



Adaptación de la Inteligencia Artificial a los Derechos Humanos

Adapting Artificial Intelligence to Human Rights

Ana Yurany Carrillo Pardo

Universidad Nebrija. Madrid (España)

acarrillop1@alumnos.nebrija.es

ORCID: 0009-0006-3804-7834

Resumen

La inteligencia artificial (IA) es una realidad que ha venido incursionando en la vida cotidiana más rápido de las percepciones diarias, sin embargo, sólo hasta la aparición del Chat Gpt nos hemos percatado de esta realidad. Las investigaciones que datan desde 1950, han avanzado al punto de que estamos en mora de regular cómo la IA debe respetar los derechos básicos del ser humano como la dignidad, la privacidad y no la discriminación. Conforme a algunas de las fuentes consultadas para este texto, pude evidenciar que se hace necesaria una sensibilización de quienes están avanzando en la revolución tecnológica, mención que se hace reconociendo que a este sector no le corresponde el desarrollo de derechos, pero sí conocer su alcance. Este trabajo tiene como propósito hacer un análisis de los dilemas del uso de la Inteligencia Artificial -IA- y algunos derechos personales, que desde su promulgación como todos los derechos de este tipo, han tenido alcance de acuerdo con el momento social de afectación universal en el que nos encontremos; por eso, con la revolución tecnológica y con la excusa de la misma, el núcleo esencial de los derechos humanos - DDHH- ha sido afectado y por tanto es obligatorio avanzar en la forma como estos se deben tutelar por parte de los estados.

Palabras clave: Inteligencia artificial; privacidad; no discriminación; Dignidad Humana.

Abstract

Artificial intelligence (AI) is a reality that has been permeating daily life faster than we realize; however, it wasn't until the emergence of GPT chat that we truly grasped this reality. Research dating back to 1950 has progressed to the point where we are overdue in regulating how AI should respect fundamental human rights such as dignity, privacy, and non-discrimination. According to some of the sources consulted for this text, it is evident that raising awareness among those driving the technological revolution is necessary. This is made while acknowledging that this sector is not responsible for developing rights, but rather for understanding their scope. The purpose of this paper is to analyse the dilemmas surrounding the use of Artificial Intelligence (AI) and certain personal rights, which, since their enactment—like all rights of this kind—have had a scope that corresponds to the current social context of universal impact in which we find ourselves; Consequently, with the technological revolution and under the pretext of it, the very core of human rights has been affected, and it is therefore imperative to advance the way in which these rights are safeguarded by states.

Keywords: Artificial intelligence; privacy; non-discrimination; human dignity.

Cómo citar este trabajo: Carrillo Pardo, Ana Yurany. (2026). Adaptación de la Inteligencia Artificial a los Derechos Humanos. *Cuadernos de RES PUBLICA en derecho y criminología*, (8), 01–10. <https://doi.org/10.46661/respublica.13516>

Recepción: 27.05.2026

Aceptación: 12.06.2026

Publicación: 23.06.2026



Este trabajo se publica bajo una licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional.

1. Introducción

Este trabajo tiene como propósito hacer un análisis de los dilemas del uso de la Inteligencia Artificial -IA- y algunos derechos personales, que desde su promulgación como todos los derechos de este tipo, han tenido alcance de acuerdo con el momento social de afectación universal en el que nos encontremos; por eso, con la revolución tecnológica y con la excusa de la misma, el núcleo esencial de los derechos humanos -DDHH- ha sido afectado y por tanto es obligatorio avanzar en la forma como estos se deben tutelar por parte de los estados.

El dilema que vamos a esbozar en este escrito es el de cómo la privacidad, se ve vulnerada con el tratamiento de datos con la necesidad de alimentar los sistemas de IA y como desde el derecho internacional se requiere poner límites; aunado a la transgresión de este derecho se encuentran otros afectados – sin ser una descripción taxativa ni excluyente - como la dignidad, la no discriminación, que se ven vulnerados cuando de forma selectiva se alimentan y se procesan datos en el marco de la creación de algoritmos por la IA (ACNUCH, 2025).

Para ello, de forma sucinta se hará mención al núcleo esencial de los derechos a la privacidad, la dignidad y la no discriminación para aterrizar en los dilemas de estos frente a la IA; cabe señalar que es la IA la que se debe adaptar al núcleo esencial de los DDHH, sin embargo, como se entenderá a lo largo del escrito, los DDHH requieren ser protegidos frente a la IA en general y claro está, y de forma específica de acuerdo al tema que trata cada herramienta que finalmente es creada por seres humanos.

2. Derechos menoscabados

2.1. La Privacidad

Empezaremos por este derecho porque en la era digital es el más perjudicado; la privacidad, está consagrada en el artículo 12 de la Declaración Universal de los Derechos Humanos -DUDDHH- Cabe señalar que este

derecho está reconocido en el PICP, la Declaración Americana, la Convención Americana y otros instrumentos internacionales; en cada uno de ellos su alcance varía, sin embargo, para este escrito sólo se hará alusión a la DUDDHH. y es el derecho que tienen todas las personas a mantener en custodia, lejos de los demás, las características de su vida, sin que nadie puede menoscabar esta posibilidad con una intromisión arbitraria a sus relaciones personales y familiares, la forma como lleva su vida, sus comunicaciones y mucho menos perjudicar el buen nombre con el que cuente. Es por ello que el derecho a la privacidad cuenta con subcategorías de protección: la privacidad informativa, la privacidad de las decisiones, la privacidad corporal y la privacidad comunicativa o asociativa (Nations, n.d.).

Este derecho cuenta con límites legales los cuales deben ser debidamente justificados; cualquier injerencia debe ser proporcional y necesaria al fin que se busca, y estar expresa de forma taxativa en la ley, la cual debe estar en concordancia con las demás normas internacionales, en especial la presunción de inocencia, libertad de expresión, libre desarrollo de la personalidad, tratamiento de datos, buen nombre e intimidad familiar, entre otros (ACNUDH-USAID, 2007).

Los organismos internacionales encargados de amparar los derechos humanos han evidenciado desde la década del 2000, la necesidad de compatibilizar el derecho a privacidad con el avance en el tratamiento de datos, sin embargo, esta necesidad se había quedado corta para entonces, más aún cuando para el 2007 ya se contaba en la telefonía celular con la posibilidad de conocer la ubicación del usuario del dispositivo con la inclusión en estos de los sistemas GPS. Si bien, dicha ubicación se puede restringir por el usuario, la sola posibilidad de que un teléfono celular pueda transmitir la ubicación, ya es una ventana abierta para infringir el derecho a la privacidad.

Como antecedente es preciso mencionar que para el año 2000 la vulneración del derecho a la privacidad se veía ya evidenciado por los avances en internet y por ende, la Relatoría para la Libertad de Expresión de la ONU señaló al hacer el análisis del principio de habeas data, o tratamiento de datos en la era digital, que el ejercicio del mismo cuenta con tres ámbitos:

- (i) el primero refiere a que las personas no sean perturbadas en su privacidad, es decir el derecho a que los asuntos propios sean administrados por su propietario y sea él quien decida qué es o no de dominio público;
- (ii) que el afectado conozca la información sobre sí mismo que este contenida en los registros públicos y privados, y por ende a modificarlos, anularlos o rectificarlos en caso de ser necesario para salvaguardar su derecho; y finalmente
- (iii) utilizar su derecho para revisar la información en bases de datos públicas y privadas.

Atendiendo al contexto del momento, la relatoría fue clara en mencionar que el control de los datos *“...adquiere una importancia aún mayor con el avance de nuevas tecnologías. Con la expansión en el uso de la computación e internet, tanto el Estado como el sector privado tienen a su disposición en forma rápida una gran cantidad de información sobre las personas”*

Sin embargo y ante lo expuesto, la protección de la privacidad, y siendo sólo una de las referencias al respecto, ¿cómo se ve amparado cuando son las mismas personas las que divulgan sus datos en la web por voluntad propia o por obligación cuando un trámite, como, por ejemplo, la expedición de un pasaporte, así lo exige? La vulneración a la privacidad por la IA no solo se vio evidenciada por los académicos sino también por los propios desarrolladores de la naciente IA en la década del 2000 (HAO, EL IMPERIO DE LA IA, 2025), en especial después de evidenciar que los dispositivos celulares podrían y serían equipados con GPS.

2.2. Derecho a la dignidad

Este principio, en cabeza de la persona humana es el segundo en importancia después de la vida, se encuentra enunciado en el preámbulo de la DUDDHH como intrínseco y por lo tanto se enuncia en su artículo primero como atributo personal. Por ser un principio este no se encuentra conceptualizado en la DUDDHH, sin embargo, su significado es determinante. Como principio, este derecho hace alusión al trato como humanos con que cuentan todas las personas y en virtud de ello, no les es permitido a otro cosificar, humillar o discriminar negativamente a otra persona.

Este principio por ser intrínseco permea a todo el conjunto de derechos, se accede en condiciones de igualdad y no se limita sólo a la libertad de desarrollar la vida conforme al querer subjetivo (Sánchez, 2020, pp. 228-231). Este derecho cuenta con antecedentes desde el derecho romano y se ha desarrollado con los siglos ante creencias de superioridad de algunos colectivos frente a otros por diversas razones (Sánchez, 2020)

Ahora bien, en un intento por conceptualizarlo el jurista Bernardo Carvajal Sánchez señala: *“Entonces, si entendemos que la dignidad humana es el “valor eminentemente atado a toda persona por el hecho de pertenecer al género humano”, (...). En efecto, lo anterior quiere decir en el lenguaje de los juristas que, en razón de la dignidad humana (causa), todos los destinatarios de la Constitución -sean poderes públicos o privados- tienen la obligación de respetar a toda persona en su humanidad, pues es lo que ella merece, al igual que cualquier otro miembro de la especie humana (consecuencia). Es entonces respetando a la persona humana como se respeta el principio de dignidad humana, (...) respetar a la persona humana en razón de su dignidad derivada del solo hecho de pertenecer a la comunidad humana.”* (negritas fuera de texto) (Sánchez, 2020, p. 147).

En consecuencia, de lo anterior se debe entender que según la RAE, merecer es tener

mérito o corresponderle algo y respeto es sinónimo de consideración.

¿Pretender que la IA pueda reemplazar las capacidades del ser humano, no va en contra de la dignidad de las personas?

¿No estaremos ante una forma de vulneración de la dignidad humana cuando se extraen datos en zonas marginadas sin mediar conocimiento previo del cómo y para qué de dicha información?

Sobre esto se hará mención más adelante.

2.3. La No Discriminación

Este principio se traduce en el derecho que tienen todos los seres humanos a ser tratados sin distinción alguna en el ejercicio de sus derechos. Este principio es complementario al derecho de igualdad y se encuentra consagrado en los artículos primero y segundo de la DUDDHH y cuenta con instrumentos de derecho internacional público de materialización específica, como en la Convención Internacional sobre la eliminación de todas las formas de discriminación racial (1965) y la Convención sobre la eliminación de todas las formas de discriminación contra la mujer (1979).

Si bien parece que el derecho a la no discriminación y la igualdad son similares, la existencia de instrumentos internacionales específicos y la misma doctrina aclaran que son conceptos diferentes, sin embargo, es igualmente claro que cuando un hecho no permite la materialización de un derecho en pro de la raza, la religión, orientación política, u otro factor no justificado estamos ante una distinción, es decir ante una discriminación; (ACNUDH-USAID, 2007, págs. 911-976) impedir o menoscabar la materialización de un derecho por cualquier razón sin mediar una justificación razonable es discriminación (ACNUCH, 2025).

Ahora, a qué se refiere un factor justificado, se refiere a las acciones de discriminación positiva o inversas, siendo las medidas que se tomen en favor de un grupo marginado, desprotegido o de especial atención que por

sus condiciones se encuentran en mayor grado de vulnerabilidad y por tanto no pueden acceder a ciertos derechos y es por ello que se requieren medidas preferenciales, con el fin de eliminar barreras de acceso a sus derechos que impiden su acceso igualitario (Corte Constitucional Colombiana, 2017)

Al igual que otros derechos, la no discriminación cuenta con dos ámbitos para su materialización, el primero un ámbito de respeto de acción, es decir del libre ejercicio de este, traducido en no mediar diferenciación en la realización de un derecho a menos que la diferenciación sea inversa; y el segundo el ámbito de la abstención de intromisión por parte del estado o de privados, esto últimos que deben ser controlados por el primero.

Ahora bien, sobre este punto cabe mencionar que la no discriminación cuenta con otro factor a tener en cuenta, y es que se encuentran derechos que son precisos de acuerdo a la situación en que se encuentre el sujeto; por ejemplo, un extranjero no se discrimina porque no ejerza el derecho al sufragio en el país en donde se encuentre de visita, pues su condición es diferente.

3. Qué es la inteligencia artificial?

Conceptualizar de forma profunda la IA es tan amplio como los datos que esta puede recoger, sin embargo, se hará una aproximación para el lector. La IA es un conjunto de sistemas tecnológicos que buscan mediante la alimentación y procesamiento de datos dar resultados y los ámbitos de aplicación son tantos “*como ámbitos tiene el ser humano*”: reconocimiento facial, de imágenes, búsqueda de bibliografía, toma de notas, agendamiento, sistemas lingüísticos, entre otras (Nations, n.d.).

La IA busca la imitación al procesamiento de datos para llevar a un resultado que permita arrojar información precisa; algunas vertientes van más allá y busca que se replique el pensamiento y los procesos cognitivos de los seres humanos, sin embargo, dista mucho de ello a pesar de los esfuerzos.

Para obtener cualquier resultado es necesario que las herramientas tecnológicas se alimenten y utilicen información que le es sustraídas de interacciones de las personas con el fin de poder imitar los procesos cognitivos de los seres humanos (UNESCO, 2021) Además de ello, requiere un avance en la capacidad operativa y de almacenamiento de los dispositivos de cómputo. El procesamiento de datos es denominado algoritmo, cuyo concepto básico es el conjunto de pasos para operaciones lógicas matemáticas de puntos de información que llegan a un resultado finito (KANUTH, 2021)

Como se mencionó, para que la IA funcione requiere de información, datos, que permiten el procesamiento de los algoritmos para llegar a un resultado (UNESCO, 2021) los cuales son entregados conforme a la programación que se haga sobre la herramienta de IA. El desarrollo de la IA data de 1950 cuando Alan Turing publicó el artículo científico titulado *"Computing Machinery and Intelligence"*, en donde, paradójicamente, puso a pensar a la comunidad académica sobre la posibilidad de que las máquinas pudieran hacer procesos lógicos similares a los de la inteligencia humana (TURING, 1950).

A pesar de la muerte del autor tres años después de la publicación, en 1956 se llevó a cabo una reunión en Dartmouth College, de hombres cuyo color de piel es blanca, hecho que ante el contexto del entonces podría crear algunos sesgos discriminatorios, en donde se obtuvo como resultado la creación de una nueva disciplina que buscaba el desarrollo de la idea de avanzar en *"máquinas que piensan como los humanos"*, concepto que inicialmente no era muy atractivo para el público objetivo -financiadores e investigadores-, pero que con el fin de llegar a patrocinadores se bautizó como *inteligencia artificial*.

Se considera por parte del ecosistema de la IA que este nombre fue un error porque a pesar de la atracción derivada de la sofisticación del enunciado, esto permitió *"humanizar"* las máquinas y facilitó la financiación que

requerían algunos investigadores necesitados de resaltar con resultados que respondieran al mercado. (HAO, EL IMPERIO DE LA IA, 2025, pp. 138-139).

Esta semejanza, ha venido avanzando y provocando múltiples daños que a continuación se enuncian. Es bastante ambicioso y pasa la línea de la pertinencia hacer investigaciones para que una máquina piense como un ser humano, e insistir en ello ha causado daños ampliamente documentados.

Los científicos de IA establecieron dos líneas de profundización, la primera, la simbólica la cual se basa en el saber y procesamiento de datos; y la segunda, la conexionista que se basa en el aprendizaje autónomo a partir de los datos que se ubican principalmente en las redes y que busca dotar a las herramientas de IA con pensamiento propio (Marcus, 2019).

Estos dos conceptos fueron recogidos en la línea neurosimbólica, buscando beneficios altruistas. A pesar de las argumentaciones, la línea del conexionismo dominó las empresas de Silicon Valley, en donde se busca recrear las redes neuronales humanas, sin embargo, este ejercicio, si bien ha creado resultados onerosos para el mercado ha traído desde acciones de discriminación de grupos poblacionales hasta la pérdida de vidas humanas (HAO, EL IMPERIO DE LA IA, 2025, pp. 136-175).

El avance más prominente y popular de la IA fue la creación de Chat GPT en noviembre de 2022. A la fecha se conocen otros desarrollos de IA de uso diario, algunos altamente útiles, otros menos, y todo caso mayoritariamente comercial, dejando atrás algunas herramientas para mejorar situaciones que aquejan a la humanidad.

4. Daños de la inteligencia Artificial

En análisis adversos, sin embargo no poco realistas, los desarrolladores de esta tecnología han intentado y siguen intentado *"un reemplazo"* de la inteligencia humana; no cabe negar que la IA trae beneficios para la humanidad, en especial como soporte de las

actividades diarias y las investigaciones de todo tipo como para por ejemplo analizar como disminuir brechas sociales, avanzar en cura de enfermedades, mejorar la calidad de vida de las personas, buscar alternativas para disminuir los gases de efecto invernadero, entre otros, sin embargo, la alta percepción como medio lucrativo ha permitido desconocer este tipo de investigaciones y ha traído al contrario acciones generalizadas que van en contra de los derechos de privacidad, dignidad humana y la no discriminación, entre otros derechos.

Esta realidad ha sido evidenciada tanto por Karen Hao, Paula Ramírez y por diferentes instancias de la ONU citadas en este escrito, y por científicos que han incomodado a la industria y en ocasiones han sido relegados de sus posiciones (HAO, "Fue terrible y despiadado", los últimos días de Timnit Gebru en Google, n.d.). El afán por la alimentación de datos, el sólo hecho de buscar que una maquina reemplace el 100% de las actividades pensantes del ser humano y el sesgo de los datos aportados permiten vulnerar los derechos mencionados, enunciación que dista mucho de ser exhaustiva. Es claro que los sesgos nacen por la información que es entregada por las personas biológicas y pretender que un ente no biológico tome "*conciencia*" mediante la racionalización de datos es bastante contraproducente como se ha evidenciado.

La línea de desarrollo que ha imperado es la conexionista que busca simular las redes neuronales para hacer análisis, lo que a su vez le permite "*ganar experiencia*" para construir "*su propio criterio*". Con este fin, se requiere gran volumen de datos, variedad y valor en grado de aporte.

Alimentar estos datos y permitir que una herramienta de IA tome decisiones con base en su análisis es arriesgado en varias dimensiones; una de ellas es la imprevisibilidad en el momento de hacer el desarrollo de IA, lo que podría en riesgo a las personas. Al respecto, en el artículo *Concrete Problems in IA Safety* escrito por

investigadores de IA señalaron: "*the problem of accidents in machine learning systems, defined as unintended and harmful behavior that may emerge from poor design of real-world AI systems.*" "El problema de los accidentes en las máquinas de sistemas de aprendizaje, definido como un comportamiento involuntario y prejudicial que podría emerger desde un pobre diseño del sistema real de la IA" (Dario Amodei, 2016); otro ámbito tiene que ver con el tratamiento de datos, en donde la ambición provocan figuras como fraude para adquirir datos -phishing- y creación de sitios web para capturar datos -pharming- información que saldría del dominio de las personas, constituyéndose así una de las forma de vulnerar el derecho a la privacidad y la dignidad de las personas, esto sin contar el uso indebido de los datos obtenidos de forma ilícita (Barbosa, 2023, pp. 108-113).

Actividades llevadas a cabo en China y Colombia dan cuenta de acciones de obtención de datos de forma discriminatoria y en contra de la dignidad humana; en dichos países se llevaron a cabo pruebas de una startup sobre la actividad cerebral de niños a cambio de información, esto termino siendo, en primer lugar, un reflejo del uso de datos personales indebidamente gestionados, en donde no sólo se vulnera la privacidad (especialmente en una población vulnerable y de especial protección como los niños y niñas) sino también parece el retorno de un patrón relacionado con los tiempos más crudos de superioridad de un pueblo sobre otro (s). (HAO, EL IMPERIO DE LA IA, 2025)

El mismo Comisionado de las Naciones Unidas para los derechos humanos identificó que estas situaciones, aplicación de nuevas tecnologías no probadas en poblaciones sin capacidad de oponerse, representa una vulneración al derecho a la privacidad con base en acciones discriminatorias. (Alto Comisionado de las ONU para los DDHH, 2025)

Pero la obtención de datos no sólo afecta el tratamiento de estos, sino también derechos laborales de quienes se encargan de subirlos a

la web; en Kenia se documentó actividades que constituyen explotación laboral en procesos de análisis de datos con el fin de alimentar herramientas de IA (Perrigo, 2022)

La enumeración de casos sobre defectos en el diseño puede ser larga: un Uber sin conductor provocó la muerte de una transeúnte en Tempe Arizona, al no identificarla como una persona; en 2024 se reportó en 200 accidentes la participación de vehículos Autopilot de Tesla; ¿falla en el sistema? La representación demográfica en los entrenamientos de IA también advierte otro punto de discriminación, pues pueden descartar ciertas características y desconocerlas cuando no se encuentran representadas en tales grupos. Adicionalmente, en el manejo de bots un periodista del *The New York Times* intercambió mensajes con chat Bing de Microsoft durante más de dos horas y entre las respuestas del sistema estuvo “estoy enamorada de ti” (HAO, EL IMPERIO DE LA IA, 2025, pp. 136-175).

Otro de los daños que aparea la IA es el uso de herramientas para la selección de personal, en el contexto empresarial se ven expuestas tanto las empresas como los potenciales candidatos al ser los criterios de selección formados por datos extraídos de las redes sociales permitiendo exclusiones (Barbosa, 2023), lo que conduce a utilizar información limitada por opiniones que la máquina toma para llegar a una decisión. Al igual que la información tomada de las redes, las que son suministradas por personas en la nueva industria de anotación de datos, cuentan con criterios de valoración subjetiva que carecen de neutralidad.

En casos más graves en procesos de selección según los datos suministrados por la web, los sistemas de reconocimiento biométrico pueden hacer discriminaciones en el momento de hacer el análisis de ciertos grupos poblaciones conforme a su color de piel, sexo o características faciales, si bien hacer reconocimiento facial es legítimo, los sesgos que permite el desarrollo tecnológico

puede conllevar graves riesgos al ejercicio de los derechos. Categorizar ciertas características a un perfil criminal constituye señalamientos discriminatorios con consecuencias considerablemente graves (AMERICAN CIVIL LIBERTIES UNION, 2024) La falta de representatividad de ciertos grupos poblaciones, que, por deficiencia en el acceso a servicios digitales derivados de brechas sociales, conllevaría que la IA los desconozca y por ende los excluya de ciertas actividades bajo un desarrollo tecnológico de este tipo.

Igualmente el medio ambiental es perjudicado por la IA, la cantidad de emisión de carbono que se produce para realizar el análisis de datos es bastante alta y se encuentra demostrado por una investigación de la universidad de Massachusetts, en donde sólo entrenar un modelo *Bidirectional Encoder Representations from Transformers* -BERT- para que una red neuronal conozca el contexto y lenguaje humano, permite la emisión de carbono en el grado equivalente a un vuelo trans americano (Anna Korhonen, 2019). Cabe anotar que los BERT son modelos mediante los cuales se lee el sentido de una oración, comprendiendo, no solo la palabra sino tratando de comprender el sentido de un enunciado, para alimentar con esta información los modelos de IA.

Estos daños también se transfieren al campo de la investigación, científicas del medio de la IA como Timmit Gebru, Joy Boulamvini y Emily Bender, entre otras, han escrito sobre la poca inclusión de mujeres y personas afro en los procesos de investigación, e incluso de cómo se excluyen del mercado laboral a científicos que cuestionan las formas de las grandes empresas de IA. (Boulamvini, 2020)

El procesamiento de datos podría de igual forma inducir la selección de una vertiente de opinión conforme al análisis que haga la IA. El análisis de los daños macro y casos específicos requeriría otro escrito (Barbosa, 2023, p. 169)

5. Conclusiones

Es claro que al ser humano con habilidad de adaptación natural le ha costado trabajo reconocer la dignidad del otro;

¿cómo podría hacer esto una maquina con posibilidad de tener daños en el sistema desde su diseño?

¿Cómo se piensa que una maquina podría tener conciencia similar al ser humano?

Es por ello que a lo largo de este escrito se han dejado varias expresiones en comillas pues pensar que una maquina puede remplazar al ser humano es desproporcionado y proviene de las conclusiones de algunos sectores del campo de la investigación, por la necesidad de dominar el mercado promovido por sus financiadores.

Esta intención es la principal razón para que el derecho internacional regule la IA. (Barbosa, 2023, pp. 12-16). Dicha regulación debe abarcar todas las partes de la cadena de la industria, no sólo se reduce a la formación de herramienta adecuadas y útiles más allá de las ganancias financieras, requiere la transversalización de derechos como los laborales y otras sub industrias como la de recopilación de datos. Además de ello, requiere la promoción de investigaciones que permitan avances en temas relacionados en el ámbito de la salud y ambientales, por mencionar algunos, que se han visto relegados ya que son pocos los científicos talentosos del sector que prefieren fines más altruistas. (HAO, EL IMPERIO DE LA IA, 2025, pp. 123-124)

Lejos de que algunos señalen que los sistemas internacionales están en declive, este debate llama la atención sobre la necesidad de fortalecerlos para poder regular el alcance de los derechos conforme al avance de la sociedad; el derecho es eso, los derechos humanos no desaparecen, se les da alcance conforme al momento social que se vive, el ser humano está llamado adaptarse y las IA llamadas a asumir esa adaptación, con análisis realistas y menos apocalípticos en momentos en que los desarrollos de la IA van a mayor velocidad. Estamos sólo en el inicio de lo que

debería ser la regulación universal, que como se ha visto a lo largo de la historia de la humanidad, quienes se encuentran construyendo desarrollos tecnológicos consideran que cuentan con ventajas que les permite considerar superioridad frente a sectores menos favorecidos.

Si bien esto parece más dirigido al mundo empresarial y al respecto se encuentran como doctrina jurídica los Principios Rectores Sobre Las Empresas y Los Derechos Humanos, estos son insuficientes, pues no se encuentran adaptados a industrias como la de recopilación de datos, que trascienden la jurisdicción territorial.

Algunos sistemas regionales de protección de los DDHH, como la Unión Africana y la U.E. por vía del Consejo de Europa, han emitido herramientas que regulan la sistematización de datos, sin embargo, el sistema universal se ha quedado en recomendaciones y líneas auxiliares del derecho internacional de los derechos humanos -DIDH-, herramientas insuficientes para la muestra de los daños personales y comunitarios que puede ocasionar la falta de regulación de las actividades de IA en su fase comercial; esto significa que además de regular la actividad empresarial se requieren reglas que incentive a los países y a las empresas a promover investigaciones con fines más altruistas en beneficio de lo que la carta de la ONU el interés común.

Además de reglamentar la consolidación de datos, es necesario que las normas internacionales integren otros ámbitos de regulación de IA. Uno de ellos es la transparencia y la rendición de cuenta, pues la competitividad del mercado hace que la dinámica de los desarrollos y contenidos tecnológicos se encuentren en reserva por las empresas del sector, sin ningún tipo de análisis que permita establecer si respeta o no derechos básicos como la privacidad, la dignidad y la prohibición de no discriminación. (Alto Comisionado de las ONU para los DDHH, 2025)(HAO, EL IMPERIO DE LA IA, 2025, pp. 256-274)

Delimitar hasta qué punto la IA puede predecir el comportamiento humano, pues de allí se deriva que ciertos grupos poblacionales sean categorizados de forma discriminatoria, al no contar con representación suficiente en la muestra poblacional o al contar con sesgos de la información que se incluye para la creación de un algoritmo.

Incluir líneas claras de concientización a la población en general de las consecuencias de exponer sus datos y los derechos que tienen (adminístralos, supervisar las bases de datos y modificarlos la información)

Establecer un procedimiento para la prueba de un nuevo desarrollo que eviten pruebas en poblaciones que no tengan la capacidad de reusarse, en especial porque hacen parte de poblaciones vulnerables.

Con los límites que se deriva de la soberanía de los pueblos, se requiere regular las relaciones entre el sector privado del cambio de la IA y los Estados, para disminuir la capacidad de influencia en los primeros en pro de acciones que pueden ir en contra del DIDH.

Incentivar la investigación de la IA simbólica la cual se basa en el saber, más no en la intención de darle "conciencia" a los desarrollos tecnológicos.

Finalmente es necesario establecer sanciones para las empresas, a los individuos y a los estados por no garantizar el ejercicio de derechos en el marco de la implementación de herramientas IA.

Referencias

- ALEXANDER, Robert. (2010). A theory of constitutional rights. New York: Oxford
<https://doi.org/10.2202/1938-2545.1042>
- ACNUDH. (2025). Promoción y protección de todos los derechos humanos, civiles, políticos, económicos, sociales y culturales, incluido el derecho al desarrollo.
<https://www.ohchr.org/es/documents/thematic-reports/ahrc6045-right-privacy-digital-age-report-office-united-nations-high>: Naciones Unidas.
- ACNUDH-USAID. (2007). Derecho a la intimidad, honra, reputación y personalidad. En D. O'Connell, Derecho Internacional de los DDHH (págs. 519-549). Bogotá: Nuevas Ediciones Ltda.
- AMONEI, Dario, Chris OLAH, Jacob STEINHARDT, Paul CHRISTIANO, John SCHULMAN, & Dan MANÉ (2016) Concrete Problems in AI Safety. arXiv.
<https://doi.org/10.48550/arXiv.1606.06565>
- Alto Comisionado de las Naciones Unidas para los Derechos Humanos, (2025). El derecho a la privacidad en la era digital A/HRC/60/45 ONU
- American Civil Liberties Union, (2024)
<https://www.aclu.org/documents/aclu-comment-facial-recognition-and-biometric-technologies-eo-14074-13eACLU>
- BUOLAMWINI, Joy, "Gender Shades" (2020).
<https://www.media.mit.edu/projects/gender-shades/overview/> MIT.
- GANGI GUILLÉN, Giuseppe. Kodjack. VII. (2025). United States' strategic shift and International Humanitarian Law: Implications for the Russia-Ukraine war. *Revista Científica General José María Córdova*, 23(49), 299–315.
<https://doi.org/10.21830/19006586.1467>
- GANGI GUILLÉN, Giuseppe. Kodjack. VII. (2023). Dinámicas migratorias en la frontera colombo venezolana y su relación con la criminalidad transnacional. *Revista Científica General José María Córdova*, 21(44), 907-924.
<https://doi.org/10.21830/19006586.984>
- GANGI GUILLÉN, Giuseppe. Kodjack. VII. & Delgado Morán, Juan José. (2025). Derechos Humanos y Terrorismo. El terrorismo en Europa: la salvaguarda de la seguridad y la protección de las víctimas. Editorial Síndesis. ISBN: 979-13-87929-25-1
- GANGI GUILLÉN, Giuseppe. Kodjack. VII. (2026). Innovación en la enseñanza del Derecho Internacional Humanitario: simulaciones para asesores jurídicos operacionales. "Revolución en las aulas: cómo la innovación está reescribiendo la

- universidad". Ed. Colex. ISBN: 979-13-7011-516-6
- GANGI GUILLÉN, Giuseppe. Kodjack. VII. (2023). An exploration of socio-cultural and linguistic issues for a sustainable migration in the global north. Aranzadi
- GANGI GUILLÉN, Giuseppe. Kodjack. VII. (2023). La migración venezolana en España: protección internacional, inclusión social y errores de la administración general del Estado. An exploration of socio-cultural and linguistic issues for a sustainable migration in the global north (Aranzadi), pp. 225-241. ISBN: 978-84-9135-116-0
- GANGI GUILLÉN, Giuseppe. Kodjack. (2025). Derechos humanos y derecho penal en la era de la inteligencia artificial: retos y propuestas. *Cuadernos de RES PUBLICA en derecho y criminología*, <https://doi.org/10.46661/respublica.11635>
- HAO, Karen. (2025). El imperio de la IA. Ed: Ariel.
- HAO, Karen. (2023). "Fue terrible y despiadado", los últimos días de Timnit Gebru en Google, n.d., MIT Technology
- KANUTH, Donald. (2021) Algoritmos Fundamentales, Ed. Revertés S.A.
- KORHONEN, Anna. (2019) Energy and Policy Considerations for Deep Learning. Revista Reverté S.A.
- LIZ RIVAS, Lenny (2018). Algunas bases neurológicas sobre la violencia y la agresión, en ;"Conflictos y diplomacia, desarrollo y paz, globalización y medio ambiente " coord. Por Emilio José García Mercader, Claudio Payá Santos; César Augusto Giner Alegría (dir.), Juan Jose Delgado Morán (dir.), Thomson Reuters/Aranzadi, pp. 943-955. <https://doi.org/10.5281/zenodo.14559664>
- MARCUS, Gary. (2019 de Septiembre de 2019). We Can't Trust IA Systems Built on Deep Learning Alone". (K. Hao, Entrevistador)
- ONU (2025) , Inteligencia Artificial.
- O'DONELL. Daniel. (2007) ACNUDH-USAID Derecho Internacional de los Derechos Humanos .Ed. Nuevas Ediciones.
- PERRIGO, Billy. (2022). Inside Facebook's African Sweatshop (2022). <https://time.com/6147458/facebook-africa-content-moderation-employee-treatment/>.
- RAMÍREZ BARBOSA, Paula Andrea. (2024), La inteligencia artificial y las nuevas fronteras jurídicas. Ed. Tirand lo blanch.
- SÁNCHEZ CARVAJAL, Bernardo. (2019). La Dignidad Humana como norma de derecho fundamental. Ed. U Externado de Colombia.
- Sentencia C115 de 2017. Corte Constitucional Colombiana.
- TURING, Alan. (1950), Computing Machinery And Intelligence. Revista: The Mind Association. <https://doi.org/10.1093/mind/LIX.236.433>
- UNESCO (2021) Inteligencia Artificial y educación, guía para las personas a cargo de formular políticas.
- UNESCO, (2021). Recomendaciones sobre la Ética de la Inteligencia Artificial, Ed. UNESCO.