generen efectos jurídicos sobre personas naturales o jurídicas domiciliadas en el país (Ecuador. Asamblea Nacional, 2024). Así, la segunda se enfoca en identificar áreas de alto riesgo para el uso de la inteligencia artificial. En tanto la tercera pretende garantizar el uso seguro de internet, redes sociales y plataformas digitales para niñas, niños y adolescentes, así como regular el desarrollo, introducción y uso de sistemas de inteligencia artificial en relación con la niñez y adolescencia (Guzmán et al., 2025).

Por su parte su proyecto de política asigna responsabilidades para emitir políticas, promover investigaciones y fomentar la modernización de procesos tecnológicos; el desarrollo económico y social para incentivar la creación de empresas de inteligencia artificial, fomentar *startups*, alianzas con el sector privado, y habilitar recursos de computación en la nube para su desarrollo (Naranjo Godoy, 2024).

3.2. Derechos digitales e inteligencia artificial

Además del espacio físico conocido, en la actualidad se ha consolidado un nuevo ámbito de interacción social llamado indistintamente como ciberespacio, mundo digital, espacio en línea o entorno virtual. Este espacio ha sido definido por Santana-Soriano y Báez Vizcaíno (2024) como "un espacio artificial y ficcional emergente en el que tienen lugar relaciones sociales entre personas, organizaciones y máquinas, y que no tiene existencia independiente del conjunto de equipos y programas informáticos que le posibilitan" (p. 49). Es decir, este ámbito excede los límites del internet para configurarse, en palabras de los citados autores, en "un ente conceptual" que incluye, no solo, *hardware, software* y sistemas de intercambio de datos, sino también una interacción social compleja entre personas y organizaciones.

En el entorno digital actual, la protección de los derechos humanos adquiere una importancia singular y urgente, dado que el ciberespacio se ha convertido en un escenario fundamental para la interacción social, el acceso a la información y la comunicación. Diversas organizaciones internacionales y la comunidad global han afirmado que los derechos que se reconocen y garantizan en el mundo físico deben igualmente aplicarse y respetarse en el espacio digital. Esto implica un enfoque especial en la protección de la libertad de expresión, el acceso equitativo a la información, la privacidad y la salvaguardia de los datos personales, asegurando que los derechos fundamentales no se vean erosionados al trasladarse al ámbito virtual.

En ese espacio surge el concepto de derechos digitales que engloba los derechos de los ciudadanos en el entorno digital o ciberespacio, ya sean derechos fundamentales o derechos ordinarios (Barrio Andrés, 2021). Se puede plantear, en consecuencia, que esta corresponde a una cuarta generación de derechos, la cual se enfoca en los desafíos actuales derivados de áreas como la informática, la biomedicina, la tecnología, la genética y la identidad, así como en nuevas situaciones que emergen a partir del avance y desarrollo científico y tecnológico, que responde a la necesidad de adaptarse a los cambios sociales y tecnológicos contemporáneos.

La finalidad de esta nueva generación de derechos radica en abordar la carencia de una regulación jurídica adecuada que acompañe la revolución digital. Pues ello ha facilitado la proliferación de monopolios que afectan negativamente la competencia en el mercado de ese tipo. Asimismo, ha contribuido al aumento cualitativo y cuantitativo de las desigualdades, la instauración de mecanismos de vigilancia en tiempo real sobre las conductas de los individuos en el ciberespacio, y, especialmente, la progresiva sustitución de la identidad real por la imagen proyectada a través de la huella digital y las correlaciones generadas por la inteligencia artificial. En suma, el uso inadecuado de esta tecnología, ya sea intencional o no, representa un riesgo para los derechos fundamentales de las personas.

Entre los más vulnerables al impacto de la inteligencia artificial se encuentran la privacidad, cuya protección se ve amenazada por la recopilación masiva y el análisis de datos personales sin un consentimiento pleno o informado, y la igualdad, afectada por sesgos algorítmicos que pueden generar discriminación en ámbitos como la selección laboral o el acceso a servicios digitales. La opacidad, inherente a muchos sistemas de inteligencia artificial, por otro lado, dificulta la transparencia y rendición de cuentas, lo que puede aumentar la desconfianza y limitar la capacidad de los usuarios para controlar su información (Grigore, 2022).

Destacan además los derechos a la dignidad, a la libertad de expresión, a la protección de datos personales, al interés superior de niñas, niños y adolescentes, así como los derechos del autor sobre su creación.

A este respecto, según el Alto Comisionado de las Naciones Unidas para los Derechos Humanos, el uso masivo de datos para alimentar algoritmos genera perfiles detallados que pueden perpetuar sesgos discriminatorios y promover la difusión de información errónea, afectando la reputación y el respeto hacia los individuos. Al mismo tiempo la inteligencia artificial puede imponer una vigilancia constante que erosiona la autonomía personal, mientras la falta de transparencia en su funcionamiento limita la capacidad de las personas para comprender y controlar el uso de su información (Organización de las Naciones Unidas, 2019). Condiciones que contribuyen a una pérdida de control sobre la identidad digital y a situaciones donde las personas son tratadas como meros datos, lo que representa un desafío ético crítico para la preservación de la dignidad humana en el entorno digital.

Así, al facilitar la inteligencia artificial procesos automatizados que pueden restringir, censurar o manipular la visibilidad de determinados contenidos sin suficiente transparencia ni rendición de cuentas esta influye negativamente en el derecho a la libertad de expresión en el espacio digital. Según informes de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (2023), la implementación de algoritmos para la moderación de contenido en plataformas digitales, aunque busca combatir la desinformación y el discurso de odio, también puede derivar en la eliminación injustificada de expresiones legítimas, afectando así la pluralidad y diversidad de opiniones en línea.

Esta situación genera un desequilibrio en el ecosistema digital, en donde las decisiones tomadas por sistemas de inteligencia artificial

relevantes para la expresión pública carecen de supervisión adecuada y errores de sesgo pueden perpetuar injusticias. A su vez, la concentración de poder en pocos actores del sector tecnológico y la opacidad de sus sistemas algorítmicos dificultan el control democrático y la participación inclusiva en la toma de decisiones sobre cómo se regula y modera la información.

Por otra parte, la mencionada tecnología facilita, como ya se ha planteado en este texto, la recopilación masiva, automatizada y rápida de grandes volúmenes de información personal, muchas veces sin el consentimiento explícito o informado de los titulares, lo que contradice principios básicos de privacidad y control sobre los datos. Además, la inteligencia artificial generativa puede concebir perfiles detallados y personalizados que influencian decisiones y conductas, aumentando riesgos de manipulación y discriminación (Europagune Research Group, 2023).

La falta de marcos jurídicos especializados y actualizados para regular estos avances tecnológicos agrava la situación, dejando espacio para usos indebidos o ilegales del tratamiento de datos. Casos como la manipulación mediante *deepfakes* y la vigilancia constante en plataformas digitales evidencian consecuencias tangibles en la vida social, laboral y personal de las personas, subrayando la necesidad urgente de fortalecer normativas que garanticen la transparencia, responsabilidad y protección efectiva de los derechos digitales de privacidad y confidencialidad en la era de la inteligencia artificial.

De la misma forma, el interés superior de niñas, niños y adolescentes en el entorno digital se ve frecuentemente vulnerado debido a una serie de riesgos inherentes a la naturaleza misma de ese espacio. La exposición temprana y constante a contenidos inapropiados, asociada a la violencia digital, el ciberacoso y la explotación sexual son amenazas significativas que impactan su desarrollo emocional, psicológico y físico. Además, la privacidad de los menores está en constante riesgo debido a prácticas como

la recogida indiscriminada de datos, la suplantación de identidad y la creación de perfiles digitales que pueden ser utilizados sin consentimiento, afectando la integridad y dignidad de este grupo (Organización de las Naciones Unidas, 2023).

Ante esa situación, la responsabilidad parental y la regulación estatal se perfilan como factores esenciales para garantizar entornos digitales seguros y libres de violencia, sin embargo, es necesario fortalecer marcos legales y mecanismos de protección que contemplen la autonomía progresiva de los menores, su derecho a la imagen y a la confidencialidad de sus datos personales. Estos elementos son cruciales para preservar el bienestar y garantizar un acceso equitativo y protegido a las oportunidades que ofrece el mundo digital, salvaguardando siempre su interés superior.

Finalmente, debido al uso masivo de inteligencia artificial, con la generación, en primer lugar, de obras a partir de datos protegidos sin la autorización de los titulares, los derechos asociados al autor en el entorno digital también sufren una serie de vulneraciones. La capacidad de la inteligencia artificial para procesar y replicar contenidos existentes ha diluido las fronteras tradicionales del concepto de autoría, complicando la atribución y protección legal de las creaciones (Wigwe & Partners, 2025). Por ello, aun cuando algunos tribunales han reconocido ciertos usos como transformativos o bajo el concepto de uso justo (Skadden, 2025), persisten preocupaciones sobre el impacto que estos modelos pueden tener en el mercado creativo, especialmente por la posibilidad de que herramientas generativas produzcan obras en competencia directa con los creadores humanos.

Por último en este análisis, es pertinente apuntar que en función del abordaje de las preocupaciones relacionadas con la inteligencia artificial y los derechos humanos, por extensión, también los derechos digitales, un conjunto de organizaciones de derechos humanos y tecnología emitieron una declaración histórica que

propone estándares de derechos humanos para el aprendizaje automático, presentando la Declaración de Toronto sobre la protección del derecho a la igualdad y la no discriminación en los sistemas de aprendizaje automático (Amnesty International, 2018).

En ella se insta a gobiernos y empresas a asegurar que las aplicaciones de aprendizaje automático e inteligencia artificial (IA) respeten los principios de igualdad y no discriminación. La declaración establece un marco normativo basado en los derechos humanos que deben observar tanto sectores públicos como privados. Este marco busca garantizar que los algoritmos se utilicen de forma equitativa y justa, ofreciendo a las personas afectadas por posibles violaciones mecanismos efectivos de reparación. Aunque se ha avanzado en diálogos éticos sobre la inteligencia artificial, la declaración destaca que la ley de derechos humanos es fundamental para proteger y remediar daños causados por esta tecnología.

Si bien su enfoque principal está en el aprendizaje automático y los derechos a la igualdad y no discriminación, muchos de los principios expuestos se aplican de manera más amplia a otros sistemas de inteligencia artificial. Además, la inteligencia artificial impacta diversos derechos humanos, incluyendo la privacidad, la libertad de expresión, la participación cultural, la reparación y el derecho a la vida, evidenciando la amplitud y profundidad de su influencia en el ámbito de los derechos fundamentales.

La precitada declaración está dividida en tres partes: en primer lugar, asigna a los Estados la responsabilidad de prevenir la discriminación en el diseño e implementación de sistemas de aprendizaje automático, ya sea en contextos públicos o en colaboraciones público-privadas. Se incluyen directrices para identificar riesgos relacionados con estos sistemas, garantizar la transparencia y la rendición de cuentas, asegurar la supervisión independiente y promover la igualdad. A nivel internacional,

se enfatiza la necesidad de construir una visión colectiva para el desarrollo tecnológico con enfoque en derechos humanos, proponiendo que la política exterior contemple una declaración actualizada de derechos humanos para la era de la inteligencia artificial.

En segundo lugar, se señalan las obligaciones de los actores privados, fundamentadas en el marco de diligencia debida en derechos humanos de la ONU para empresas. Estas responsabilidades incluyen la detección y evaluación de riesgos discriminatorios, la adopción de medidas para prevenir y mitigar tales riesgos, realizados mediante auditorías externas independientes cuando existan riesgos significativos de violaciones a los derechos humanos, y la transparencia respecto a especificaciones técnicas y datos empleados.

Y finalmente, afirma el derecho a contar con recursos efectivos para responsabilizar a quienes violen los derechos. Insta a los gobiernos a establecer estándares de debido proceso para la utilización del aprendizaje automático en el sector público, a manejar con precaución su uso en el sistema penal, a definir claramente las responsabilidades legales en el desarrollo e implementación de estos sistemas y a identificar a los sujetos jurídicos responsables de las decisiones derivadas de su uso.

Luego, lo analizado en este epígrafe permite concluir que el avance tecnológico ha transformado radicalmente el entorno social y jurídico, generando un nuevo espacio digital donde los derechos humanos deben ser reinterpretados y adaptados. En ese entorno la irrupción de la inteligencia artificial amplía la complejidad de proteger los derechos fundamentales en un contexto donde la tecnología impone retos inéditos en cuanto a transparencia, participación y equidad.

Este nuevo escenario demanda no solo la actualización normativa, sino también el diseño de mecanismos efectivos de rendición de cuentas y supervisión que aseguren un uso responsable y

ético de la inteligencia artificial, especialmente ante riesgos de discriminación, vigilancia masiva y opacidad en los procesos automatizados. En definitiva, la gobernanza en el ámbito digital requiere un compromiso conjunto de Estados, sociedad civil y sector privado para construir una administración tecnológica que respete y promueva los derechos digitales, fortaleciendo la justicia y la inclusión en la era digital.

3.3. Protección jurídica de la inteligencia artificial. Acercamiento a la Propiedad Intelectual

Es fundamental partir de que los sistemas de inteligencia artificial no solo se configuran con *softwares* sino que pueden estar integrados por distintos elementos y componentes, variables en función de su tipología, como pueden ser además *hardwares*, algoritmos y bases de datos (Muñoz Vela, 2021). Cada uno de ellos puede comportar independientemente una forma de tutela jurídica, de manera que la defensa legal de todo el sistema dependerá de la conjunción de formas de protección aplicables a los mismos.

El principal elemento, alrededor del cual gira todo el engranaje del sistema de inteligencia artificial, es el *software* o programa de ordenador, para el cual en el campo jurídico se escindieron dos modalidades de protección principales: por un lado, a través del Derecho de Autor y por otro mediante las patentes por el Derecho de Propiedad Industrial.

Los programas de ordenador reciben protección por medio del Derecho de Autor al calificarse como obras, tomando en cuenta que la elaboración de todas sus partes: algoritmos, el código fuente, el código objeto, una carta de flujo u organigrama y los manuales de uso (Flórez-Acero et al., 2017), requiere de un complejo proceso creativo de desarrollo de ideas, diseño de estructuras y configuración.

Esta afirmación sustenta el reconocimiento que en ese sentido hacen el Tratado de la OMPI sobre Derecho de Autor (WCT) y el Acuerdo sobre los Aspectos de los Derechos de Propiedad Intelectual relacionados con el Comercio (ADPIC) en sus artículos 4 y 10 respectivamente, en correspondencia con el artículo 2.1 del Convenio de Berna, equiparándolos a las obras literarias en función de su expresión escrita (Ríos, 2002). Siguiendo esa premisa varios son los países que han adecuado sus ordenamientos legales en función de otorgar protección al *software* bajo los mismos términos que las obras literarias, confiriéndole similares derechos a su titular como parte de la protección del *copyright*.

Por solo citar algunos ejemplos está España que lo reconoce en su Ley de Propiedad Intelectual, en el artículo 10.1 i), y que además adopta la normativa especial en materia de protección jurídica de programas de ordenador dictada por el Parlamento Europeo; igualmente La India define al programa de ordenador en su Ley de Derechos de Autor incluyéndolo como parte de las obras literarias, por lo que otorga con relación a este objeto las mismas prerrogativas dispuestas para las obras literarias. Por otro lado, se encuentran China, y Japón que en lugar de categorizar al programa de computador como obra literaria, lo sitúan en igualdad de condiciones con respecto a ellas.

Los algoritmos por su parte, por regla general, figuran dentro del *software* y para que las ideas que lo integran no queden fuera de la protección que brinda el Derecho de Autor a aquel, estas "deben estar plasmadas literalmente en el código fuente" (Flórez-Acero et al., 2017). A su vez, los algoritmos que no se incluyan en el código fuente podrán encontrar amparo legal en los marcos jurídicos de competencia desleal o los marcos penales reguladores de los delitos contra el secreto empresarial, según se establece en el artículo 39 apartados 1 y 2 del Acuerdo sobre los ADPIC en correspondencia con el 10 *bis* del Acta de París.

Por otro lado, están las bases de datos de las cuales el *software* extrae información para su posterior utilización en los sistemas de inteligencia artificial y que alcanzan resguardo legal como obras de acuerdo con el artículo 2.5 del Convenio de Berna, el artículo 10.2 del Acuerdo sobre los ADPIC y el 5 del WCT. Con la salvedad de que la protección no se hace extensible al material compilado, que siempre podrá encontrar resguardo legal como parte del secreto empresarial por las normas que regulan la competencia desleal; misma tutela que puede otorgársele a la agrupación de datos cuando esta no implique originalidad ni creatividad alguna que justifique su protección como obra.

Asimismo, puede el *software* estar implementado en *hardwares*, cuya protección legal se asienta en las disposiciones relativas a las patentes de invenciones, modelos de utilidad y/o diseños industriales previstas por el Derecho de Propiedad Industrial en el Convenio de París para la Protección de la Propiedad Industrial, que reconoce estos objetos en su artículo 1, apartado 2.

Se concluye que el *software* es requisito *sine qua non* en la configuración del sistema de inteligencia artificial, pudiendo constituir por sí solo uno de estos sistemas; al tiempo que pueden darse otras clases de acuerdo a la combinación, junto a él, de todos o parte de los elementos descritos con anterioridad. De ahí que su tutela quede simplemente establecida por medio del Derecho de Autor o del otorgamiento de patentes de *software*, según regulen los ordenamientos nacionales la salvaguarda legal de este objeto; o se produzca una protección acumulada de derechos por la coincidencia de varias modalidades de tutela legal.

Como pudo corroborarse no se abren grandes brechas en cuanto a las vías de reconocimiento legal de los sistemas de inteligencia artificial. En realidad en la relación IA-Derecho lo más controversial no ha sido la protección del sistema en sí mismo sino la tutela jurídica de los recientes resultados de su uso que

pueden comportar, como se explicaba anteriormente, caracteres especiales capaces de convertirlos en verdaderos objetos merecedores del amparo legal que brinda el Derecho de Autor, siempre y cuando estos ostenten todos los requisitos exigidos en la norma para otorgar su protección a esa situación jurídica concreta.

Un sistema de inteligencia artificial, como se apuntaba, sostiene su funcionamiento en la implementación de un variado conjunto de técnicas computacionales, que pueden agruparse: técnicas destinadas al razonamiento y la toma de decisiones; técnicas destinadas al aprendizaje; y técnicas destinadas a la acción y a la capacidad de interacción (Fernández, 2021) que al aplicarse condicionan su calificación en múltiples tipologías, a saber existen sistemas que tratan de emular el pensamiento humano; sistemas que tratan de imitar el comportamiento humano; y sistemas que tratan de realizar ambas operaciones con un cierto grado de lógica y racionalidad (Russell y Norving, 2009).

De este catálogo, existen un grupo de sistemas de inteligencia artificial que se corresponden con el grupo de aquellos que buscan simular el pensamiento del hombre: de forma lógico-racional o no; mediante un proceso semejante al de sus funciones cognitivas, el cerebro humano corresponde al de un sistema altamente complejo que puede realizar muchas operaciones simultáneamente y esto se debe a que funciona por medio de una red neuronal, compuesta por un número inmenso de neuronas con interconexiones masivas entre ellas, vinculadas mediante sinapsis por la emisión de señales electroquímicas, que son la causa, entre otros aspectos, de las funciones cognitivas humanas (Velásquez et al., 2009), que se lleva a cabo utilizando, entre otras, las tecnologías del *machine learning* o aprendizaje automático y del *deep learning* o aprendizaje profundo (Fernández, 2021), con lo cual son capaces de generar creaciones, invenciones y otros objetos protegidos por los Derechos de Propiedad Intelectual.

Resultados de estas técnicas se tienen un conjunto de creaciones que se dividen en dos grupos: las creaciones asistidas por computadoras y las creaciones algorítmicas. Las creaciones asistidas por computadoras son aquellas en las que el programa de ordenador es utilizado como una mera herramienta por el autor, al tiempo que las creaciones algorítmicas son las generadas por la inteligencia artificial independientemente, con una intervención humana mínima o meramente técnica (Fernández, 2021; Muñoz, 2021; Saíz, 2019). Alrededor de estas han surgido varias interrogantes cuyo punto de partida común es su aptitud real para ser efectivamente un objeto tutelable por el Derecho de Autor, idóneo para hacer surgir en cabeza de su creador la titularidad de los derechos que por regla general derivan de las obras.

En ese sentido, valorando los requisitos de protección de las obras: originalidad y participación humana en el resultado, se puede concluir que las creaciones generadas por el sistema de inteligencia artificial con una relativa autonomía donde la participación humana se reduce a programar el *software* y alimentarlo con datos no terminan siendo obras propias de un autor que lleven su toque personal, ergo tampoco portadoras de originalidad tal como la demanda el Derecho de Autor que hoy rige. Debido a que el sistema por sí mismo no toma las elecciones que conducen al resultado de forma libre y creativa y por otro lado porque las acciones realizadas por la persona física no poseen la entidad suficiente que se demanda para que sean creativas.

Mientras que, por el reverso, las creaciones en que el hombre utiliza al sistema de inteligencia artificial como apoyo en la creación, donde la implicación de la persona natural es relevante, está claro que gozan del carácter original que se exige para otorgar protección jurídica, una vez que el factor humano se halla presente en ambos casos plasmando su impronta tanto en el proceso de creación como en la creación misma.

Así, al respecto de la atribución de autoría y correlativa titularidad de las creaciones fruto de sistemas de inteligencia artificial relativamente independientes al programador, Fernández Carballo-Calero en armonía con Ginsburg y Budiardjo advierte que en todo caso la condición de autor y titular de los derechos correspondientes le vendría dada por el hecho de aportar el factor humano relevante a la creación de las correspondientes obras, y no por ser el creador del sistema (Fernández, 2021). Se refiere el autor al criterio de que los diseñadores de máquinas completamente generativas formulan un plan creativo, manifestado en los algoritmos y procesos que conducen a la creación, cumpliendo así con el requisito de concepción exigido por el Derecho de Autor. Sin embargo, quien escribe se muestra en desacuerdo atendiendo al hecho de que la mera programación sin la estructuración de órdenes específicas a la máquina le deja a esta un margen de libertad decisional bastante amplio que no se corresponde con la exigencia contemporánea de concepción ligada a un control notable del proceso creativo, con lo cual al no desarrollarse ese control no podría hablarse de relevancia de la participación del programador en el proceso creativo.

Por lo que tratando de solucionar el inconveniente que supone la falta de relevancia del comportamiento humano en ese proceso creativo se plantea entonces, a modo de conclusión, que sería más correcta la protección de las creaciones generadas con una cierta autonomía del sistema inteligente por medio de otro Derecho de Propiedad Intelectual, diferente al Derecho de Autor (Saíz, 2019). Adviértase que no se valora aquí la idea de dejar estas creaciones en el dominio público en vistas de las consecuencias negativas que traería esto para el autor humano pues supondría que los costos de acceso a ellas sean menores que los de las obras de creación humana en exclusiva, lo que a la larga traería consigo la desincentivación de la creación por parte del hombre. En cambio, la atribución de derechos de autor, en cualquier caso a la máquina o a la persona que realice los arreglos necesarios

para que esta, de forma relativamente independiente, produzca el resultado creativo, no debería ser posible ya que de ningún modo se adecuaría al fundamento del Derecho de Autor.

Por otra parte, se ha demostrado que la ya mencionada inteligencia artificial generativa, al ser capaz de imitar el proceso de creación del hombre, puede utilizar el conocimiento previo con que fue entrenada para generar soluciones novedosas que puedan llegar a ser consideradas invenciones desde el punto de vista del Derecho de Propiedad Industrial. Nótese que en la actualidad ya existen sistemas de inteligencia artificial a los que se les ha reconocido con la categoría de inventor. Una muestra de ello es el sistema DABUS, un inventor artificial basado en un sistema de redes neuronales artificiales creado por Stephen Thaler (Kokane, 2021), capaz de generar ideas novedosas y reforzarlas según su valor percibido que generó dos invenciones, en cuya solicitud de patentes se listó a esta inteligencia artificial como inventor.

El surgimiento de sistemas de inteligencia artificial capaces de generar invenciones autónomas ha desafiado los fundamentos tradicionales de la Propiedad Industrial, que históricamente atribuyen la titularidad de las patentes a personas naturales o jurídicas. El caso DABUS ha puesto en evidencia los límites del marco legal actual al proponer a una inteligencia artificial como inventora de un recipiente para bebidas basado en geometría fractal. La Oficina Europea de Patentes (OEP) rechazó la solicitud al exigir que el inventor sea una "persona natural con capacidad jurídica", argumentando que la inteligencia artificial carece de personalidad legal. Sin embargo, tribunales como el australiano han reconocido la posibilidad de atribuir inventiva a sistemas de inteligencia artificial, aunque manteniendo la titularidad en humanos.

Recibiendo diferentes análisis en las distintas jurisdicciones (Rivera, 2024). Las oficinas de patentes en Estados Unidos (EE.UU), la Unión Europea (UE) y Sudáfrica discutieron si una

inteligencia artificial podría ser reconocida como inventor. Sin embargo, esta solicitud fue rechazada tanto por la Oficina de Propiedad Intelectual del Reino Unido (UKIPO) como por la Oficina Europea de Patentes (EPO) (Schulze, 2023), alegando que el inventor debe ser una persona natural, aunque reconocieron que las invenciones eran novedosas, inventivas y aplicables industrialmente.

En este sentido se debe destacar que el surgimiento de la inteligencia artificial generativa ha acelerado la necesidad de replantear los sistemas de patentes, ya que estos sistemas pueden producir invenciones sin intervención humana directa, desafiando conceptos jurídicos tradicionales como "inventor humano" y "actividad inventiva" (Fernández Gil, 2021). Este avance tecnológico obliga a reconsiderar los marcos legales actuales, especialmente en un ecosistema de inteligencia artificial, entorno de colaboración entre la academia, empresas, gobierno y sociedad civil que facilita la cooperación y el intercambio de conocimientos, permitiendo que las soluciones tecnológicas se implementen de forma más eficiente y con mayor impacto social y económico, donde las creaciones surgen de algoritmos entrenados con grandes conjuntos de datos, muchas veces sin supervisión explícita (Tauk y Salomão, 2023).

La generación de normas específicas dirigidas a evaluar y controlar la compatibilidad de los desarrollos tecnológicos con los marcos jurídicos existentes resulta imprescindible para garantizar la protección efectiva de los derechos de propiedad intelectual en un entorno donde la creación automatizada se multiplica. Este proceso implica una actualización y armonización legislativa a nivel internacional, que considere las particularidades técnicas y éticas de la inteligencia artificial, permitiendo un equilibrio entre la innovación y la protección de los derechos, para que la tecnología sirva al interés público sin comprometer los derechos individuales y colectivos. En definitiva, un marco jurídico sólido y adaptable es

fundamental para promover un desarrollo tecnológico responsable y garantizar la justicia en el ámbito intelectual en la era digital.

3.4. Los sistemas de responsabilidad legal aplicables ante las consecuencias de la inteligencia artificial

Los sistemas de Inteligencia artificial pueden ser susceptibles de provocar daños y perjuicios en el proceso de generación de productos y servicios, durante su comercialización y en el uso que hace el usuario de inteligencia artificial que puede manipular o emplear el producto con finalidad de dañar, de modo intencional. Entre seres humanos y los sistemas autónomos se producen interacciones cada vez más frecuentes que representan riesgos de diversa naturaleza sobre la persona y su patrimonio.

Esto pone la inteligencia artificial ante varias interrogantes que no encuentran respuesta en el sistema tradicional del derecho de daños, es preciso establecer la solución legal para la configuración de daños, cómo actuaría la víctima del resultado dañoso y qué sistema de atribución de responsabilidad resultaría más efectivo o ajustado a las características de la inteligencia artificial.

Cabría preguntar, a quién atribuir la responsabilidad: al fabricante o la larga cadena de sujetos que intervienen en el ciclo de vida de generación del producto, a los sujetos comercializadores, a los prestadores de servicios o a quienes lo utilizan y pueden cometer errores o actuar con negligencia sin adoptar las mayores precauciones.

Otra cuestión a valorar sobre el tema de responsabilidad es la autonomía de los sistemas de inteligencia artificial, en su vertiente generativa, esta puede llegar a tomar decisiones imprevisibles y a veces no determinadas de manera previa, sobre todo cuando esta es capaz de desarrollarse por sí misma sin intervención del factor humano, lo que lleva a cuestionarse los límites para la exigencia de responsabilidad, hasta dónde alcanzaría la implicación humana, cuando el sistema de inteligencia artificial

autónomo es capaz de generar determinadas consecuencias dañinas.

La opacidad de los datos puede llevar a la incomprensión de la configuración del sistema de inteligencia artificial y su funcionamiento, constituyendo un obstáculo complejo para las víctimas y la posibilidad de establecer la relación causal y aportar el material probatorio necesario en términos de responsabilidad objetiva, este es un fenómeno conocido como black box o caja negra (Ayo Ferrándiz, 2025).

Frente a estas interrogantes la doctrina, partiendo de las soluciones que tradicionalmente ofrece el derecho de daños sobre la autonomía y capacidad, ha ofrecido posibles respuestas.

Una de las vertientes a valorar es la responsabilidad por productos defectuosos entendiendo en esta categoría los sistemas de inteligencia artificial, a estos efectos debería considerarse la inteligencia artificial como un producto de naturaleza física, lo que no es problema en cuanto a vehículos autónomos, los robots y otros equipos que poseen aplicaciones de inteligencia artificial. Sin embargo, aquellos que no poseen corporeidad o soporte físico, como los softwares podrían no considerarse como productos, no obstante, se apunta sobre esto su condición de bienes muebles, dado su carácter intangible es discutible esta postura. En este caso cabria cuestionar si los datos suministrados por un tercero son los causantes del daño y por ende no depende del fabricante sino del suministrador (Sotomayor, 2024).

Se sugiere entender responsable al fabricante y a todos lo que contribuyen al ciclo de vida de elaboración del producto asumiendo la responsabilidad solidaria frente a la víctima que solo debe demostrar el defecto y el consiguiente daño, resultante de estos. Sobre este aspecto la Directiva Europea 85/374/CEE de 28 de septiembre de 2022, que regula la responsabilidad civil por productos defectuosos, ha sido adaptada a la exigencia de responsabilidad al fabricante que debe indemnizar a los

consumidores cuando los productos ocasionen daños o perjuicios al consumidor del producto.

Esta Directiva europea de 2022 será sustituida pronto por la Directiva de productos defectuosos de 2024, que entrará en vigor en 2026, la que contempla la responsabilidad civil objetiva al fabricante ampliando su ámbito objetivo a los programas informáticos incluida en esta categoría la IA.

Cabe analizar la postura que considera hacer extensivo el sistema de responsabilidad civil siguiendo el criterio de imputación objetiva a partir de la existencia de un riesgo tolerado a la inteligencia artificial en este contexto, se considera que una actividad que genera beneficios y tiene un determinado riesgo permitido generando beneficios para el titular si es susceptible de provocar daños, cabe la responsabilidad a quien utiliza el producto peligroso; el obtener beneficios hace surgir un deber inherente de supervisar y utilizar el producto de inteligencia artificial, para asegurar que su funcionamiento se mantenga dentro de determinado límites de seguridad. Son las llamadas actividades que generan riesgos, que de manera absoluta cuando provoquen daños, traen consigo la obligación de indemnizar al perjudicado.

Tal idea está sujeta a objeciones porque la tecnología de inteligencia artificial está centrada en minimizar los riesgos, de producción de daños, y no todos los sistemas de inteligencia artificial pueden ser catalogados de peligrosos, por ello solo podría imputarse al usuario responsabilidad cuando se compruebe el mal funcionamiento o una actitud negligente y que exista un comprobado nexo causal entre el daño y la conducta negligente (Sotomayor, 2024).

La Unión Europea ha tenido avances en este tema enfocados en la armonización de los países europeos, en torno a la resolución de conflictos vinculados a la exigencia de responsabilidad civil, que penden de las regulaciones internas de los estados

miembros, a esta intención responde la Propuesta de Directiva sobre responsabilidad en materia de inteligencia artificial (Unión Europea, 2022), dispone su art. 1, normas comunes sobre la exhibición de pruebas relativas a los sistemas de inteligencia artificial de alto riesgo, a fin de permitir a los demandantes (persona que interponga una demanda por daños y perjuicios y que se ha visto perjudicada por la información de salida de un sistema de inteligencia artificial, o por la no producción por parte de dicho sistema de una información de salida que debería haber producido; también aborda la carga de la prueba en las demandas de responsabilidad civil extracontractual subjetiva. No establece que debe ser considerado indemnizable.

La Propuesta de directiva también contempla prescripciones sobre la carga de la prueba, que como principio procesal debe recaer sobre la víctima o perjudicado y así lo contemplan los ordenamientos jurídicos nacionales, sin embargo, ello resulta complejo para la parte afectada por las características que rodean la aplicación de inteligencia artificial, su complejidad y el efecto caja negra, a lo que se añaden los altos costos que introducen un claro factor de inequidad procesal (Vallespín Pérez, 2025).

La pregunta de si la normativa general sobre responsabilidad es suficiente o no, en estos momentos parece obligada. Todo parece apuntar a que se requieren normas y principios específicos que aporten claridad sobre la responsabilidad jurídica de los distintos agentes y su responsabilidad por los actos y omisiones de los sistemas de inteligencia artificial, cuya causa no pueda atribuirse a un agente humano concreto, y de si los actos u omisiones de los sistemas de inteligencia artificial causantes de daños que podría haberse evitado. No es un asunto acabado la determinación de la forma en que se exigirá responsabilidad en esta materia (Iturmendi Morales, 2020).

El Reglamento de inteligencia artificial de la Unión Europea (Parlamento Europeo, 2024), no contiene normas sobre

responsabilidad de la inteligencia artificial, se posiciona en la fijación de niveles de riesgos y la prevención, determina la exigencia de medidas precautorias para cada actividad según el nivel de riesgo, establece las actividades peligrosas, prohibidas, define requisitos, obligaciones, garantías y responsabilidades de los distintos operadores en función de su actividad, ya sean fabricantes, proveedores distribuidores, comercializadores y usuarios, con sanciones y formas de cumplimiento. Este Reglamento es complementario de las reglas de responsabilidad, posibilitando que se tomen en cuenta las normas de debida diligencia y el incumplimiento de las obligaciones establecidas en la ley.

La propuesta de Directiva de responsabilidad en materia de inteligencia artificial de la Unión Europea quedó pospuesta, ante la aprobación del Reglamento europeo para u uso, por lo que se desecha un valido intento de regulación de responsabilidad civil objetiva, con una alta protección a la víctima y un fuerte componente procesal, a esos efectos Ayo Ferrandiz et al. (2025), refieren que:

> La PD-RIA, por tanto, no pretendía una introducción de nuevas normas sustantivas ni armonizar aspectos generales de la responsabilidad civil, como la definición de la culpa o la causalidad, los diferentes tipos de daños que pueden dar lugar a reclamaciones, la distribución de la responsabilidad entre varios sujetos causantes de los daños, la concurrencia de culpas, el cálculo de los daños y perjuicios o los plazos de prescripción, cuestiones todas ellas que se regirían fundamentalmente por la legislación nacional de cada Estado miembro. Se trataba, además, de una Directiva de armonización mínima, por lo que se permitía a los Estados miembros la adopción de normas más estrictas.

Por esta razón, cuando no sea de aplicación el Reglamento que regula la responsabilidad por productos defectuosos que incluye

los productos informáticos incluidos la inteligencia artificial en el contexto europeo habría que acudir a las normas sobre responsabilidad de los ordenamientos jurídicos internos que no se ajustan a las novedades y su complejidad técnica, lo que plantea un escenario donde quedan pendientes los estudios y consensos doctrinales sobre responsabilidad y la necesidad de reparar los daños y perjuicios con un régimen especial en materia de inteligencia artificial sobre todo en sistemas jurídicos regionales menos avanzados en términos de marco jurídico sobre su uso.

3.5. Sesgos algorítmicos, discriminación y derechos humanos

El desarrollo de los sistemas de inteligencia artificial y los algoritmos de aprendizaje automático ha llevado a la digitalización de datos a escala masiva. Los algoritmos emplean esos datos para tomar decisiones y acometer acciones en diferentes sectores. Sin embargo, los conjuntos de datos con los que se entrenan los algoritmos suelen ser incompletos o infrarrepresentan a determinados grupos de personas. Si determinados grupos están excesivamente representados o son menos representados en los conjuntos de entrenamiento, incluidos los grupos raciales, de diversas identidades de género, étnicos, puede producirse un sesgo algorítmico. Del mismo modo, si los conjuntos de entrenamiento incluyen datos ya sesgados, pueden producir resultados sesgados (Organización de las Naciones Unidas, 2024b).

Un elemento importante en este tema son los sesgos históricos, estos pueden afectar a los propios datos. Un elemento central del aprendizaje automático es hacer predicciones sobre el futuro a partir de datos del pasado. Si los datos del pasado están sesgados en contra de ciertos grupos, en particular por brechas de género, criterios raciales y étnicos, los modelos informáticos pueden reproducir y amplificar esos sesgos (Calvete, 2024).

La última cuestión relacionada con los datos es la privacidad. Los datos utilizados en los sistemas de inteligencia artificial suelen incluir información personal de las personas a las que pertenecen. La recogida y el tratamiento de datos sin consentimiento vulnera el derecho a la intimidad. Para las personas pertenecientes a grupos marginados por motivos raciales, los problemas de derechos humanos relacionados con el derecho a la intimidad pueden verse amplificados

Las entidades creadoras de inteligencia artificial están obligadas a salvaguardar los principios de equidad, justicia social y no discriminación. Este mandato plantea una perspectiva integradora para garantizar la disponibilidad y accesibilidad equitativas de los beneficios de la tecnología de la inteligencia artificial para todos, teniendo en cuenta las distintas necesidades de grupos demográficos dispares, como los delimitados por la edad, la cultura, el idioma, la capacidad física o cognitiva, el género y la situación socioeconómica.

Existen varios principios éticos vinculados al tema en cuestión (Morandím-Ahuerma, 2023):

1. Proporcionalidad e inocuidad

Según la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, el principio de proporcionalidad subraya la importancia de garantizar que la inteligencia artificial se desarrolle y utilice de forma que se ajuste a su finalidad prevista, evitando al mismo tiempo cualquier exceso o peligro innecesarios. En esencia, esto implica que su aplicación debe mantenerse dentro de límites para lograr el objetivo predeterminado, sin aventurarse en una utilización excesiva o innecesaria que supere el objetivo establecido. Por ejemplo, si se utiliza la para detectar y prevenir delitos, se debe asegurar que las medidas tomadas sean proporcionales al riesgo de cometerlo y no utilizar demasiada fuerza o recursos en un objetivo más allá de lo establecido.

Por ello, se debe aplicar la evaluación de riesgos y resultados colaterales

La inocuidad, por su parte, se refiere a la necesidad de evitar el uso de la inteligencia artificial para fines que puedan causar daño o perjuicio a las personas o al medio ambiente. La inteligencia artificial no debe ser utilizada para discriminar a ciertos grupos, para la vigilancia masiva o para la manipulación psicológica. Para garantizar una mejor comprensión, el método elegido de inteligencia artificial debe ser adecuado y equilibrado para lograr un objetivo legítimo específico. No debe vulnerar los principios fundamentales de los derechos humanos y debe adaptarse bien a las circunstancias concretas que se presenten, basándose en principios científicos fiables

El Convenio de Europa establece requisitos de transparencia y supervisión adaptados a contextos y riesgos específicos, incluida la identificación de contenidos generados por sistemas de inteligencia artificial:

> "Las partes tendrán que adoptar medidas para identificar, evaluar, prevenir y mitigar posibles riesgos y evaluar la necesidad de una moratoria, una prohibición u otras medidas apropiadas en relación con el uso de sistemas de IA, cuando sus riesgos puedan ser incompatibles con las normas de los derechos humanos. También tendrán que garantizar la rendición de cuentas y la responsabilidad por el impacto negativo, y que los sistemas de IA respeten la igualdad, incluida la igualdad de género, la prohibición de la discriminación y el derecho a la intimidad."

> "Además, las partes del tratado tendrán que garantizar la disponibilidad de recursos legales para las víctimas de violaciones de los derechos humanos relacionadas con el uso de sistemas de IA y de garantías procesales, incluida la notificación a todas las personas que interactúen con sistemas de IA de que

están interactuando con estos sistemas. En cuanto a los riesgos para la democracia, el tratado requiere a las partes que adopten medidas para garantizar que los sistemas de IA no se utilicen para minar las instituciones y procesos democráticos, incluido el principio de separación de poderes, el respeto de la independencia judicial y el acceso a la justicia."

El uso generalizado de la inteligencia artificial por parte de los Estados incluidas la elaboración de perfiles, la adopción de medidas de vigilancia y las tecnologías de aprendizaje automático, afecta al disfrute del derecho a la intimidad y los derechos conexos.

Existen cuatro aspectos clave:

1. Los sistemas de inteligencia artificial se basan en grandes conjuntos de datos, con información sobre las personas recopilada, compartida, fusionada y analizada de múltiples formas, a menudo opacas.
2. Los datos utilizados para informar y guiar los sistemas de inteligencia artificial pueden ser defectuosos, discriminatorios, obsoletos o irrelevantes, afirma, y añade que el almacenamiento de datos a largo plazo también plantea riesgos particulares, ya que en el futuro los datos podrían ser explotados de formas aún desconocidas.
3. Dado el rápido y continuo crecimiento de la IA, una de las cuestiones más urgentes en materia de derechos humanos a las que nos enfrentamos es colmar la inmensa laguna que existe en la rendición de cuentas sobre cómo se recopilan, almacenan, comparten y utilizan los datos.
4. Deben plantearse serias dudas sobre las inferencias, predicciones y supervisión que realizan las herramientas de inteligencia artificial, incluida la búsqueda de patrones de comportamiento humano.

El Informe del Alto Comisionado de las Naciones Unidas para los Derechos Humanos" El derecho a la intimidad en la era digital". Michelle Bachelet, concluyó que "los conjuntos de datos sesgados en los que se basan los sistemas de inteligencia artificial pueden dar lugar a decisiones discriminatorias, lo que supone un grave riesgo para los grupos ya marginados. Por eso es necesario evaluar y vigilar sistemáticamente los efectos de los sistemas de inteligencia artificial para identificar y mitigar los riesgos para los derechos humanos".

"El poder de la IA para servir a las personas es innegable, pero también lo es su capacidad para alimentar violaciones de derechos humanos a una escala enorme y prácticamente sin visibilidad" (Organización de las Naciones Unidas, 2021).

Aplicación de la inteligencia artificial, seguridad y sistema de justicia penal

Las aplicaciones de inteligencia artificial a los sistemas de seguridad y justicia penal son variadas y entrañan riesgos de discriminación. A continuación, se examinarán varias de estas.

Identificación automática: Las fuerzas policiales emplean herramientas de identificación automatizadas para relacionar lo que observan en un entorno concreto con una posible "coincidencia" en una base de datos. Uno de los tipos más comunes de herramientas de identificación automatizada es la tecnología de reconocimiento facial. Las herramientas de reconocimiento facial toman imágenes de vídeo o fotografías de una persona y las comparten con algoritmos. A continuación, los algoritmos comparan las imágenes con una base de datos de fotografías policiales, fotografías de carnés de conducir, bases de datos de identificación u otras imágenes con el objetivo de identificar a la persona. Los algoritmos de clasificación de género se entrenan con conjuntos de datos que reúnen, con una alta prevalencia de rostros masculinos blancos. La falta de diversidad

racial, de género y cultural en los conjuntos de entrenamiento de las herramientas de inteligencia artificial conduce a uno de los típicos problemas de datos.

Los sistemas de detección de disparos son otro tipo común de herramienta de detección automatizada que utilizan agentes del orden en varios países. Un sistema, denominado ShotSpotter, consiste en colocar en los barrios sensores que contienen un micrófono, un sistema GPS, dispositivos de memoria y procesamiento y capacidad celular Cuando los sensores detectan un ruido que podría ser un disparo, un algoritmo triangula la ubicación de lo que haya causado el ruido. La colocación de sistemas de detección de disparos en comunidades en las que viven grupos raciales y étnicos marginados y las imprecisiones de los sistemas de detección de disparos exacerban los prejuicios sistémicos de las propias fuerzas del orden.

Algoritmos policiales predictivos: Otra forma de tecnología de inteligencia artificial utilizada habitualmente por las fuerzas de seguridad es la vigilancia policial predictiva. Las herramientas de predicción policial evalúan quién podría cometer delitos en el futuro y dónde podrían producirse, basándose en datos personales y de localización. La vigilancia policial predictiva puede exacerbar el exceso histórico de vigilancia de comunidades según criterios raciales y étnicos a partir de los datos suministrados a los sistemas.

En el ámbito de la salud también son empleadas las aplicaciones de inteligencia artificial que entrañan sesgos discriminatorios con frecuencia. Así se pueden relacionar:

Puntuaciones de riesgo sanitario. La inteligencia artificial puede utilizarse con miras a establecer puntuaciones de riesgo para la salud con diversos fines sanitarios, como el diagnóstico médico y la planificación de los cuidados. Cuando estos algoritmos se utilizan para asignar recursos sanitarios pueden producirse efectos de discriminación racial, debido al diseño algorítmico y

a los datos utilizados para entrenar los sistemas de inteligencia artificial.

Detección de enfermedades: Otra aplicación de las tecnologías de inteligencia artificial es la detección de enfermedades, entre ellas el cáncer. Los sistemas de inteligencia artificial entrenados con amplios conjuntos de datos que comprenden miles o millones de imágenes, incluidas exploraciones radiológicas, imágenes patológicas y fotografías, pueden aprender a distinguir entre lesiones normales y cancerosas. Este despliegue de inteligencia artificial puede contribuir significativamente a la detección precoz del cáncer, lo que podría salvar vidas y mejorar la eficiencia del sistema de salud. Sin embargo, es posible que las personas pertenecientes a grupos raciales y étnicos marginados no se beneficien en igual medida de estos avances debido a que los algoritmos no son extensibles a poblaciones de pacientes que no están adecuadamente representadas en los datos de entrenamiento.

Se sugiere que las empresas de salud y los Estados tomen distancia de la adopción de un enfoque con "daltonismo racial" respecto a la gobernanza y la regulación de las tecnologías emergentes, incluida la inteligencia artificial.

Riesgos de sesgo y discriminación en inteligencia artificial

Los sesgos en la Inteligencia Artificial pueden llevar a discriminación y dañar la reputación empresarial. Estos sesgos surgen de errores en diseño de modelos y datos sesgados. Se recomiendan prácticas como revisión de datos, pruebas de equidad y monitoreo continuo para mitigar estos riesgos y asegurar un funcionamiento ético de la inteligencia artificial.

I- Uno de los mayores riesgos a los que se exponen las empresas en el empleo de la Inteligencia Artificial (IA) es la obtención de

resultados sesgados (lo que se denomina "el sesgo", y si este sesgo se traslada a los datos personales, se podría terminar en un riesgo de discriminación/prejuicios (bien por sexo, raza, procedencia, creencia, religión origen étnico, etc.) con el consecuente impacto no deseado sobre ciertos grupos de personas, afectando en la reputación e imagen de las empresas. Así, "el sesgo" podría definirse como la discriminación o prejuicio sobre la realidad que afectan a la toma de decisiones (o resultados obtenidos) de un sistema de inteligencia artificial.

II- Se pueden producir resultados desiguales que favorezcan a ciertos grupos sobre otros. Un sistema de reconocimiento facial podría ser más preciso a la hora de identificar rostros de personas de determinadas razas en comparación con otras razas. Se pueden tomar decisiones dirigidas a ciertos grupos de personas y discriminar otras de forma errónea.

III-Los algoritmos utilizados en la contratación pueden revelar preferencias de género basadas en el contenido educativo. Se puede distorsionar o influenciar en las decisiones. Un algoritmo publicitario podría mostrar artículos que favorezcan ciertas ideas predominantes, mientras que excluyese otros artículos relacionados con otras perspectivas

Los sesgos de género son prejuicios, estereotipos o desviaciones sistemáticas basados en el género que se reflejan en los resultados arrojados por esta tecnología. Los datos con los que se entrenan a los modelos de inteligencia artificial son la fuente a partir de la cual aprenden y crean nueva información. Si estos datos no son equitativos, se pueden perpetuar estos sesgos y dar lugar a resultados desiguales, distorsionados y discriminatorios. Por ejemplo, podrían exacerbar desigualdades existentes tanto en la forma en la que se nombra el género, como en la asociación estereotipada de habilidades y profesiones.

Existen distintos tipos de sesgos de género:

- **Sesgo en los datos de entrenamiento:** si los datos utilizados para entrenar un modelo de inteligencia artificial están sesgados hacia ciertos géneros, el modelo aprenderá y replicará esos sesgos. Por ejemplo, si un modelo de procesamiento de lenguaje natural se entrena principalmente en textos escritos por hombres, puede tener dificultades para comprender o responder adecuadamente al otro sexo.
- **Sesgo en las etiquetas o anotaciones:** las etiquetas o anotaciones se refiere a la información adicional proporcionada a un modelo de inteligencia artificial para entrenarlo y evaluar su rendimiento. Esta información puede reflejar estereotipos de género, dando como resultado que el modelo reproduzca estos estereotipos en sus predicciones o respuestas.
- **Sesgo en las presuposiciones de diseño:** los algoritmos y modelos de inteligencia artificial a menudo están diseñados con ciertos supuestos sobre el comportamiento humano o las características de los datos. Si estas hipótesis están sesgadas hacia una perspectiva de género particular, el modelo puede mostrar sesgos en sus resultados.
- **Amplificación de prejuicios sociales:** si los datos históricos reflejan prejuicios o desigualdades de género, los modelos de inteligencia artificial pueden aprender y replicar estos patrones, lo que puede llevar a la amplificación de estos prejuicios en las predicciones o decisiones del modelo.
- **Falta de diversidad en los equipos de desarrollo:** Si los equipos que desarrollan y entrenan los modelos de inteligencia artificial carecen de diversidad de género, es posible que no se aborden adecuadamente los sesgos de género en el proceso de desarrollo.

La Inteligencia Artificial (IA) con su impacto transformador en casi todos los aspectos de la vida y el trabajo, determina que las brechas

de género adquieran una relevancia crítica. A medida que esta tecnología avanza, la representación equitativa de géneros en su desarrollo y aplicación se vuelve esencial para evitar sesgos y garantizar que los sistemas de inteligencia artificial sean justos y efectivos para todos. Una brecha de género pronunciada en este campo puede llevar a la creación de sistemas de inteligencia artificial que perpetúen estereotipos de género y discriminación, afectando desde decisiones de contratación hasta diagnósticos médicos y preferencias de consumo.

"Los impactos de la inteligencia artificial son desiguales, entre los sectores, entre los géneros y entre los países. "El impacto de la tecnología sobre el empleo se manifiesta de manera diferente entre mujeres y hombres debido a que la distribución del empleo por sectores y ocupaciones no es equitativa por sexo" (Castaño, 2016, p. 44).

Se identifican tres brechas digitales de género:

La primera es el acceso material a la TIC;

- La segunda es "en torno a cómo y para qué se usan las tecnologías e incluye dos dimensiones esenciales: las habilidades digitales, imprescindibles para extraer el mayor provecho de ellas, y los patrones de actividades realizadas" (Castaño, 2019, p. 47).

- La tercera brecha digital de género se refiere a los beneficios que las mujeres obtienen del uso de las tecnologías.

La discriminación y el racismo mediados por la inteligencia artificial

El Informe de la Relatora Especial sobre las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia, Ashwini K. P., sobre "Formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia" (Organización de las Naciones Unidas, 2024b), resalta aspectos importantes sobre el reforzamiento de la inteligencia artificial en estas formas.

Establece los nexos estratégicos entre inteligencia artificial y racismo, resalta que los efectos de la inteligencia artificial deben verse a través de la mirada del racismo sistémico, que se define como "el funcionamiento de un sistema complejo e interrelacionado de leyes, políticas, prácticas y actitudes en las instituciones del Estado, el sector privado y las estructuras sociales que, combinadas, dan lugar a una discriminación, distinción, exclusión, restricción o preferencia, directa o indirecta, intencionada o no, de hecho o de derecho, por motivos de raza, color, ascendencia u origen nacional o étnico" (Organización de las Naciones Unidas, 2024b).

Un ejemplo de problemas relacionados con los datos es el siguiente:

En un estudio centrado en las bases de datos de imágenes de las fuerzas del orden en los Estados Unidos se demostró que las personas afrodescendientes tenían más probabilidades de ser señaladas erróneamente en las redes de reconocimiento facial utilizadas por los agentes del orden. Esto se debía a errores en la identificación facial de ese grupo y a la sobrerrepresentación de afrodescendientes en las bases de datos de fotografías policiales, que refleja pautas históricas de racismo sistémico.

En algunos casos, la inteligencia artificial puede utilizarse con fines explícitamente racistas realizando el despliegue selectivo contra grupos específicos, lo que da lugar a resultados discriminatorios.

Otro ejemplo de cómo la inteligencia artificial es utilizada con fines discriminatorios se describe a continuación:

Hay informes de que las fuerzas del orden utilizan intencionadamente la inteligencia artificial para vigilar y sobrevigilar a determinadas comunidades, siguiendo criterios de discriminación racial. Además, la discriminación intencionada puede producirse cuando los gobiernos y otras entidades

aprovechan las capacidades de la tecnología para vigilar o perfilar a grupos o individuos específicos, o centrarse en ellos, en función de su identidad racial o étnica.

La difusión de desinformación es otra forma de utilizar la inteligencia artificial con fines explícitamente racistas. Los lideres políticos y de opinión pueden utilizar la inteligencia artificial para generar textos, imágenes y vídeos con el fin de manipular la opinión pública y los procesos políticos a su favor y socavar la confianza en las instituciones, entre otros por motivos raciales. También se informa de gobiernos que han utilizado la inteligencia artificial para sembrar la discordia y facilitar la censura en línea.

Por otro lado, sobre el racismo y la inteligencia artificial el Informe de la Relatora Especial sobre las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia, Ashwini K. P. Identifica:

Cuatro formas transversales en las que la inteligencia artificial puede contribuir a las manifestaciones de discriminación racial:

- Los problemas relacionados con los datos.
- Las cuestiones de diseño de algoritmos.
- El uso intencionadamente discriminatorio de la inteligencia artificial.
- Las cuestiones relativas a la rendición de cuentas.

Sobre homofobia y racismo, los estudios también revelaron que los modelos de lenguaje tendieron a generar contenido negativo sobre personas homosexuales y ciertos grupos étnicos (Organización de las Naciones Unidas, 2024b).

3.6. Derecho penal y ciberdelitos vinculados a la inteligencia artificial

La inteligencia artificial plantea diversos retos ante el derecho penal y la exigencia de responsabilidad, la inteligencia artificial

generativa que consiste en la producción de patrones y algoritmos a partir del suministro de datos con cierto grado de autonomía, depende de la la intervención humana, que condiciona la actuación de la maquina y la generación de decisiones similares a la actuación humana, lo que determina que no varie el sistema de atribución de responsabilidad al quedar identificado el autor del producto y su autoría plena por acción u omisión.

Sin embargo, el desarrollo vertiginoso de los sistemas de inteligencia artificial y sus productos con alto grado de autonomía, díganse los vehículos autónomos, los robots, drones, internet de las cosas y otras aplicaciones como son los software y sus funciones de análisis de imágenes, motores de búsqueda, sistemas de reconocimiento facial y de voz, asistentes de voz, traducción de textos, generación de subtítulos, identificación y bloqueo de spam, hacen necesario repensar el sistema tradicional de responsabilidad penal e incluir implicaciones éticas que pondría límites a la producción de máquinas o equipos que podrían generar conductas transgresoras con un nivel de daño sobre la economía y las personas entre otros bienes jurídicos.

Sobre el nivel de autonomía y el grado de intervención humana según el modelo desarrollado: Harbers, Peeters y Neerincx, existirían cuatro modelos de inteligencia artificial en la actualidad según el grado de interacción hombre-máquina (Miró Llinares, 2018):

1. Man in the loop, cuando la inteligencia artificial necesita aportes humanos a intervalos de tiempo regulares para poder llevar a cabo sus acciones;
2. Man on the loop, si la máquina es capaz de actuar por sí misma a partir de una programación previa, pero el humano puede intervenir interrumpiendo o modificando las acciones del robot en cualquier momento; y
3. Man out of the loop, en este la máquina actúa de manera independiente durante ciertos períodos de tiempo y, en estos

intervalos, el ser humano no tiene influencia sobre las acciones del robot. Pero la posibilidad de interacción humana en tiempo real tan solo es una parte de lo que podría denominarse autonomía.

4. «No man on the loop», en el que la maquina es totalmente autónoma, donde o bien el aprendizaje por el que se toma la decisión no hubiera devenido de una acción humana, sino de la propia máquina que ha aprendido por sí misma o de otra máquina anterior, o bien el comportamiento de la máquina no depende de ese aprendizaje anterior y de las decisiones atribuidas previamente en la inteligencia artificial sino de algo ajeno a ello y propio del mismo sistema.

Aun no existe el nivel de autonomía 4, pero resulta una preocupación lograr deslindar entre el nivel de responsabilidad humana y la imposibilidad de exigencia de responsabilidad a los sistemas de inteligencia artificial. De esta manera, la autonomía, que puede comportar la causación, por parte de los sistemas de inteligencia artificial, de resultados no previstos y no deseados por sus diseñadores, fabricantes y/o usuarios, hace necesaria la exigencia de un control de evaluación de riesgos con carácter previo a la comercialización y de un mecanismo de control posterior, cuando el sistema ya se halle en funcionamiento.

Cuatrecasas Monforte señala que el grado de autonomía del sistema de inteligencia artificial es el que deberá determinar el grado de responsabilidad, tanto civil como, en su caso penal. Y es que, está claro que, si este ha sido diseñado para actuar sin supervisión humana, por ejemplo, cualquier fallo deberá atribuirse a su diseñador o, en su caso, fabricante, o incluso a los dos, en función de lo que un perito independiente concluya sobre el origen de tal error (puesto que no es lo mismo un incorrecto diseño del software que un defecto de fabricación, debiendo determinar quién habría podido haber previsto y evitado el resultado dañoso (Monforte, 2022).

Es necesario valorar otro elemento sustancial, es la opacidad de los datos, señalada por Januário (2023), que valora que la inteligencia artificial y los algoritmos que utiliza se consideran opacos, tal es su complejidad técnica que es muy difícil para el ser humano comprender sus procedimientos internos, las razones que subyacen en determinadas tomas de decisiones e incluso los datos que se utilizan como input y su relación con un determinado output.

En otras palabras, aunque tengamos acceso a una decisión concreta, entender el "cómo" y el "por qué" es muy complicado. Ello genera dudas sobre la legalidad y credibilidad de los datos utilizados como input, lo que demanda transparencia y las medidas para evitar la invasión de privacidad y del derecho a la intimidad. El uso de datos erróneos, insuficientes o sesgados, que se emplean en el sistema sin previa corrección, sumado a la falta de transparencia en cuanto a la configuración y funcionamiento de los algoritmos utilizados en el sistema, trae consigo importantes riesgos a los derechos fundamentales.

A partir del potencial y elevado daño (accidental o intencionado) que puede ocasionarse con los sistemas de inteligencia artificial, debe existir robustez y solidez técnica de los mismos, ello es necesario para evitar cualquier posible error o fallo, para ello es preciso la implementación de mecanismos de reparación eficaces y accesibles desde el punto de vista físico y psicológico.

La inteligencia artificial en la prevención de los delitos

La inteligencia artificial en los últimos años se ha introducido en la actividad policial y de la seguridad con el uso de datos estadísticos para asistir en la toma de decisiones y en l prevención del delitos lo cual se conoce como el fenómeno del "predictive policing" o policía predictiva, definido por Martin Degeling, como aquella "variedad de técnicas utilizadas por los cuerpos policiales

para generar probabilidades delictivas, a menudo denominadas predicciones, y actuar en consecuencia" (Monforte, 2022).

Con estos sistemas se procura optimizar recursos, e incrementar la efectividad de la labor policial en la prevención de delitos. Las herramientas de inteligencia artificial analizan los datos históricos estadísticos que constan en las bases de datos policiales para predecir en qué áreas geográficas hay una mayor probabilidad de actividad criminal, qué perfiles de personas tienen más posibilidades de delinquir en el futuro, o qué tipo de gente cuenta con mayor predisposición a ser víctima de un delito, entre otros.

La Agencia de los derechos fundamentales de la Unión Europea (2021) establece mediante un cuadro comparativo con la investigación tradicional sobre las ventajas de la inteligencia artificial predictiva, el contexto en el que se desenvuelve es cuando aún no se ha cometido ningún delito, es proactiva, su objetivo es prever dónde y cuándo pueden cometerse delitos o contra quién, los datos que utiliza la inteligencia artificial predictiva contienen información genérica proveniente de varios casos y se basa fundamentalmente en procesos que involucran gran cantidad de datos. (Unión Europea, 2021)

Se configura en distintas variantes, están aquellos que tienen como objetivo pronosticar los lugares y momentos en que se presenta un mayor riesgo de comisión de delito, de otro lado,, los que permiten predecir identidades delictivas, que tienen por misión la creación de perfiles delictivos, con carácter general, para identificar a los posibles delincuentes del futuro, los que permiten predecir víctimas de delitos, que tienen como propósito la identificación de aquellos individuos o grupos de individuos que es más probable que resulten víctimas de un delito; y los que permitente predecir delincuentes, que tienen como finalidad identificar el riesgo concreto de que un determinado individuo delinca en el futuro (Monforte, 2022).

Existen dos sistemas de predicción: los sistemas place- based y los sistemas person-based" (Fluja, 2022). El sistema o sistemas "place-based", representa el primer nivel, o nivel menos invasivo con los derechos de las personas, de la policía predictiva. En esencia es el análisis de los datos disponibles para tratar de establecer en qué lugares y momentos es más probable que se cometa un delito, a lo que se puede añadir, generalmente, la referencia a un tipo o tipos concretos de delitos.

Puede decirse que hay tres elementos comunes a todos los sistemas "place-based", el primero, ya visto, es que se focalizan en predecir lugares, tiempos y tipos de crímenes, y los otros dos son, respectivamente, que reconocen que el riesgo criminal tiene conexión con las vulnerabilidades sociales, y que creen que la interacción criminal puede reducir las tasas de delincuencia. Ello puede conducir a la estigmatización de determinados lugares y poblaciones y a etiquetarlos como marginales.

Los sistemas "person-based", se dirigen a predecir de forma individualizada y concreta qué personas van a cometer delitos con más alto grado de probabilidad. En este hay mayores motivos de preocupación, porque las herramientas predictivas pueden representar amenazas de alto grado para los derechos y garantías de las personas, amenaza que puede extenderse luego, en ciertos casos, a las propias de un proceso penal. Tales sistemas pueden generar discriminaciones grupales o individuales, más cuando ofrecen la posibilidad de predecir posibles colectivos con tendencias mayores a cometer delitos o personas pertenecientes a esos colectivos en términos de raza, etnia, o cualquier otra característica específica identificadora de unos colectivos frente a otros.

Este sistema entraña mayores riesgos por la invasión en la esfera de la protección de datos, con falsos positivos y la posible criminalización de personas, además genera perfilados criminales no siempre efectivos. A ello se agrega la sospecha que se genera

a partir de una multiplicidad de datos afectando las garantías de los derechos en el proceso penal, y en primer término sobre el principio de presunción de inocencia-

Ejemplos de IAP es el algoritmo ROSSMO, cuya formulación permite estimar el área geográfica donde, con mayor probabilidad, reside un presunto agresor serial en función de la ubicación de los delitos que previamente se le atribuyen, este permite priorizar sospechosos y concentrar los recursos policiales en determinados individuos que encajan con la geografía del crimen. El instrumento The Lethality Screen, se están utilizando de forma conjunta con herramientas de valoración del riesgo de violencia doméstica para proporcionar información útil a los equipos de intervención temprana. En España, se ha desarrollado el sistema de seguimiento integral VioGén para, a través de un algoritmo que analiza la información actuarial disponible en el protocolo de valoración policial del riesgo (VPR) tratar de identificar aquellos casos en los que la victimización por violencia doméstica es más probable (Miró Llinares, 2018).

Otra herramienta de inteligencia artificial para minimizar la actividad delictiva, es COPKIT, desarrollado entre 2021 y 2023, cuyo objetivo es desarrollar, con base en el marco legal vigente, aquellas herramientas de inteligencia artificial que empleará la policía del futuro, sin vulnerar los principios de libertad, igualdad y justicia, centrándose principalmente en el objetivo de analizar, investigar, mitigar y prevenir el uso de las nuevas tecnologías de la información y la comunicación por parte del crimen organizado y los grupos terroristas.

COMPAS es una herramienta de inteligencia artificial utilizada con el mismo fin de PRISMA en Estados Unidos. COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) fue desarrollada por la empresa estadunidense Northpointe Inc, es una herramienta de inteligencia artificial de Sistemas de Análisis de Evaluación de Riesgo, la cual predice la probabilidad de que

un reo pueda reincidir en un delito. Su importancia en el debate jurídico y ético que se da respecto a estos sistemas no se debe precisamente a su exactitud y eficiencia su desarrollo en las cortes de Estados Unidos, sino a su baja fiabilidad y sesgos raciales. Pro Publica, un medio investigativo de Estados Unidos descubrió la poca fiabilidad que tenía este sistema partiendo del análisis de casos en los cuales había sido utilizada (Sarmiento, 2024).

No obstante se cuestiona la efectividad de estos sistemas, y el impacto que tiene en el uso de los datos personales, la colisión con otros derechos fundamentales como la presunción de inocencia, la protección de los datos personales, la libertad y la no discriminación, la presencia de datos sesgados que influyen en la discriminación de determinados grupos de personas, el uso desmedido de estas herramientas por gobiernos y cuerpos de seguridad pueden traer aplicaciones injustas y desprovistas de ética.

Proceso penal e inteligencia artificial

En el ámbito del derecho procesal, los sistemas de inteligencia artificial revisten utilidad en varios campos. En primer término, con la descongestión de la actividad judicial al optimizar la redacción de escritos. También incide en la tramitación y gestión, en la actividad de archivo y documentación, notificaciones, actos de citación y comunicación, la verificación de la integridad documental, las actas de vistas y audiencias, el tratamiento de los datos, incidiendo en la realización del derecho a la tutela judicial efectiva.

Otra de las aplicaciones en el derecho penal está vinculada al derecho probatorio, así la inteligencia artificial puede ser aplicada en la revisión de los parámetros de valoración de la prueba, la elaboración de hipótesis y, eventualmente, la concreción con menor subjetividad de los llamados “estándares probatorios”.

Una tercera aplicación está relacionada con la predicción de riesgos, lo que determina la configuración del perfil del imputado, si es posible que reincida en la comisión de delitos, llegando el juez incluso a una convicción razonable de culpabilidad a veces permeado de excesivos criterios subjetivos y la experiencia del juez que puede a veces estar cargada de perjuicios. A este tenor una herramienta concebida para esta finalidad es COMPAS una auténtica herramienta de inteligencia artificial que asiste, no sustituye a los jueces en la predicción del riesgo a fin de evaluar este factor en un reo a efectos de imponer medidas cautelares, decisiones de libertad en el ámbito penitenciario o incluso como acompañamiento de juicios de culpabilidad (Nieva-Fenoll, 2022).

Sin embargo, en cualquiera de sus usos, la tecnología puede estar sujeta a sesgos discriminatorios a veces son evidentemente racistas, xenófobos o contrarios a la diversidad de género, los datos que conforman el algoritmo pueden estar sesgados y tener resultados contrarios a los deseados por el sistema de justicia. La técnica se nutre de datos existentes en el pasado y provistos del mismo trato discriminatorio debiendo superarse esta barrera y trabajar los desarrolladores en algoritmos que respeten los principios límites de inteligencia artificial.

3.7. La inteligencia artificial y la protección de datos personales

Los sistemas de inteligencia artificial requieren para su desarrollo del suministro de datos, estos permiten como puerta de entrada conformar algoritmos para su funcionalidad, se requiere entrenar modelos que puedan reconocer patrones, hacer predicciones o tomar decisiones, la inteligencia artificial de involucra la recolección, el almacenamiento, análisis y procesamiento o interpretación de gran cantidad de datos (big data), aplicada para generar resultados que impliquen aprendizajes de los sistemas y por consiguiente determinados comportamiento de las maquinas.

No siempre los datos que se emplean para diseñar modelos de inteligencia artificial son personales, se estos incluyen información personal, lo que plantea preocupaciones sobre la privacidad y la seguridad de los individuos.

Los datos personales aparecen conceptualizados en los Estándares de protección de datos personales de la Red Iberoamericana de Protección de Datos (2017), como cualquier información concerniente a una persona física identificada o identificable, expresada en forma numérica, alfabética, gráfica, fotográfica, alfanumérica, acústica o de cualquier otro tipo. Se considera que una persona es identificable cuando su identidad pueda determinarse directa o indirectamente, siempre y cuando esto no requiera plazos o actividades desproporcionada.

Los estándares supra mencionados, conceptualizan los datos personales sensibles como aquellos que se refieren a la esfera intima por el su titular o cuya utilización indebida pueda dar origen a discriminación o conlleva un riesgo grave para este, estos pueden conllevar aspectos como origen racial o étnico, creencias o convicciones religiosa, filosóficas y morales, afiliación sindical, opiniones políticas, datos relativos a la salud, ala vida, preferencias u orientación sexual, datos genéticos o datos biométricos dirigidos a identificar de manera univoca a una persona física.

En la Declaración relativa a la ética y protección de datos en inteligencia artificial se reconoce el vínculo entre la recolección, uso y revelación de información personal y el desarrollo de ciertas aplicaciones de inteligencia artificial, reconociendo que los datos son una categoría jurídica, no todo su desarrollo implica el tratamiento de los datos personales y los datos personales no son la única fuente de información que se recolecta y usa para estos efectos.

El uso de los datos personales implica riesgos de vulnerabilidad de los derechos humanos afectándose la protección de los mismos,

entrando en contradicción con los principios reconocidos en esta materia. La recopilación masiva de los datos personales que exige la inteligencia artificial pone en riesgo la seguridad y privacidad de los datos estos pueden emplearse de manera indebida y filtrarse, los algoritmos de inteligencia artificial pueden estar sesgados debido a los datos utilizados para que sean entrenados, en el procesamiento de datos en función de la inteligencia artificial puede haber falta de transparencia que genera la desconfianza de los titulares de los datos.

Los Estándares relacionados arriba, establecen que los principios son:

Legitimación, artículo 11: el responsable solo podrá tratar datos personales cuando se presente alguno de los siguientes supuestos: a. El titular otorgue su consentimiento para una o varias finalidades específicas. b. El tratamiento sea necesario para el cumplimiento de una orden judicial, resolución o mandato fundado y motivado de autoridad pública competente. c. El tratamiento sea necesario para el ejercicio de facultades propias de las autoridades públicas o se realice en virtud de una habilitación legal. El tratamiento sea necesario para el reconocimiento o defensa de los derechos del titular ante una autoridad pública. e. El tratamiento sea necesario para la ejecución de un contrato o precontrato en el que el titular sea parte. f. El tratamiento sea necesario para el cumplimiento de una obligación legal aplicable al responsable. g. El tratamiento sea necesario para proteger intereses vitales del titular o de otra persona física. h. El tratamiento sea necesario por razones de interés público establecidas o previstas en ley. i. El tratamiento sea necesario para la satisfacción de intereses legítimos.

Principio de licitud, artículo 14: El responsable tratará los datos personales en su posesión con estricto apego y cumplimiento de lo dispuesto por el derecho interno del Estado Iberoamericano que resulte aplicable, el derecho internacional y los derechos y libertades de las personas.

Principio de lealtad, artículo 15: El responsable tratará los datos personales en su posesión privilegiando la protección de los intereses del titular y absteniéndose de tratar éstos a través de medios engañosos o fraudulentos. Se considerarán desleales aquellos tratamientos de datos personales que den lugar a una discriminación injusta o arbitraria contra los titulares.

Principio de transparencia, artículo 16: El responsable informará al titular sobre la existencia misma y características principales del tratamiento al que serán sometidos sus datos personales, a fin de que pueda tomar decisiones informadas al respecto.

Principio de finalidad, artículo 17: Todo tratamiento de datos personales se limitará al cumplimiento de finalidades determinadas, explícitas y legítimas. El responsable no podrá tratar los datos personales en su posesión para finalidades distintas a aquéllas que motivaron el tratamiento original de éstos, a menos que concurra alguna de las causales que habiliten un nuevo tratamiento de datos conforme al principio de legitimación. El tratamiento ulterior de datos personales con fines archivísticos, de investigación científica e histórica o con fines estadísticos, todos ellos, en favor del interés público, no se considerará incompatible con las finalidades iniciales.

Principio de proporcionalidad, artículo 18: El responsable tratará únicamente los datos personales que resulten adecuados, pertinentes y limitados al mínimo necesario con relación a las finalidades que justifican su tratamiento.

Principio de calidad, artículo 19: El responsable adoptará las medidas necesarias para mantener exactos, completos y actualizados los datos personales en su posesión, de tal manera que no se altere la veracidad de éstos conforme se requiera para el cumplimiento de las finalidades que motivaron su tratamiento. Cuando los datos personales hubieren dejado de ser necesarios para el cumplimiento de las finalidades que motivaron su tratamiento, el responsable los suprimirá o eliminará de sus archivos, registros, bases de datos,

expedientes o sistemas de información, o en su caso, los someterá a un procedimiento de anonimización.

Principio de responsabilidad, artículo 20: El responsable implementará los mecanismos necesarios para acreditar el cumplimiento de los principios y obligaciones establecidas en los presentes Estándares, así como rendirá cuentas sobre el tratamiento de datos personales en su posesión al titular y a la autoridad de control, para lo cual podrá valerse de estándares, mejores prácticas nacionales o internacionales, esquemas de autorregulación, sistemas de certificación o cualquier otro mecanismo que determine adecuado para tales fines.

Principio de seguridad, artículo 21: El responsable establecerá y mantendrá, con independencia del tipo de tratamiento que efectúe, medidas de carácter administrativo, físico y técnico suficientes para garantizar la confidencialidad, integridad y disponibilidad de los datos personales. Los principios antes relacionados pueden verse afectados cuaado se emplean los datos personales como insumos de inteligencia artificial.

La Ley de inteligencia artificial europea (IA Act) prohíbe sistemas de inteligencia artificial, siempre que implique la violación de los derechos fundamentales incluida la protección de datos personales y recoge la obligación de realizar Evaluación de Impacto de privacidad (PIA- Privact impact assesssment).

Las Recomendaciones establecen que dicha evaluación debe contener un descripción detallada de las operaciones de tratamiento de estos datos, una evaluación de riesgos específicos para los derechos y libertades de los titulares de los datos personales y la medidas previstas para afrontar los riesgos, incluidas garantías medidas de seguridad diseños de software, tecnologías y mecanismos que garanticen la protección de datos personales , todo ello permite aplicar de manera adecuada el principio de privacidad (Red Iberoamericana de Protección de Datos, 2019).

El uso de la inteligencia artificial representa riesgos para la privacidad y la protección de datos personales, tales como la retención y divulgación no intencionada, esta puede retener información personal empleada para el entrenamiento y en ocasiones puede ser recuperable y divulgable. Otro riesgo es la falta de transparencia, la opacidad en los procesos de toma de decisiones de la inteligencia artificial, que pueden dificultar que las personas comprendan como se utilizan sus datos. Un riesgo importante es la discriminación y sesgo si existe sesgo en los datos utilizados para entrenamiento de modelos tales sesgos discriminatorios pueden perpetuarse afectando en forma negativa a ciertas poblaciones y grupos. Finalmente, también se puede producir el acceso no autorizado a los datos, generando brechas de seguridad.

Dentro de los derechos de los titulares de los datos personales se establece que estos tienen derecho a no ser objeto de decisiones que le produzcan efectos jurídicos o le afecten de manera significativa que se basen únicamente en tratamiento automatizados sin intervención humana en determinados aspectos personales, que puedan analizar o predecir su rendimiento profesional situación económica, estado de salud, preferencia sexuales o comportamiento, salvo que exista un consentimiento demostrable del titular.

Solo se establecen tres excepciones a la prohibición del tratamiento automatizado estos son: que sea necesario para la celebración de un contrato entre titular y responsable; que este autorizado por el derecho interno de los Estados iberoamericanos, y por último, que se base en el consentimiento del titular, que debe ser demostrable.

Se demuestra que existe un interés especial por el respeto a los derechos humanos al prohibir que se empleen los datos personales con finalidad de discriminación en el artículo 29.4.

El consentimiento del titular sobre uso de datos sensibles contiene especiales prescripciones en los Estándares sobre uso de los datos personales, estos establecen que el consentimiento es según el artículo 2 inciso b) la manifestación de la voluntad, libre, específica, inequívoca e informada, del titular a través de la cual acepta y autoriza el tratamiento de los datos personales que le conciernen.

Las condiciones para obtener el consentimiento del titular previstas en el artículo 12.2) son: que el responsable demuestre de manera indubitable que el titular otorgó su consentimiento, ya sea a través de una declaración o una acción afirmativa clara. En segundo lugar, siempre que sea requerido el consentimiento para el tratamiento de los datos personales, el titular podrá revocarlo en cualquier momento, para lo cual el responsable establecerá mecanismos sencillos, ágiles, eficaces y gratuitos.

En el caso de que se trate de los datos de niños, niñas y adolescentes la obtención del consentimiento por el responsable dependerá de la autorización del titular de la patria potestad o tutela, conforme a lo dispuesto en las reglas de representación previstas en el derecho interno de los Estados Iberoamericanos, o en su caso, solicitará directamente la autorización del menor de edad si el derecho interno de cada Estado Iberoamericano ha establecido una edad mínima para que lo pueda otorgar directamente y sin representación alguna del titular de la patria potestad o tutela, tomando siempre en cuenta que prevalece el interés superior del menor.

La anonimización es una de las medidas relativas a la protección de datos, cuando estos se empleen en proyectos de inteligencia artificial, esto implica que la información que va a ser utilizada debe ir dirigida o estar asociada a una persona, se recomienda que este anonimizada, y que por ende se desconozca la identidad del titular mitigando los riesgos sobre el uso masivo de datos

personales, prevista en la Recomendación 9 (Red Iberoamericana de Protección de Datos, 2019).

Las Recomendaciones generales para el tratamiento de los datos en la inteligencia artificial, en conexión con los Estándares que se han venido valorando determinan la necesidad de garantizar la seguridad en el tratamiento de los datos personales, por lo que es fundamental adoptar medidas tecnológicas, humanas, administrativas, contractuales o de cualquier índole cumpliendo los siguientes objetivos: evitar accesos indebidos o no autorizados a la información, evitar manipulación de la información, evitar la destrucción de la información, evitar usos indebidos o no autorización de la información y evitar circular o suministrar la información a personas no autorizadas.

Los Estándares de RIDP obligan a los responsables a establecer medios y procedimientos sencillos, expeditos, accesibles y gratuitos al titular para ejercer sus derechos de acceso, rectificación, cancelación, oposición y portabilidad". Por esta razón deben prever desde el inicio del desarrollo de los sistemas de inteligencia artificial de alternativas o procedimientos para que los titulares d ellos datos puedan ejercer estos derechos.

Se establece en el artículo 43 el régimen de reclamaciones y de imposición de sanciones por lo que todo titular tendrá derecho a presentar su reclamación ante la autoridad de control aso como a recurrir a la tutela judicial pata hacer efectivos sus derechos conforma a la legislación nacional de estado donde sea aplicable. Por lo que puede presentar la reclamación ante la autoridad administrativa o de control e incluso solicitar tutela judicial.

Los Estándares señalan que la legislación nacional establecerá un régimen que permita la adopción de medidas correctivas y sancionar las conductas que contravengan lo dispuesto en las legislaciones nacionales correspondientes indicando al menos el límite máximo y los criterios objetivos para fijar las correspondientes sanciones, a partir de la naturaleza, gravedad, duración de la infracción y sus consecuencias.

REFERENCIAS

Agencia de los Derechos Fundamentales de la Unión Europea. (2021). Construir correctamente el futuro, la inteligencia artificial y los derechos fundamentales. https://fra.europa.eu/sites/default/files/fra_uploads/fra-2021-artificial-intelligence-summary_es.pdf

AlgorithmWatch. (2025). AlgorithmWatch is a non-profit research and advocacy organization that is committed to watch, unpack and analyze automated decision-making (ADM) systems and their impact on society. https://algorithmwatch.org/en/

Álvarez Villajuana, D. T. (2025). Importancia de la informática jurídica y la inteligencia artificial en el derecho. *Revista Iberoamericana de Derecho, Cultura y Ambiente, (7).* https://aidca.org/ridca7-cyc-alvarez-villajuana-importancia-de-la-informatica-juridica-y-la-inteligencia-artificial-en-el-derecho/

Arana Hernández, A. (2023). Criptoanálisis de las máquinas ENIGMA [Trabajo de Fin de Grado, Universidad Politécnica de Madrid].

Arellano Toledo, W. (2019). El derecho a la transparencia algorítmica en Big Data e inteligencia artificial. Revista General de Derecho Administrativo, 50, https://www.iustel.com/v2/revistas/detalle_revista.asp?id_noticia=421167

Aristóteles. (2016). Ética a Nicómaco. Imprenta Nacional.

Arriagada-Bruneau, G. (2024). Una mirada crítica a la ética de la IA: de preocupaciones emergentes y principios orientadores a un desvelar ético. Resonancias, (17), 101-120. https://doi.org/10.5354/0719-790X.2024.74438

Asimov, I. (2004). Yo, robot. Edhasa.

Ayo Ferrándiz, C., Seijo Bar, Á., Garre Anguera de Sojo, I., & González Guillén, P. (2025). Responsabilidad civil e inteligencia artificial. ***Actualidad Jurídica Uría Menéndez, 67***, 29-56. https://www.uria.com/documentos/publicaciones/9326/documento/AJUM_67-2025.pdf?id=14010&forceDownload=true

Banco Interamericano de Desarrollo. (2024). Reporte de tecnología: Inteligencia artificia. https://publications.iadb.org/publications/spanish/document/Reporte-de-tecnologia-inteligencia-artificial.pdf

Barrio Andrés, M. (2018). ***Robótica, derecho e inteligencia artificial.*** Instituto Real Elcano, (103), 7–28. https://media.realinstitutoelcano.org/wp-content/uploads/2021/11/ari103-2018-barrioandres-robotica-inteligencia-artificial-derecho.pdf

Barrio Andrés, M. (2021). Génesis y desarrollo de los derechos digitales / Origin and development of digital rights. ***Revista de las Cortes Generales,*** (110), 197–233. https://doi.org/10.33426/rcg/2021/110/1572

Benitez Vázquez, F. (2025). Decisiones Inteligentes. La Aplicación de la Inteligencia Artificial en la Toma de Decisiones Gubernamentales. Encrucijada Revista Electrónica Del Centro De Estudios En Administración Pública, (49), 19–31. https://doi.org/10.22201/fcpys.20071949e.2025.49.90109

BlogThinkBig. (2017). Neuromarketing: cómo los datos influyen en tus emociones. ***BlogThinkBig.*** https://blogthinkbig.com/el-neuromarketing-influye-en-tus-emociones

Boddington, P. (2020). Normative modes: Codes and standards. En M. Dubber, F. Pasquale, & S. Das (Eds.), ***The Oxford handbook of ethics of AI*** (pp. 124–140). Oxford University Press.

Breceda Pérez, J. A., & Castillo Lara, C. (2023). Derecho y ciencia: Entre la dignidad humana y la inteligencia artificial. *Ius et Scientia, 9*(2), 261–287. https://doi.org/10.12795/IESTSCIENTIA.2023.i02.12

Calo, R. (2015). Robotics and the lessons of cyberlaw. *California Law Review, 103*(3), 513–564. https://digitalcommons.law.uw.edu/faculty-articles/23

Campoverde Ortiz, F. J., & Lujan Johnson, G. L. (2025). Gobierno abierto comunicacional en contextos de crisis: Un modelo teórico integrador para gobiernos locales. *Maestro y Sociedad, 22*(2), 1537–1551. https://maestroysociedad.uo.edu.cu/index.php/MyS/article/view/6983

Campuzano Gallegos, A. (2019). *Inteligencia artificial para abogados.* Thomson Reuters.

Cárcar Benito, J. E. (2019). La inteligencia artificial (IA): Aplicación jurídica y regulación en los servicios de salud. *Derecho y Salud, 29*(Número extra 1), 278–290. https://www.ajs.es/sites/default/files/2020-10/vol29Extra_02_14_Comunicación.pdf

Cárdenas Gallagher, D. A., & Gómez Gutiérrez, V. (2020). *Responsabilidad civil extracontractual de la inteligencia artificial en el régimen jurídico colombiano: ¿Necesidad de un cambio normativo?* [Tesis de grado, Pontificia Universidad Javeriana].

Carrascosa Baeza, J. M. (2018). Ciencia, ética y el derecho humano a la ciencia. *Papeles de Relaciones Ecosociales y Cambio Global,* (142), 61–70. https://dialnet.unirioja.es/descarga/articulo/6524889.pdf

Castrillón Gómez, O. D., Rodríguez Córdoba, M. P., & Leyton Castaño, J. D. (2008, abril 21–25). Ética e inteligencia artificial: ¿Necesidad o urgencia? [Ponencia]. ***Congreso Internacional de Sistemas Cognitivos y Sistemas Inteligentes.*** La Habana, Cuba.

Cerrillo i Martínez, A. (2019). Com obrir les caixes negres de les administracions públiques? Transparència i rendició de comptes en l'ús dels algoritmes. ***Revista Catalana de Dret Públic,*** 58, 13–28. https://dialnet.unirioja.es/servlet/articulo?codigo=7005054

Chavarry Garrido, L. Y., Zamora Vargas, E. V., & Terrones Chávez, R. O. (2025). Uso de inteligencia artificial, drones y biometría en las operaciones policiales: una revisión sistemática de sus implicaciones éticas y legales. Revista Escpogra PNP , 4(2), 45–61. https://doi.org/10.59956/escpograpnpv4n2.4

Chávez Valdivia, A. K. (2021). Hacia otra dimensión jurídica: El derecho de los robots. ***Revista IUS, 15***(48), 325–337. https://doi.org/10.35487/rius.v15i48.2021.681

Combarros Merino, R. (2023). ***Vehículos autónomos e inteligencia artificial: Responsabilidad civil y productos defectuosos*** [Tesis de maestría, Universidad de León].

Comisión Europea para la Eficiencia de la Justicia. (2018). ***Carta ética europea sobre el uso de la inteligencia artificial en los sistemas judiciales y su entorno.*** https://protecciondata.es/wp-content/uploads/2021/12/Carta-Etica-Europea-sobre-el-uso-de-la-Inteligencia-Artificial-en-los-sistemas-judiciales-y-su-entorno.pdf

Comisión Europea. (2020). ***Libro blanco sobre la inteligencia artificial: Un enfoque europeo orientado a la excelencia y la confianza***. https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX%3A52020DC0065

Concha Flores, L. F. (2024). Inteligencia artificial, enfoque de riesgos y responsabilidad civil: Aspectos centrales para una razonabilidad práctica. Sapientia Iuris, 1, 140-169. https://doi.org/10.5281/zenodo.14735552

Consejo de Europa. (2024). *Convenio Marco del Consejo de Europa sobre inteligencia artificial, derechos humanos, democracia y Estado de Derecho.* Bruselas. https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX%3A52024PC0264

Cortina Orts, A. (2000). *Ética mínima: Introducción a la filosofía práctica.* Tecnos.

Cortina Orts, A. (2013). *¿Para qué sirve realmente la ética?* Paidós.

Cortina Orts, A. (2019). Ética de la inteligencia artificial. *Anales de la Real Academia de Ciencias Morales y Políticas,* 96, 379–394. https://dialnet.unirioja.es/servlet/articulo?codigo=7426666

Cotino Hueso, L. (2019). Ética en el diseño para el desarrollo de una inteligencia artificial, robótica y big data confiables y su utilidad desde el derecho. *Revista Catalana de Dret Públic,* (58), 29-48. https://dialnet.unirioja.es/servlet/articulo?codigo=7005055&orden=0&info=link

Council of Europe. (2024). *Framework Convention No. 225 (CETS-225) on Artificial Intelligence and Human Rights, Democracy and the Rule of Law.* https://www.coe.int/en/web/artificial-intelligence/the-framework-conventionon-artificial-intelligence

Cruz, G., Riobó, A., Pfeifer, M., & Duarte, D. (2024). *IA desde los cimientos: Desafíos y oportunidades en el contexto de América Latina y el Caribe.* https://doi.org/10.18235/0013275

Cuatrecasas Monforte, C. (2022). *La inteligencia artificial en el proceso penal de instrucción español: Posibles beneficios y potenciales riesgos* [Tesis doctoral, Universitat Ramon Llull].

De Zan, J. (2004). *La ética, los derechos y la justicia*. Mastergraf.

Delgado Díaz, C. J. (2021). Reconfigurar la bioética ante la complejidad tecnológica y social. En S. Vidal (Coord.), Manual de educación en bioética (Vol. 1) (pp.41-49). UNAM UNESCO.

Delgado Díaz, C. J. (2023). Bioética y filosofía de la tecnología: Emergencias biopolíticas. En J. R. Acosta Sariego (Ed.), *Bioética y biopolítica* (pp. 123–145). Félix Varela.

Domingues Villarroel, M. P. (2021). *La responsabilidad civil derivada del uso de la inteligencia artificial* [Trabajo Fin de Grado, Universidad de La Laguna].

Drnas de Clément, Z. (2022). Inteligencia artificial en el derecho internacional, Naciones Unidas y Unión Europea. *Revista de Estudios Jurídicos, (22)*, 8–12. https://doi.org/10.17561/rej.n22.7542

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., & Galanos, V. (2019). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management, 57*. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

España Pérez, J. A. (2024). El uso de la IA en el control del tráfico y sus implicaciones en la protección de datos. En P. Valcárcel Fernández & F. L. Hernández González (Coords.), *El derecho administrativo en la era de la inteligencia artificial* (pp. 371–381). Instituto Nacional de Administración Pública.

Europagune Research Group. (2023). La inteligencia artificial y el derecho fundamental a la protección de datos de carácter personal. *Investigaciones Jurídicas, 30*(1), 112–130. https://www.ehu.eus/en/web/europagune/liburuen-kapituluak/-/asset_publisher/OLsbCGu6uXy8/content/la-inteligencia-artificial-y-el-derecho-fundamental-a-la-proteccion-de-datos-de-caracter-personal

European Union. (2024). ***AI Act: Shaping Europe's digital future***. European Union. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Fernández Carballo-Calero, P. (2021). ***La propiedad intelectual de las obras creadas por inteligencia artificial***. Aranzadi.

Fernández Gil, I. (2021). ***Inteligencia artificial y el concepto de inventor en el sistema de patentes*** [Tesis doctoral, Universidad de Alicante].

Figueroa Jácome, L., & Barrios Lira, J. C. (2014). ¿Es necesaria la ética? ***Ética Judicial Cuaderno,4***(3), 26–48. https://eticayvalores.poderjudicial.go.cr/images/ConsejoNotables/Cuadernos/cuaderno04.pdf

Fitsilis, F., von Lucke, J., & De Vrieze, F. (2024). ***Directrices para el uso de la IA en los parlamentos europeos***. Westminster Foundation for Democracy. https://www.wfd.org/sites/default/files/2025-04/wfd-ai-guidelines-for-parliaments-2024-spanish.pdf

Flores Salgado, L. L., & Martínez Xelhuantzi, D. A. (2024). La inteligencia artificial y el Derecho: Una mirada a un futuro presente. ***Revista de Investigación en Ciencias Jurídicas, (57), 4-21.*** https://dialnet.unirioja.es/descarga/articulo/9889644.pdf

Flórez-Acero, G. D., Salazar, S., Durán, M. A., Rodríguez-Flórez, J. C., & Sierra-Marulanda, O. R. (2017). ***Propiedad intelectual, nuevas tecnologías y derecho del consumo: Reflexiones desde el moderno derecho privado***. Editorial Universidad Católica de Colombia.

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People: An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Fluja, V. C. (2022). Ideas para un debate sobre la predicción del crimen. En S. C. López (Coord.), *Inteligencia artificial legal* (pp. 289–338). Aranzadi.

Future of Life Institute. (2017). Principios de IA de Asilomar. https://futureoflife.org/es/open-letter/ai-principles/

Gabriel, M. (2021). *Ética para tiempos oscuros: Valores universales para el siglo XXI* (G. García, Trad.). Pasado & Presente.

Galindo Ayuda, F. (2019). Inteligencia artificial y acceso a documentación jurídica: Sobre el uso de las TIC en la práctica jurídica. *Democracia Digital e Governo Eletrônico, 1*(18), 144–166. https://buscalegis.ufsc.br/revistas/index.php/observatoriodoegov/article/view/319

Gamero Casado, E. (2021). El enfoque europeo de inteligencia artificial. *Revista de Derecho Administrativo – CDA, 20*, 268–289. https://dialnet.unirioja.es/servlet/articulo?codigo=8510535

González Moreno, L. (2023). *Inteligencia artificial y derecho: Retos jurídicos* [Trabajo de fin de grado, Universidad Pontificia Comillas].

Goñi Sein, J. L. (2019). Innovaciones tecnológicas, inteligencia artificial y derechos humanos en el trabajo. *Documentación Laboral,* (117), 57–72. https://dialnet.unirioja.es/servlet/articulo?codigo=7095888

Grañó Calvete, M. (2024). ***La brecha de género en la era de la inteligencia artificial.*** OBS Business School. https://marketing.onlinebschool.es/Prensa/Informes/Informe%20OBS%20Brecha%20de%20genero%20en%20la%20era%20IA.pdf

Grigore, A. E. (2022). Derechos humanos e inteligencia artificial. ***Ius et Scientia, 8***(1), 164–175. https://doi.org/10.12795/IETSCIENTIA.2022.i01.10

Guerrero Arévalo, W. D. (2021). Los alcances de la inteligencia artificial (IA) y su responsabilidad frente al derecho y la ética [Trabajo de investigación, ***Universidad Libre].***

Guzmán, D., Guachilema, A., & Cataña, E. (2025). Ecuador: Tres iniciativas de regulación de la IA. ***Universidad Adolfo Ibáñez.*** https://centrocompetencia.com/ecuador-tres-iniciativas-diferentes-de-regulacion-de-la-ia/

Haenlin, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. ***California Management Review, 61***(3), 5–14. https://doi.org/10.1177/0008125619864925

Han, B.-C. (2012). ***La sociedad del cansancio.*** Herder.

Healthvisors. (2020). ***IDx-DR: Sistema autónomo de inteligencia artificial para detección precoz de retinopatía diabética.*** Healthvisors. https://www.healthvisors.com/en/idx-dr/

Henman, P. (2020). Improving public services using artificial intelligence: Possibilities, pitfalls, governance. ***Asia Pacific Journal of Public Administration, 42***(1), 1–13. https://doi.org/10.1080/23276665.2020.1816188

High-Level Expert Group on Artificial Intelligence. (2019). ***A definition of AI: Main capabilities and scientific disciplines.*** European Commission. https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

HP Tech Takes. (2024). Vehículos autónomos: el camino de la IA hacia un futuro sin conductor. https://www.hp.com/cl-es/shop/tech-takes/vehiculos-autonomos-e-inteligencia-artificial

Ienca, M., & Malgieri, G. (2021). Mental data protection and the GDPR. *Journal of Law and the Biosciences, 9*(1). https://doi.org/10.1093/jlb/lsac006

International Business Machines Corporation. (2024). *What is artificial intelligence (AI)?* IBM. https://www.ibm.com/think/topics/artificial-intelligence

Iturmendi Morales, G. (2020). Responsabilidad civil por el uso de sistemas de inteligencia artificial. https://laleydigital.laleynext.es/Content/Documento.aspx?params=H4sIAAAAAAAEAMtMSbF1CTEAAmNDCwsjI7Wy1KLizPw8WyMDIwNDQyMDkEBmWqVLfnJIZUGqbVpiTnEqANKvjtI1AAAAWKE

Jalomo-Aguirre, F. (2012). Sobre las leyes y su dimensión ética. *Vniversitas, 61*(124), 147–167. https://doi.org/10.11144/Javeriana.vj61-124.slde

Januário, T. F. (2023). Inteligencia artificial y responsabilidad penal de personas jurídicas: Un análisis de sus aspectos materiales y procesales. *Estudios Penales y Criminológicos, 40.* https://doi.org/10.15304/epc.44.8902

Kant, E. (2003). *Fundamentación de la metafísica de las costumbres*. Encuentro.

Kelsen, H. (1982). *¿Qué es la justicia?* Planeta-Agostini.

Kokane, S. (2021). The intellectual property rights of artificial intelligence-based inventions. *Journal of Scientific Research, 65*(2), 116-118. https://www.bhu.ac.in/research_pub/jsr/Volumes/JSR_65_02_2021/23.pdf

Kouvchinova, A., & Voronine, M. (2024). *Análisis y aplicación de técnicas de machine learning* [Trabajo de fin de grado, Universidad de Cantabria].

Manet, L. (2014). Modelos de desarrollo regional: Teorías y factores determinantes. *Nóesis. Revista de Ciencias Sociales y Humanidades, 23*(46), 18-56. https://www.redalyc.org/pdf/859/85930565002.pdf

Martínez Bahena, G. C. (2012). La inteligencia artificial y su aplicación al campo del Derecho. Alegatos, (82), 827-846. https://alegatos.azc.uam.mx/index.php/ra/article/view/205

Martínez Martínez, R. (2019). Inteligencia artificial desde el diseño: Retos y estrategias para el cumplimiento normativo. *Revista Catalana de Dret Públic, 58*, 64-81. https://doi.org/10.2436/rcdp.i58.2019.3317

Martínez, M. V. (2024). De qué hablamos cuando hablamos de inteligencia artificial. *UNESCO*. https://www.unesco.org/es/articles/de-que-hablamos-cuando-hablamos-de-inteligencia-artificial

Medina Peña, R., Valarezo Román, J., & Romero Romero, C. D. (2021). Fundamentos epistemológicos del neoconstitucionalismo Latinoamericano. Aciertos y desaciertos en su regulación jurídica y aplicación práctica en Ecuador. Sociedad &Amp; Tecnología, 4(S1), 213–225. https://doi.org/10.51247/st.v4iS1.130

McCarthy, J., Shannon, C. E., Minsky, M. L., & Rochester, N. (1955). *A proposal for the Dartmouth Summer Research Project on Artificial Intelligence.* The Dartmouth Summer Research Project on Artificial Intelligence. http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf

Medina Rojas, C. (2025). Robótica, inteligencia artificial y derecho: nuevas dimensiones jurídicas en el siglo XXI. *Ciencia Latina, Revista Científica Multidisciplinar, 9*(1), 9488–9513. https://doi.org/10.37811/cl_rcm.v9i1.16574

Mellado, R., Escobar, E., De la Mora, H., Díaz, E. J., & Rodríguez, F. L. (2023). Estudio de concordancia entre el sistema Watson for Oncology y la práctica clínica en pacientes con cáncer de mama dentro del Hospital Ángeles Pedregal. *Acta Médica Grupo Ángeles, 21*(4), 338-342. https://doi.org/10.35366/112643

Meza Ruiz, I. V. (2024). El camino hacia una regulación de la IA en México. https://turing.iimas.unam.mx/~ivanvladimir/posts/camino_regulacion_ia_mexico/

Ministerio de Ciencia y Tecnología de la República Popular China. (2021). *Código de ética de la inteligencia artificial de nueva generación.* https://www.most.gov.cn/kjbgz/202109/t20210926_177063.html

Miró Llinares, F. M. (2018). Inteligencia artificial y justicia penal: Más allá de los resultados lesivos causados por robots. *Revista de Derecho Penal y Criminología, 20*, 87-130. https://doi.org/10.5944/rdpc.20.2018.26446

Mitelman, C. Z. (2024). Inteligencia artificial en el ámbito del derecho de la salud. *Revista Derecho y Salud, 8*(9). https://revistas.ubp.edu.ar/index.php/rdys/article/view/494

Molina, E., & Medina, E. (2025). *La revolución de la inteligencia artificial en la educación superior: Lo que hay que saber.* Banco Mundial. https://documents1.worldbank.org/curated/en/099809404152514027/pdf/IDU-91d6e888-fcbd-4694-ac88-18bcae998934.pdf

Morandín-Ahuerma, F. (2023). Principios normativos para una ética de la inteligencia artificial. *Consejo de Ciencia y Tecnología del Estado de Puebla.* https://philarchive.org/archive/MORDDM-2v1

Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism.* PublicAffairs.

Munn, L. (2023). The uselessness of AI ethics. *AI and Ethics,* 3, 869–877. https://doi.org/10.1007/s43681-022-00209-w

Muñoz Vela, J. M. (2021). ***Derecho de la inteligencia artificial: Un enfoque global de responsabilidad desde la ética, la seguridad y las nuevas propuestas reguladoras europeas*** [Tesis doctoral, Universidad de Valencia].

Naranjo Godoy, L. (2024). Proyecto de política para el desarrollo y fomento del uso ético y responsable de la inteligencia artificial en el Ecuador – 2024. ***El Comercio.*** https://www.elcomercio.com/opinion/proyecto-politica-inteligencia-artificial-lorena-naranjo-columnista.html

Nieva-Fenoll, J. (2022). Inteligencia artificial y proceso judicial. En S. C. Arjona (Ed.), ***Inteligencia artificial legal y administración de justicia*** (pp. 417-435). Aranzadi.

Nodals García, C. L. (2021). Sobre la necesidad de unificación de las iniciativas para un uso ético de la inteligencia artificial. ***Socialium: Revista Científica de Ciencias Sociales y Humanidades,*** 5(2), 318-334. https://doi.org/10.26490/uncp.sl.2021.5.2.880

Oremus, W. (2025). Senators revive bill to restrict TSA facial recognition at airports. ***The Washington Post.*** https://www.washingtonpost.com/politics/2025/05/08/tsa-face-recognition-airports-privacy-merkley/

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2003). ***Declaración de Principios.*** Cumbre Mundial sobre la Sociedad de la Información. https://www.itu.int/net/wsis/outcome/booklet/declaration_A-es.html

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2021). ***Recommendation on the ethics of artificial intelligence.*** UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137-spa/

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2023). ***Libertad de expresión, inteligencia artificial y elecciones.*** UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000393473_spa

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2025). ***Inteligencia artificial en la educación.*** UNESCO. https://www.unesco.org/es/digital-education/artificial-intelligence

Organización de las Naciones Unidas. (2019a). ***Derechos humanos en la era digital.*** ONU. https://www.ohchr.org/es/2019/10/human-rights-digital-age

Organización de las Naciones Unidas. (2021). ***El derecho a la privacidad en la era digital*** (Informe de la Alta Comisionada de las Naciones Unidas). ONU. https://docs.un.org/es/A/HRC/48/31

Organización de las Naciones Unidas. (2023). ***La seguridad de la infancia y la juventud en la red.*** https://www.un.org/es/global-issues/child-and-youth-safety-online

Organización de las Naciones Unidas. (2024a). ***Principios rectores y funciones de la gobernanza internacional de la IA.*** https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_es.pdf

Organización de las Naciones Unidas. (2024b). ***Informe de la Relatora Especial sobre las formas contemporáneas de racismo, discriminación racial, xenofobia y formas conexas de intolerancia, Ashwini K. P.*** https://docs.un.org/es/A/HRC/56/68

Organización de las Naciones Unidas. (2025). Declaración sobre inteligencia artificial inclusiva y sostenible para las personas y el planeta. *Cumbre de Acción sobre IA*. https://onu.delegfrance.org/IMG/pdf/02_11_-_statement_on_inclusive_and_sustainable_artificial_intelligence_for_people_and_the_planet_002_.pdf

Ospina Díaz, M. R., Mora Pabón, R., & Maya Ceballos, A. B. (2025). Percepción de la inteligencia artificial en la lucha contra la corrupción: Una exploración al caso del Estado de Colombia. *Revista Opera, (36)*, 7-45. https://dialnet.unirioja.es/descarga/articulo/9852203.pdf

Palencia, O. (2020). Ética y derecho: Aciertos y desaciertos. *Revista Utopía y Praxis Latinoamericana, 25(91)*, 217-231. https://www.redalyc.org/journal/279/27965041021/html/

Price, W. N., Gerke, S., & Cohen, I. G. (2024). Liability for use of artificial intelligence in medicine. En B. Solaiman & I. G. Cohen (Eds.), *Research handbook on health, AI and the law* (cap. 9). Edward Elgar Publishing Ltd.

Puyol-Cortez, J. (2023). Tecnologías emergentes en la educación del siglo XXI. *Multidisciplinary Collaborative Journal, 1*(4), 40-55. https://doi.org/10.70881/mcj/v1/n4/25

Ramírez-García, H. S. (2008). Derecho y ética: Convergencias para la formación jurídica. *Revista Díkaion, 22*(17), 49-69. http://www.redalyc.org/articulo.oa?id=72011607004

Randstad. (2019). *Lineamientos globales*. https://www.randstad.com.mx/s3fs-media/mx/public/migration/downloads/principios-globales-ia.pdf

Real Academia Española. (2024). En *Diccionario de la lengua española* (23.ª ed.). https://dle.rae.es

Red Iberoamericana de Protección de Datos. (2017). *Estándares de protección de datos personales para los Estados iberoamericanos.* http://www.redipd.org/documento/estandares-iberoamericanos-2017.pdf

Red Iberoamericana de Protección de Datos. (2019). *Recomendaciones generales para el tratamiento de datos en inteligencia artificial.* https://www.redipd.org/sites/default/files/2020-02/guia-recomendaciones-generales-tratamiento-datos-ia.pdf

Repsol. (2025). *Inteligencia artificial o IA: qué es y cuáles sus ventajas.* https://www.repsol.com/es/energia-avanzar/innovacion/inteligencia-artificial/index.cshtml

Reyes Vásquez, P. A. (2023). Ética de la inteligencia artificial. Recomendación de la UNESCO, noviembre 2021. *Compendium, 26*(50), 1–6. https://doi.org/10.5281/zenodo.10271853

Ríos Ruiz, W. R. (2002). El derecho de autor en la protección jurídica de los programas de ordenador: soporte lógico (software) y los bancos o bases de datos. *Revista La Propiedad Inmaterial, 5*, 81–92. https://revistas.uexternado.edu.co/index.php/propin/article/view/1165

Rivera García, A. (2024). Análisis de la propiedad industrial en las invenciones patentables generadas por inteligencias artificiales. *Sapientia & Iustitia, (9)*, 137–158. https://sapientia.ucss.edu.pe/index.php/sei/article/view/analisis-propiedad-industrial-invenciones-patentables-generadas-

Russell, S., & Norvig, P. (2009). *Artificial intelligence: A modern approach* (3rd ed.). Prentice Hall.

Saiz García, C. (2019). Las obras creadas por sistemas de inteligencia artificial y su protección por el derecho de autor. *InDret, (1)*, 1–45. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3365458

Salado García, J. P. (2018). ***Inteligencia artificial en China.*** Oficina Económica y Comercial de España en Cantón. http://www.iberchina.org/files/2018/ai_china.pdf

Salardi, S. (2020). Robótica e inteligencia artificial: retos para el derecho. ***Derechos y Libertades: Revista de Filosofía del Derecho y Derechos Humanos, (42)***, 203–232. https://doi.org/10.14679/1158

Sánchez Vásquez, C., & Toro-Valencia, J. (2021). ***El derecho al control humano: Una respuesta jurídica a la inteligencia artificial***. ***Revista Chilena de Derecho y Tecnología, 10***(2), 211–228. https://doi.org/10.5354/0719-2584.2021.58745

Santana-Soriano, E., & Báez Vizcaíno, K. (2022). Ciberespacio y cibermundo: delimitaciones conceptuales desde el materialismo sistémico. ***Ciencia y Sociedad, 47***(1), 45–57. https://www.redalyc.org/journal/870/87070563004/87070563004.pdf

Sarmiento Arango, A. (2020). ***La aplicación de la inteligencia artificial en el derecho penal: consecuencias éticas y jurídicas frente a la medida de aseguramiento*** [Trabajo de grado, Universidad de los Andes].

Schulz, P. C. (2005). La ética en ciencia. ***Revista Iberoamericana de Polímeros, 6***(2), 120–156. https://reviberpol.org/wp-content/uploads/2019/08/2005-2-schulz.pdf

Schulze, C. (2023). UK IPO suspended its guidelines on AI patents after High Court judgment. ***JUVE Patent*** [Blog]. https://www.juvepatent.com/cases/uk-ipo-suspended-its-guidelines-on-ai-patents-after-high-courtjudgment/

Schwab, K. (2017). ***La cuarta revolución industrial.*** Penguin Random House.

Sebio Martín, M. (2020). ***Inteligencia artificial y ética*** [Trabajo de Fin de Grado, Universidad Pontificia Comillas].

Serna, A., Acevedo, E., & Serna, E. (2017). Principios de la inteligencia artificial en las ciencias computacionales. En E. Serna M. (Ed.), ***Desarrollo e innovación en ingeniería*** (pp. 161–172). Editorial Instituto Antioqueño de Investigación.

Sheikh, H., Prins, C., & Schrijvers, E. (2023). ***Mission AI: The new system technology.*** Springer.

SIENNA. (2024). ***Technology, ethics and human rights.*** https://www.sienna-project.eu/w/si

Silva, M. E., & Cuevas, A. C. (2024). ***El uso de la inteligencia artificial para alcanzar la seguridad vial***. Instituto Mexicano del Transporte, ***Publicación Externa***, (211). https://imt.mx/resumen-boletines.html?IdArticulo=618&IdBoletin=212

Skadden. (2025). ***Copyright Office weighs in on AI training and fair use.*** https://www.skadden.com/insights/publications/2025/05/copyright-office-report

Solar Cayón, J. I. (2020). La inteligencia artificial jurídica: nuevas herramientas y perspectivas metodológicas para el jurista. ***Revus: Revista de Teoría Constitucional y Filosofía del Derecho, (41)***, 1–27. https://doi.org/10.4000/revus.6547

Solar Cayón, J. I. (2024). La inteligencia artificial jurídica como herramienta para promover el acceso al Derecho y a servicios jurídicos básicos. ***Derechos y Libertades: Revista de Filosofía del Derecho y Derechos Humanos, 51***, 423–431. https://doi.org/10.20318/dyl.2024.8588

Taddeo, M., & Floridi, L. (2021). How AI can be a force for good: An ethical framework to harness the potential of AI while keeping humans in control. En L. Floridi (Ed.), ***Ethics, governance, and policies in artificial intelligence*** (pp. 91–106). Springer International Publishing.

Tapia Hermida, A. J. (2021). La responsabilidad civil derivada del uso de la inteligencia artificial y su aseguramiento. ***Revista Ibero-Latinoamericana de Seguros, 30(54)***, 111–146. https://doi.org/10.11144/Javeriana.ris54.rcdu

Tauk, C. S., & Salomão, L. F. (2023). Inteligência artificial no judiciário brasileiro: Estudo empírico sobre algoritmos e discriminação. ***Diké, 22***(23), 2–32. https://periodicos.uesc.br/index.php/dike/article/download/3819/2419/

Tejada Núñez, Á. D., & Balbin Ramos, W. R. (2025). Transformación digital imperativa: reconfigurando la gestión pública Latinoamericana para la ciudadanía del siglo XXI. ***Revista de Historia, Ciencias Humana y Pensamiento Crítico, (10)***, 984-1005. https://doi.org/10.5281/zenodo.15337401

The Path to Singularity. (2023). ***The path to singularity: The destiny of artificial intelligence?*** [Video]. YouTube. https://www.youtube.com/watch?v=W8ZLx7ulpc0

Toledo García, J. A. (Coord.). (2013). ***Pensamiento político: De la Antigüedad hasta la Modernidad*** (Tomo I). Félix Varela.

Top de Impacto. (2023). ***Los riesgos de la inteligencia artificial que no todos te cuentan*** [Video]. YouTube. http://www.youtube.com/topdeimpacto/Los_RIESGOS_de_la_Inteligencia_Artificial_que_no_todos_te_cuentan

Torreblanca Gómez de las Heras, S. (2024). El impacto de las tecnologías disruptivas en la administración pública: Nuevos retos en los modelos de gestión. European Public & Social Innovation Review, 9, 1–21. https://doi.org/10.31637/epsir-2024-850

Turing, A. M. (1950). Computing machinery and intelligence. ***Mind, 59***(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

Unión Europea. (2017). *Normas de Derecho civil sobre robótica: Resolución del Parlamento Europeo, de 16 de febrero de 2017, con recomendaciones destinadas a la Comisión (2015/2103(INL)).* https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52017IP0051

Unión Europea. (2022). *Propuesta de Directiva del Parlamento Europeo y del Consejo sobre responsabilidad en materia de inteligencia artificial* (COM(2022)496final). https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52022PC0496

Unión Europea. (2024). *Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, por el que se establecen normas armonizadas en materia de inteligencia artificial y por el que se modifican varios reglamentos y directivas.* https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX%3A32024R1689

Uribe Loaiza, M. I. (2024). *Explorando las tipologías, motivaciones, beneficios y desafíos de la inteligencia artificial en la innovación de procesos empresariales* [Trabajo Fin de Máster, Universidad de Cantabria].

Uribe, R. (2008). Ética y derecho en la posmodernidad. *Scientia, (145)*, 221–241. https://www.redalyc.org/articulo.oa?id=651769533010

Valcárcel Fernández, P., & Hernández González, F. L. (Coords.). (2025). *El Derecho Administrativo en la era de la inteligencia artificial: Actas del XVIII Congreso de la Asociación Española de Profesores de Derecho Administrativo.* Universidad de Vigo, España.

Valero Torrijos, J. (2019). Las garantías jurídicas de la inteligencia artificial en la actividad administrativa desde la perspectiva de la buena administración. *Revista Catalana de Dret Públic, 58*, 82–96. https://doi.org/10.2436/rcdp.i58.2019.3307

Vallespín Pérez, D. (2025). Responsabilidad civil extracontractual en materia de IA: Especial referencia a la carga de la prueba y la aplicación de presunciones. *Revista de Internet, Derecho y Política, (42).* https://doi.org/10.7238/idp.v0i42.432054

Velásquez Burgos, B. M., Remolina de Cleves, N., & Calle Márquez, M. G. (2009). El cerebro que aprende. *Tabula Rasa, (11)*, 329–347. https://www.redalyc.org/pdf/396/39617332014.pdf

Véliz, C. (2022). *Privacy is power: Why and how you should take back control of your data.* Penguin Random House.

Veracruz, A. (2024). Aplicación de la inteligencia artificial en el derecho. Anáhuac. https://veracruz.anahuac.mx/licenciaturas/blog/aplicaci%C3%B3n-de-la-inteligencia-artificial-en-el-derecho

Villalba, J. F. (2020). Algor-ética: La ética en la inteligencia artificial. *Anales de la Facultad de Ciencias Jurídicas y Sociales, 17*(50), 679–698. https://doi.org/10.22529/rbia.2020.1(1)02 Vogl, T. M., Seidelin, C., Ganesh, B., & Bright, J. (2019). Algorithmic bureaucracy: Managing competence, complexity, and problem solving in the age of artificial intelligence. *SSRN Electronic Journal.* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3327804

Waltz, D. L. (1997). Artificial intelligence: Realizing the ultimate promises of computing. *AI Magazine, 18*(3), 49–52. https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1305/1206

Webster, G., Creemers, R., Triolo, P., & Kania, E. (2017). Full translation: China's "New Generation Artificial Intelligence Development Plan" (2017). *DigiChina, Stanford University.* https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/

Wigg Sotomayor, I. (2024). Panorama del fenómeno de la inteligencia artificial en relación con la responsabilidad civil. *Actualidad Jurídica, 50,* 245–263. https://derecho.udd.cl/actualidad-juridica/files/2024/09/9-isabel-wigg-sotomayor.pdf

Wigwe & Partners. (2025). *Copyright infringement and generative artificial intelligence: Emerging legal challenges.* https://wigweandpartners.com/copyright-infringement-and-generative-artificial-intelligence-emerging-legal-challenges/

Zambrano, O., Talavera-Ruz, M., & Nino Peñalosa, J. A. (2025). Desafíos regulatorios de la inteligencia artificial en la comunidad andina y oportunidades para el emprendimiento digital. *Revista Ciencia UNEMI, 18*(49), 122–136. https://doi.org/10.29076/issn.2528-7737vol18iss49.2025pp122-136p

Rolando Medina Peña.

Investigador en el proyecto de investigación: Fundamentos epistemológicos del neoconstitucionalismo latinoamericano. Aciertos y desaciertos en su regulación jurídica y aplicación práctica en Ecuador, de la Universidad Metropolitana Sede Machala. Doctor en Ciencias Jurídicas por la Pontificia Universidad Católica Argentina "Santa María de los Buenos Aires". Postdoctorado en Metodología de la investigación científica, socioformación y desarrollo humano el Centro Universitario CIFE (México). Coordinador de la Comisión de investigaciones UMET sede Machala. Con experiencia en programas de postgrados en Cuba, Ecuador, Argentina y México, en tutorías de tesis de maestría y doctorales en la actualidad (cinco defendidas y cuatro en proceso). Autor de varios artículos científicos publicados en revistas científicas indexadas en bases de datos regionales y de alto impacto internacionales de: Cuba, México, España, Venezuela y Ecuador. Director de las revistas científicas: Metropolitana de Ciencias Aplicadas (REMCA), Transdisciplinaria de Estudios Sociales y Tecnológicos (RTEST) y Episteme & Praxis. Autor de varios textos científicos y capítulos de libros publicados en prestigiosas editoriales de Cuba, Bélgica y España. Ha dictado

varias conferencias en diversos eventos tantos nacionales como internacionales (México, Argentina, Cuba, Ecuador). Ha sido objeto de varias condecoraciones, entre las que destaca: "Académico correspondiente" de la Academia internacional de Ciencias, tecnologías, Educación y Humanidades, de valencia España. Correo: rolandormp74@gmail.com; rmedina@umet.edu.ec . ORCID: https://orcid.org/0000-0001-7530-5552

Yisel Muñoz Alfonso

Master en derecho privado por la Universidad de Valencia, España y la Universidad de La Habana. Doctora en Ciencias Jurídicas por la Universidad de La Habana, Profesora Titular del Departamento Civil y Asesoría de la Universidad Central " Marta Abreu" de Las Villas (UCLV). Laboró como Docente Titular y Directora de Carrera de la Universidad Metropolitana del Ecuador. Imparte docencia de pregrado y posgrado en la UCLV y en otras universidades extranjeras, es directora de un proyectó internacional VLIR e investigadora de varios proyectos de investigación nacional, sectorial y territoriales. Ha participado en eventos internacionales y nacionales y publicaciones de artículos de revistas, monografías y libros.

Yulier Campos Pérez

Licenciado en Derecho, Universidad Central “Marta Abreu” de Las Villas (UCLV), 2010. Doctor en Ciencias Jurídicas, Universidad de La Habana, 2022. Profesor Titular de Derecho de Autor, Propiedad Industrial y Derecho Cooperativo del Departamento de Derecho (UCLV). Decano de la Facultad de Ciencias Sociales de la propia universidad. Jefe del proyecto: Bases jurídicas para el uso ético y legal de las herramientas informáticas y de inteligencia artificial en las instituciones de la Educación Superior Cubana. Miembro de la Asociación Internacional Derecho e Inteligencia Artificial. Autor de varios artículos científicos en Cuba y el extranjero. Ponente en eventos nacionales e internacionales.

Daniela Lázara Hernández Domínguez

Licenciada en Derecho, Universidad Central "Marta Abreu" de Las Villas, Cuba (2023). Profesora Asistente de Derecho de Obligaciones y Contratos de la Facultad de Ciencias Sociales de la Universidad Central Marta Abreu de Las Villas. Investigadora de Proyecto Bases jurídicas para el uso ético y legal de las herramientas informáticas y de inteligencia artificial en las instituciones de la Educación Superior Cubana.

UMET UNIVERSIDAD METROPOLITANA

El libro *Inteligencia artificial, ética y derechos. Desafíos* analiza las profundas implicaciones éticas, jurídicas y sociales de las tecnologías avanzadas que automatizan procesos cognitivos y de decisión. Explora cómo estas herramientas transforman la interacción humana, los sistemas productivos y los marcos legales tradicionales, generando desafíos inéditos para la protección de derechos y la regulación de su uso. Se examinan los principios fundamentales que deben guiar su desarrollo, incluyendo la seguridad, la transparencia, la supervisión humana y la ética, enfatizando la necesidad de que estas tecnologías respeten los valores universales y los derechos de las personas. El texto reflexiona sobre los límites de su aplicación en contextos sensibles, como la medicina, la administración pública y los sistemas de reconocimiento, señalando los riesgos de sesgos, vulneraciones de privacidad y daños potenciales. Además, aborda la relación entre estas tecnologías y la propiedad intelectual, la responsabilidad frente a errores o perjuicios, y los desafíos que plantean en materia de derechos digitales y justicia. También se analiza cómo pueden contribuir al desarrollo sostenible y al bienestar social, siempre que se implementen de manera responsable y con criterios claros de gobernanza. La obra integra perspectivas interdisciplinarias, combinando análisis técnico, ético y legal, y ofrece herramientas conceptuales y prácticas para que profesionales, legisladores y académicos comprendan, evalúen y regulen estas tecnologías de manera responsable. En conjunto, constituye un aporte significativo para pensar cómo la sociedad contemporánea puede incorporar estas innovaciones sin comprometer los principios de equidad, justicia y respeto a los derechos humanos, garantizando que su desarrollo y aplicación se orienten hacia un impacto positivo y sostenible.

EDITORIAL EXCED

ISBN: 978-9942-560-11-7

9 789942 560117