



## LA INTELIGENCIA ARTIFICIAL Y SU IMPACTO EN LA SOCIEDAD

### Resumen

Hoy día se critica, pero se incluyen mecanismos con inteligencia artificial (IA) en distintos ámbitos de la vida del ser humano. El alcance de los distintos mecanismos en las organizaciones presenta ventajas y desventajas para el ser humano, incluso cognitivas, ¿quién determina o controla su uso? La IA es una palanca para transformar distintos trabajos y hábitos que realiza el ser humano, pero posee implicaciones en variados dominios. Se reflexiona sobre ese auge que toma, los riesgos, los desafíos y la ética, desde luego, mirando su evolución, aplicaciones, concepciones y perspectivas, ¿qué acción jurídica se desarrolla? Estas líneas muestran algunos hitos clave para su avance. Se explora aplicaciones de la IA en entornos médicos, banca, transporte, producción, recursos humanos y gestión de conocimiento, mostrando de manera breve la eficiencia y experiencias como usuario. En la ética se mira la privacidad, el sesgo y las implicaciones sociales, en particular en la educación. El potencial de la IA lleva a construir una nueva sociedad, pero trae efectos negativos en el empleo, la educación, la gobernanza y los desafíos sociales, entre otros. Y ante el desarrollo de las tecnologías emergentes se trata de identificar, los desafíos para la privacidad, seguridad y transparencia de los datos. Al final, las reflexiones consideran la importancia de un desarrollo responsable y ético de la IA, a sabiendas que se requiere de mayor investigación para implementar procesos cognitivos del ser humano en las máquinas, pero por ahora, es una base que establece una comprensión holística sobre la IA. Tal vez el luchar por una ética para la gobernanza de la IA, debe partir de los procesos educativos, pues muchos de ellos están quedando en la virtualidad o a distancia, sin determinar cuáles son los efectos en la salud de las personas.

**Palabras clave:** inteligencia artificial, aplicación, educación, ética, sociedad.

## ARTIFICIAL INTELLIGENCE AND ITS IMPACT ON SOCIETY

### Abstract

Nowadays it is criticized, but mechanisms with artificial intelligence (AI) are included in different areas of human life. The scope of the various mechanisms in organizations has advantages and disadvantages for humans, even cognitive, who determines or controls their use? AI is a lever to transform different jobs and habits that human beings do, but it has implications in various domains. We reflect on this boom that takes, the risks, challenges and ethics, of course, looking at its evolution, applications, conceptions and perspectives, what legal action is developed? These lines show some key milestones for its progress. It explores AI applications in medical, banking, transport, production, human resources and knowledge management, briefly showing efficiency and user experiences. Ethics look at privacy, bias and social implications, particularly in education. The potential of AI leads to building a new society, but it has negative effects on employment, education, governance, and social challenges, among others. And with the development of emerging technologies, it is important to identify challenges for privacy, security, and data transparency. In the end, the reflections consider the importance of a responsible and ethical development of AI, knowing that more research is needed to implement cognitive processes of the human being in the machines, but for now, it is a basis that establishes a holistic understanding of AI. Perhaps the fight for an ethics for AI governance must start from educational processes, as many of them are remaining virtual or remote, without determining what the effects on people's health are.

**Keywords:** artificial intelligence, application, education, ethics, society.



## A INTELIGENCIA ARTIFICIAL E O SEU IMPACTO NA SOCIEDADE

### Resumo

Hoje foi crítico, mas foram incluídos mecanismos de inteligência artificial (IA) em diferentes ambientes da vida do ser humano. O alcance dos diferentes mecanismos nas organizações apresenta vantagens e desvantagens para o ser humano, inclusive cognitivo, quem determina ou controla seu uso? A IA é uma palanca para transformar diferentes trabalhos e hábitos que realizam o ser humano, mas apresenta implicações em vários domínios. Se você refletir sobre o que faz, os riscos, os desafios e a ética, desde então, olhando para sua evolução, aplicações, concepções e perspectivas, que ação jurídica se desenvolve? Estas linhas mostram algumas chaves para seu avanço. São aplicações de IA em ambientes médicos, bancários, transporte, produção, recursos humanos e gerenciamento de conhecimento, mostrando de maneira breve a eficiência e as experiências de um usuário. Na ética se mira a privacidade, a segurança e as implicações sociais, em particular na educação. O potencial da IA leva para construir uma nova sociedade, mas traz efeitos negativos no emprego, na educação, na governança e nos desafios sociais, entre outros. E antes do desenvolvimento das tecnologias emergentes trata-se de identificar, os desafios para a privacidade, segurança e transparência dos dados. No final, as reflexões consideraram a importância de um desenvolvimento responsável e ético da IA, sabendo que é necessária uma maior investigação para implementar processos cognitivos do ser humano nas máquinas, mas por agora é uma base que estabelece uma compreensão holística sobre a IA. Tal vez que a luta por uma ética para o governo da IA, deve partir dos processos educativos, muitos deles estão caindo na virtualidade ou à distância, sem determinar quais são os efeitos na saúde das pessoas.

**Palavras-chave:** inteligência artificial, aplicação, educação, ética, sociedade.

### INTRODUCCIÓN

La inteligencia artificial (IA) desde sus inicios por la década de 1950, se desarrolla a pasos de gigante, transformando varios contextos, cambiando la forma como se aprende, se enseña, se vive, se trabaja e interactúa, además construyendo una sociedad tecnolozada por variedad de dispositivos electrónicos con IA.

Además de los avances significativos en elaborar dispositivos que faciliten hallar información sobre la profundidad de los mares o sobre el espacio, la mejora del aprendizaje automático, que llega ahora al aprendizaje profundo, considerando el buen procesamiento del lenguaje natural, se construyen herramientas para realizar tareas académicas.

El desarrollo de la IA lleva a elaborar modelos, técnicas y algoritmos sofisticados que muestran sorprendentemente similitudes con los procesos cognitivos del ser humano en la realización de diversas tareas, como también la potencialidad para el reconocimiento de imágenes, la traducción de idiomas y los distintos juegos, muchos algoritmos se construyen.

Hoy día, la IA es una energía que transforma los trabajos en variedad de organizaciones y empresas de toda índole, en particular en servicios de salud y médicos, las finanzas, la producción, el control y seguridad, el transporte y otros más.

La IA surge ante la idea de Alan Turing en 1950, de que las máquinas podrían tener inteligencia, igual o superior a la del ser humano. La reunión de Dartmouth en 1956 fue el hito para que empezar a plantearse en distintas direcciones su desarrollo, pues en esa reunión se exploró las posibilidades de construir distintos sistemas con inteligencia.



Desde entonces, la IA ha pasado por varias etapas en su avance, desde el desarrollo de los primeros sistemas para expertos, basados en reglas, hasta los inicios del aprendizaje automático mediante redes neuronales hasta llegar hoy día al aprendizaje profundo mediante redes convolucionales [Sutskever et al., 2014].

Las aplicaciones de mecanismos con IA son varias. En el ámbito médico se emplea para el análisis de distintas imágenes, para identificar anomalías y enfermedades, se realiza diagnóstico y posibles tratamientos, se generan nuevos fármacos, entre otras cosas. En la banca se emplea para detección de fraudes, la calificación crediticia y la evaluación de riesgos, entre otros.

Se construyen vehículos autónomos, sondas que viajan por el espacio con amplia autonomía, además se elaboran controles inteligentes para que los sistemas de transporte sean seguros en las distintas operaciones y las diferentes travesías. En las fábricas, la IA mejora los procesos productivos pues se automatizan las líneas de producción y el mismo control de calidad. Se elaboran distintos juegos para entretenimiento, donde la IA mejora las experiencias de los usuarios, además, se desarrollan distintos sistemas inteligentes para los procesos de enseñanza/aprendizaje.

No obstante, la utilización sin control de distintos mecanismos con IA plantea varias consideraciones éticas, en particular cuando los sistemas con IA se les delega tomar decisiones y generar predicciones que pueden impactar en la vida [Gerke et al., 2020], entonces, surge preocupación por el posible sesgo, la privacidad, la transparencia, pues se toman para los diferentes procesos volúmenes de datos que se poseen.

Se mira que el potencial de la IA no debe hacer surgir desigualdades sociales, pues se intuye que su uso debe ser responsable y ético, para lo cual se piensa en establecer marcos normativos.

El futuro de la IA es una energía potencial que va a impulsar el desarrollo de tecnologías emergentes centradas en la computación cuántica, como en la computación evolutiva, a sabiendas que es un gran reto ya que lleva a que se presenten nuevas oportunidades y desafíos para su desarrollo.

Distintas críticas e inquietudes conducen a elaborar marcos éticos, políticas y directrices que determine caminos para un desarrollo responsable y ético. Aunque los estudios tienen de por sí una alta colaboración interdisciplinaria y multidisciplinaria al considerar que varias disciplinas pueden aportar elementos teóricos y experiencias, el fin es garantizar una integración beneficiosa y ética en la sociedad.

Los distintos mecanismos con IA lleva a que se realicen estudios hacia distintas áreas del saber, lo que genera amplias posibilidades que llevan a transformar varios ámbitos, sin embargo, al realizar un análisis amplio se perciben reflexiones en distintas direcciones, que surgen con respecto a las diferentes concepciones, además de las teorías y métodos que se plantean [Cao, 2022], por tanto, el propósito solo muestra cuál es ese potencial de la IA, sus aplicaciones y consideraciones éticas, pues es necesario determinar y reducir los diferentes sesgos que puedan producirse al tener una inadecuada selección de datos, a la vez considerar en qué puede alterar la privacidad de las personas [Cao, 2022].

## DESARROLLO DE LA INTELIGENCIA ARTIFICIAL

El surgimiento de la IA es la muestra de que el ser humano aplica sus capacidades y emprende una búsqueda incansable por obtener más conocimientos, utiliza su imaginación y la creatividad, entonces, puede realizar lo que se propone.

Durante el recorrido por diferentes caminos, etapa a etapa, la IA elabora modelos, técnicas y métodos para avanzar fuertemente, y al mismo tiempo cambia de paradigma ante la necesidad de construir un gran tejido de teorías y aplicaciones.

Son varios los hitos significativos que se han dado para llegar al punto en que se halla la IA hoy



día, en esta sociedad de la información. Son varios los desarrollos clave, donde la creatividad e imaginación, como el ingenio y la dedicación de las personas que plantean los distintos conceptos teóricos y aplicaciones prácticas, para que construyan sistemas inteligentes (SI), cada día con mayor potencial.

1.- El concepto de IA se formuló en la reunión de Dartmouth en 1956. Allí John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Shannon, entre otros, acuñaron el término [McCarthy et al., 2006].

2.- Se realiza razonamiento simbólico y sistemas para expertos en las décadas de 1970 y 1980. Los investigadores Newell y Simon [2012] realizaron contribuciones significativas a la IA simbólica en la resolución de problemas complejos. Los sistemas para expertos, como MYCIN, DENDRAL, de diagnóstico médico, mostró potencial de los sistemas basados en reglas y el razonamiento simbólico [Floridi, 2019]. Fueron varios los sistemas para expertos, aunque hoy día se les denomina sistemas expertos o sistemas basados en conocimiento, pues dependen del gran conocimiento que se les incluya en las llamadas bases de conocimiento.

3.- Al nacer el conexionismo dio pie a las redes neuronales en las décadas de 1980 y 1990. El algoritmo de retropropagación, un algoritmo de aprendizaje impulsó los avances en el aprendizaje de máquina. Geoffrey Hinton y sus colegas ampliaron a arquitecturas de aprendizaje profundo, que lo poseen los sistemas modernos de IA [LeCun, 2015]. Son varios los tipos de redes neuronales artificiales (RNA) que se han elaborado.

4.- Surge Big Data, que fortalece el aprendizaje de máquina.

A partir de esto surgen herramientas que ayudan en internet, además que mejoran las tecnologías digitales. Los algoritmos de aprendizaje de máquina basados en métodos estadísticos llevaron al reconocimiento de patrones, procesamiento del lenguaje natural y visión artificial, entre otros hechos. Pero la mejora continua de las RNA conduce a elaborar el aprendizaje profundo que, desde luego es producto de la mejora de los algoritmos y las mismas RNA, por tanto, se multiplica el uso de los sistemas inteligentes ya con aprendizaje profundo [Janiesch, et al., 2021]. Se automatiza la construcción de modelos analíticos con datos de entrenamiento específicos del dominio, llevando a grandes desafíos, en particular para el mercado electrónico y la máquina humana.

5.- Se elaboran nuevas teorías sobre el aprendizaje por refuerzo (parte del aprendizaje de máquina). Su éxito, como AlphaGo emerge al poder derrotar a campeones en juegos complejos (ajedrez, damas chinas, Go, ...), a la vez se realizan sistemas para que soporten la toma de decisiones estratégicas. También se desarrollan los modelos generativos que, de cierta manera, superan varios límites de la misma IA al generar contenidos realistas (imágenes, texto y música) [Goodfellow, et al., 2014].

6.- Se realiza integración de la IA con otras tecnologías, por ejemplo, análisis de BigData, computación en la nube e Internet de las cosas (IoT) que abren alternativas para el desarrollo de aplicaciones con IA. También se elabora hardware específico para la IA, como unidades de procesamiento de gráficos (GPU), las unidades de procesamiento de tensores (TPU), aunque con dificultades por las tareas de inferencia y su entrenamiento.

7.- Surgen tendencias emergentes, pues uno de los propósitos es poder explorar la XAI, cuyo fin es proporcionar transparencia e interpretabilidad a los modelos de IA [Arrieta, et al., 2020], ya que se requiere mejorar el desarrollo de sistemas inteligentes autónomos en el área de la salud, el hogar, el control, quizá mediante robots, así como la comprensión del lenguaje natural [Gunning, 2017]. Sobre todo, se reflexiona sobre la necesidad de esos sistemas inteligentes justos, éticos y con valores humanos.

El desarrollo de los mecanismos con IA es esfuerzo colectivo de distintos investigadores (médicos, psicólogos, ingenieros, físicos, matemáticos, entre otros), por tanto, de equipos multidisciplinarios, que trabajan por ampliar facilidades de accesibilidad a los diferentes recursos computacionales. En cada nueva etapa surgen mejores mecanismos con IA, siendo siempre producto de



los anteriores logros y, hoy día, la IA es una fuerza energética que transforma diversos dominios.

El esfuerzo mancomunado de los distintos profesionales conduce a que la IA sea una tecnología esencial en diversas tareas del ser humano, implementándose en los sistemas inteligentes mayores capacidades. Se implementan para exploración del universo, en sondas o telescopios, así como para explorar las profundidades de los océanos.

## Aplicaciones de la inteligencia artificial

La IA, hoy día, es un conjunto de software (herramientas) para esta era digital, en la sociedad de la información, insertándose distintas herramientas en los dispositivos para que se utilicen en variados sectores de la economía, tal vez alterando su curso de acción. La versatilidad de las aplicaciones con IA lleva a que se cambien los enfoques comunes y rutinarios, al igual que se varíe la forma de realizar distintas tareas, pero lo esencial como soporte para la toma de decisiones es que la mejora, tanto en la persona como en las organizaciones.

Con aplicaciones de IA se profundiza en los distintos ámbitos con grandes beneficios, en especial en dominios críticos: atención médica, finanzas, gestión de recursos humanos (HRM), entretenimiento, el transporte, la fabricación, exploración riesgosa, según características propias en cada uno [Torres-Soler, 1996].

Utilizar mecanismos con IA refleja la amplia adaptabilidad y flexibilidad que presentan, así como su potencial en las distintas empresas.

– Salud: Contribuciones significativas mejoran los diagnósticos, el tratamiento personalizado y la asistencia eficiente. En dermatología mediante el aprendizaje profundo determina clases de cáncer en la piel [Esteva et al., 2017]. Se utiliza también para predecir y guiar la toma de decisiones clínicas [Rajkomar, Dean y Kohane, 2019]. Se construyen modelos para elaborar fármacos según particularidades de las enfermedades o molestias en el organismo [Tan et al., 2023], se realiza análisis de imágenes [Litjens et al., 2017], se mejora la detección oportuna y el diagnóstico de enfermedades. Se desarrollan distintos dispositivos, scanner para toma de imágenes en los distintos órganos de los seres vivos.

– Finanzas: en las entidades financieras se emplea la IA para la detección de fraudes, evaluación de riesgos, y servicios financieros personalizados, en general, mediante el aprendizaje profundo que permite identificar transacciones fraudulentas [Pozzolo et al., 2014], y empleándose redes neuronales convolucionales se hace predicción del mercado de valores [Krizhevsky et al., 2017]. Se utilizan chatbots para ofrecer servicios financieros personalizados.

– Transporte: Se desarrollan vehículos autónomos con gestión inteligente para el transporte, además SI para organizar el tráfico, e igualmente se emplea la IA para el mantenimiento predictivo de los vehículos. En las diversas aplicaciones se emplea el aprendizaje profundo para reconocer adecuadamente los caminos, además para reducir la congestión y mejorar la eficiencia en los viajes.

– Educación: se elaboran sistemas inteligentes para que las personas aprendan de forma autónoma. El sistema hace seguimiento al rendimiento, por tanto, indica al aprendiz (usuario) cuáles son las fallas, dónde debe realizar refuerzo conceptual, e induce a la persona a autoorganizarse para autoaprender.

– Producción: la IA mejora los procesos de producción en las fábricas y en distintas empresas que producen diferentes productos, mejora su eficiencia, el control de calidad y realizar mantenimiento predictivo a las distintas máquinas, esto lleva a reducir costos al optimizar el uso de las materias primas, además que se mejora la calidad del producto [Lee et al., 2018].

Se emplea distinto tipo de robots para la automatización con IA en tareas de montaje, selección y embalaje, mejorando particularmente la seguridad de las personas. La IA allí ayuda a identificar posibles



fallas de los equipos y, por tanto, se optimiza su mantenimiento.

– Gestión de recursos humanos: se automatizan y mejoran las funciones de recursos humanos, por ejemplo, para la contratación, su incorporación y el seguimiento del rendimiento. Se emplean programas con IA para análisis predictivos que respalden la toma de decisiones, se usan chatbots para atender consultas rutinarias. Sin embargo, acá es donde las personas se restringen a usar esos programas creyendo en un posible despido de su labor, por lo que genera un optimismo poco realista ante la influencia de la IA [Weber et al., 2018], pero por los beneficios que ofrece, la reducción de costos y la eficiencia en las diferentes tareas, además de las perspectivas hacia escenarios futuros, se considera que es alto el impacto de la IA en las organizaciones.

– Entretenimiento: Diferentes programas con IA se desarrollan para elaborar escenarios, que el usuario tome diferentes roles, lo que lleva a mejorar las experiencias de los usuarios y, sobre todo, por la generación de contenidos. La capacidad de la IA puede deberse a la inclusión de distintos agentes inteligentes en un SI, por lo que lleva a que el juego logre mayor nivel, pues esos agentes recomiendan acciones, preferencias y analizan el comportamiento del usuario. Además, los agentes inteligentes incluyen algoritmos de generación de contenidos para elaborar imágenes, música y personajes virtuales.

Son varias las aplicaciones de la IA, pero solo se sintetizan algunas de ellas para visualizar en algo el impacto y su expansión en diferentes sectores.

Son muchas las aplicaciones a enunciar, considerando que la actitud creativa, la inteligencia y la destreza para innovar del ser humano lo impulsa de forma desmesurada a la búsqueda de mejoras para su calidad de vida, pero a la vez, para ser parte de una transformación digital y la disrupción en varias organizaciones.

Las capacidades de la IA indican el inicio de una nueva era de mayor eficiencia, toma de decisiones inteligente y eficiencia super sin precedentes en diferentes tareas, lo que induce a que se generen cambios en las personas y en las organizaciones.

A medida que se mejoran los métodos, técnicas y los algoritmos de IA, se busca adaptarlos a contextos diversos, pues su construcción tiene el fin de mejorar el potencial para nuevas innovaciones, esto lleva a considerar que los escenarios futuros en esta sociedad estarán supeditados a las tecnologías inteligentes, las cuales darán nuevas formas al mundo en que se vive, llevará, de cierto modo, a cambiar la forma de vivir, pues se contará con estufas, neveras, televisores inteligentes, además con dispositivos que indican de manera continua qué y cómo se deben hacer ciertas tareas.

## Consideraciones éticas de la IA

La rápida evolución e integración de la IA en diversos ámbitos del quehacer humano genera ciertas inquietudes en torno a la ética y la privacidad de las personas. Ante esto surgen amplias críticas, particularmente por su uso desafortunado en que se aniquila la privacidad de las personas [Mittelstadt et al., 2016].

Es necesario esbozar una exploración a los distintos desafíos, determinando la complejidad que existe en diversas situaciones ante el sesgo en los datos que se disponen para ciertas tareas, puede depender de los algoritmos que se construyen, que no enmarcan cómo debe ser los datos, ni de qué datos no disponer, por lo que está implícito la implicación sobre la privacidad, la responsabilidad o el impacto social, en particular por las tecnologías con IA y las emergentes que se desprenden hacia un futuro más prometedor.

Existe variedad de dilemas éticos que exigen la comprensión amplia para un panorama ético que rodea a los mecanismos con IA.

– Sesgo: en general el entrenamiento de los mecanismos con IA se realiza con volumen de datos,



algunos escogidos al azar, pero que pueden incluir sesgo, los algoritmos pueden estar inacabados, se llegan a utilizar técnicas en desuso o defectuosas para los nuevos desarrollos, esto puede llevar a la discriminación y la injusticia en la toma de decisiones por parte de esos sistemas inteligentes.

Con esto pueden suceder prejuicios raciales y de género, particularmente en el reconocimiento facial, generándose tasas de error mayores para las mujeres y las personas con tono de piel oscuro [Buolamwini y Gebru, 2018], además que estos pueden propagarse generando otros [Rumelhart, et al., 1986]. Para enfrentar los sesgos, se requiere de mayor cuidado en la selección de datos, realizar entrenamientos con conjunto de datos diversos e inclusivos, además evaluar de forma continua esos sistemas, tanto en sus métodos como en los algoritmos, tarea que es compleja.

La responsabilidad en la toma de decisiones sociales debe concebirse bajo el principio democrático de la "razón pública" [Binns, 2018] para que proporcione justificación sobre lo que genera el sistema, y de esta manera, facilitar una evaluación más precisa de los esfuerzos que se realizan.

– Privacidad y protección de datos: la IA utiliza volúmenes grandes de datos, en general personales y sensibles, por tanto, preocupa sobre la privacidad y la protección de datos tanto individuales como de las organizaciones, por lo que se necesita proteger la privacidad de las personas, además dar mayor seguridad, dado que muchos de los datos se almacenan en la nube, donde surgen ciberdelinquentes.

La protección de los datos personales no solo debe contar con un consentimiento informado, sino que debe implementarse medidas de seguridad sólidas que garanticen la privacidad y se cumpla con la normativa de protección de datos.

– Transparencia: Los sistemas inteligentes a menudo funcionan como "cajas negras", lo que dificulta su comprensión y el porqué de una toma de decisiones, sin embargo, se implementa en los SI diversos métodos para que fuesen a ser sustraídos. Y esa falta de transparencia y explicabilidad puede obstaculizar la confianza. No obstante, ante esto se le da importancia a la XAI e IAG (inteligencia artificial generativa), que tiene como propósito dar explicación sobre las decisiones de los sistemas inteligentes. Para ello se mejoran las técnicas para lograr que la XAI ayude a los usuarios a comprender cuál es el razonamiento que realizan los sistemas inteligentes que se emplean [Weber et al., 2023], así se amplía la confianza y se mitigan posibles sesgos o errores.

– Responsabilidad: determinar la responsabilidad en los sistemas de IA es complejo, quizá depende de que tanta autonomía se le provea para tomar decisiones.

– Impacto social y económico: no se duda de las ventajas de la IA, pero tiene su impacto no solo en lo técnico, también en lo económico, lo social, lo cognitivo del ser humano. La IA ante su potencial en cuanto puede realizar operaciones como las hace el ser humano, realiza procesos en las industrias que alteran diferentes rutinas, en particular impactan en el empleo, exacerbando desigualdades económicas, pues esas tecnologías pueden tener costos altos según los intereses en las organizaciones.

– Colaboración humana e IA: lo ético en la IA implica comprender el papel de los sistemas inteligentes y la colaboración humana, pues debe preservarse la autonomía, la toma de decisiones y los valores humanos en sus aplicaciones. En varios SI se requiere supervisión humana, por tanto, los marcos de colaboración entre humanos y los sistemas inteligentes no se reduce a entender unas instrucciones de cómo emplearlos y cuál puede ser el mejoramiento a realizar, porque la intervención humana debe inducir a resultados responsables y beneficiosos para toda la sociedad [Floridi et al., 2021]. Al fin de cuentas, lo esencial es aprovechar las fortalezas de la IA respetando los derechos humanos, la dignidad y los valores.

Son varias las consideraciones éticas a tener en cuenta, pues ante todo se debe garantizar la equidad, la inclusión y distribución de los beneficios y riesgos asociados a la IA [Castro-Yama y Torres-Soler, 2026]. La importancia de abordar los desafíos en la sociedad planteados por la IA se percibe en el empleo, la educación y la justicia social.

Esto induce a elaborar marcos sobre lo ético, que determine bajo la necesidad de una



responsabilidad compartida entre desarrolladores y los usuarios de tales sistemas.

Es crucial establecer escenarios amplios y precisos para un futuro, en donde los posibles marcos legales y pautas éticas permitan determinar responsabilidades y, por tanto, garantizar un desarrollo y despliegue de responsables de la IA [Prem, 2023].

Es claro que existe complejidad para abordar cuáles son las consideraciones éticas a tener en cuenta en el ámbito de la IA [Torres-Soler, 1996], pero se logra, con pequeñas falencias al clasificar las que enmarcan los investigadores en los diferentes contextos, por lo que se requiere de miradas y enfoques desde múltiples disciplinas y en direcciones diferentes.

Debe considerarse que es muy necesaria la participación activa de expertos, no solo en IA, sino también en ética, derecho y ciencias sociales, entre otras. La interacción de diversos campos lleva a aprovechar múltiples conocimientos para ser eficaces y eficientes ante el intrincado panorama ético que se requiere.

Pero una vez se tengan los marcos éticos, las directrices, y los marcos normativos determinan cómo dar cumplimiento, que no es tan fácil, pues las distintas medidas serán posibles si se determina en qué momento y por qué causas se viola la privacidad, de todas formas, lo ético es pilar para un diseño responsable, un desarrollo, y un despliegue de sistemas inteligentes, y de esa forma garantizar que verdaderamente contribuyan a la sociedad, de manera ética, respetando todo un conjunto de conocimientos y creencias en la sociedad [Jobin, et al., 2019].

## Inteligencia artificial y sociedad

La integración de la IA en distintos ámbitos de la vida del ser humano marca el inicio de una era digital y transformadora que conduce a realizar distintos cambios, los cuales pueden traer posibles consecuencias para la sociedad. Quizá existe influencia alta de la IA en múltiples aspectos del cotidiano vivir del ser humano, no obstante, puede ser que todavía no esté claro. Hay que mirar las repercusiones en el empleo, pues se automatizan distintas tareas, lo que cambia el mercado laboral y desde luego el bienestar de la familia, que repercute en la sociedad.

La inclusión de la IA también toca a la educación, por lo que los procesos educativos deben realizar distintos cambios, y desde allí a una gobernanza, a varios desafíos sociales. Se requiere relacionar las distintas oportunidades y desafíos que la tecnología trae consigo, las cuales deben categorizarse según diferentes enfoques.

– Empleo: la inclusión de tecnologías de IA suscita desplazamiento de puestos de trabajo, por tanto, las organizaciones disminuyen el número de empleados, esto pone de manifiesto la necesidad de autoaprender, autoorganizarse cada persona para asumir retos en nuevas labores. Es claro que no todos los trabajos pueden automatizarse, pero los que sí, debe preverse nuevas oportunidades, por tanto, mejora de las capacidades. La transición de un puesto de trabajo a otro produce crisis y críticas, más por ser reemplazados por máquinas con IA.

– Educación: La educación sufre transformación pues la IA lleva a personalizar el aprendizaje, teniendo en beneficio de mejorar los resultados educativos. Las herramientas con IA ayudan a docentes, a automatizar las tareas administrativas y, en particular, facilitar escenarios de aprendizaje personalizados para todas las personas. Sin embargo, se debe mirar lo ético y la privacidad que debe tenerse con los datos de los estudiantes y la toma de decisiones en la educación. Se desarrollan distintas herramientas para educar con IA, pero en sí, se debe tener grupos de control. Esto induce a realizar estudios amplios sobre los sistemas inteligentes en la educación.

– Gobernanza: La IA es una gran ayuda que presenta desafíos y oportunidades para la gobernanza y la formulación de políticas, su papel está en la mejora de la prestación de servicios públicos,



la mejora de los procesos de toma de decisiones, facilitando la gobernanza predictiva [Kshetri, 2017]. Sin embargo, los problemas relacionados con el sesgo deben abordarse para garantizar que los sistemas de IA sirvan al interés público. Es decir, se necesita de desarrollo de marcos regulatorios, directrices éticas y estándares para la gobernanza [Castro-Yama y Torres-Soler, 2019].

– Desafíos sociales: la inclusión de la IA plantea desafíos sociales que requieren cuidados. En particular identificar su potencial, los posibles sesgos y la discriminación existente, por lo que se necesita abordar distintas cuestiones: a) el sesgo de los datos, b) la equidad, y c) la inclusión, el diseño y la implementación de la IA. Los desafíos sociales se marcan por preocupaciones sobre el impacto de la IA en la privacidad, la autonomía humana y las interacciones sociales.

La integración de la IA en la sociedad plantea implicaciones éticas y legales que requieren atención. Lo ético debe garantizar la equidad, la transparencia y el desarrollo y uso responsables de las tecnologías de IA. En los marcos jurídicos deben estar cuestiones como la protección de la privacidad, la responsabilidad en el contexto de la IA.

Es necesario realizar innovaciones con ética y normas legales que lleven a que los beneficios de la IA en la sociedad sean amplios, por tanto, hay que realizar grandes esfuerzos por reducir las implicaciones sociales negativas.

La colaboración entre los responsables de la formulación de políticas, los estudiosos y las partes interesadas de diferentes ámbitos es esencial para los marcos éticos, jurídicos y sociales que rodean a la IA.

## Desafíos futuros en inteligencia artificial

El área de la IA es dinámica, se halla en primera línea del progreso tecnológico. Ofrece amplias oportunidades y formidables desafíos. Esto requiere una visión prospectiva de forma perspicaz sobre los horizontes futuros y los obstáculos que se presentan y que deben analizarse en procesos investigativos sobre el desarrollo de la IA.

En la medida que la IA avanza a pasos de gigante, deben concebirse cuáles son los escenarios inminentes que permiten mayor aplicabilidad, al mismo tiempo que se deben observar los desafíos críticos los cuales exigen soluciones innovadoras que lleven a navegar por aguas inexploradas, porque su potencial puede considerarse ilimitado, aunque existe complejidad ante posibles desafíos en el futuro.

– XAI: el fin primordial es mejorar la transparencia y la interpretabilidad de los sistemas de IA, en particular para que los usuarios comprendan cómo toman las decisiones esos sistemas. Se requiere analizar los enfoques basados en reglas, explicaciones y técnicas específicas. El desarrollo robusto y los métodos fiables de XAI son esenciales para generar confianza, garantizar los procesos y abordar las preocupaciones con respecto a la toma de decisiones de IA.

– Aprendizaje: surge el aprendizaje federado que es entrenar modelos de forma colaborativa a través de dispositivos distribuidos o bases de datos, eso sí teniendo en cuenta la necesidad de privacidad de los datos. Debe elaborarse un marco para el aprendizaje federado, pues los modelos se entrenan con datos descentralizados. Este tipo de aprendizaje ofrece caminos adecuados para el avance de la IA en ámbitos sensibles a la privacidad, como la salud y las finanzas, respetando la normativa de protección de datos.

– Computación cuántica: la computación cuántica tiene el potencial de revolucionar la IA al facilitar computación y resolución de problemas complejos. Un ordenador cuántico puede resolver un problema específico más allá del alcance de los computadores clásicos [Weber et al., 2018]. Pese a que la computación cuántica todavía se encuentra en sus primeras etapas, explorar su integración con la IA es prometedor para abordar en la IA, complejos desafíos y desarrollar aplicaciones de la IA.



– IA ética y responsable: desde luego que el desarrollo y el uso de la IA debe ser ético y responsable, pese a ser un desafío crítico. Debe estipularse muy bien problemas de sesgo, protección de la privacidad y transparencia. Los marcos éticos y las directrices desempeñan un papel vital en proceso investigativos en IA y su puesta en marcha [Floridi et al., 2021], pues debe existir un equilibrio entre la innovación y la IA responsable ya que se requiere mitigar riesgos potenciales y maximizar los beneficios sociales de la IA.

– IA en dominios complejos: el desarrollo de la IA también se extiende a dominios complejos: robótica y la comprensión del lenguaje natural, que son desafíos únicos, aunque es preciso aclarar que los esfuerzos mejoran la comprensión del lenguaje [Jouppi et al., 2017], así como el diálogo y la traducción automática. Abordar la complejidad de los escenarios del mundo real lleva a la comprensión del contexto, la ambigüedad y el conocimiento del dominio, algo muy esencial para mayor rendimiento en sistemas de IA.

– Privacidad y seguridad de los datos: la IA se soporta en volúmenes de datos a gran escala, donde la privacidad de los datos y la seguridad son retos interesantes. Se necesita proteger la información confidencial, reducir las violaciones a los datos, lo que genera miradas una computación multipartita segura y a la computación homomórfica cifrada.

Abordar las distintas dimensiones y los desafíos futuros en el área de la IA requiere una estrecha colaboración entre investigadores, políticos, académicos y partes interesadas de diversos ámbitos, los que trabajan en IA, en ética, en derecho y campos específicos de las distintas aplicaciones. Ese esfuerzo colaborativo para que el impacto de la IA abarque un amplio espectro, inmiscuyéndose en diversos aspectos de la sociedad.

Para sortear eficazmente los diferentes desafíos, se necesita un compromiso para realizar estudios de forma continua y así estar a la vanguardia de ese mundo cambiante de la IA. Además, la inclusión del diálogo abierto y las iniciativas proactivas que se plantean son puntos para desarrollar tecnologías de IA responsables, inclusivas y beneficiosas, pero alineados con principios éticos y priorización del bienestar de la sociedad. De este modo la IA al avanzar es mejora para la humanidad, por la reducción de tiempo en diferentes tareas y la mitigación de riesgos potenciales y desafíos.

## CONSIDERACIONES FINALES

No es posible mostrar en pocas líneas el amplio panorama de la IA, porque desde sus inicios hasta hoy día existen diferentes hitos ante la gran evolución de sus métodos y aplicaciones, las que generan amplias perspectivas futuras, pero a la vez se percibe su profundo impacto en la sociedad que induce a tener protectores y contradictores ante la falta de un marco ético que determine que las aplicaciones no traerán consigo perjuicios para la sociedad.

El acelerado progreso de la IA revoluciona y mejora varios aspectos del quehacer diario del ser humano, no obstante, se presentan inquietudes éticas, quizá morales.

El equilibrio entre la innovación y la responsabilidad ética de la IA es esencial para garantizar que las distintas aplicaciones formen una sociedad positiva, una sociedad que se una para el mismo desarrollo tecnológico.

El futuro de la IA ofrece posibilidades fascinantes, pero también desafíos que requieren una atención diligente, desde el desarrollo responsable hasta la privacidad de los datos y las implicaciones sociales deben tener cuidados extensivos.

El desarrollo de la IA y su impacto en la sociedad no se determina sin una colaboración interdisciplinaria entre investigadores, los responsables de la formular políticas y las partes interesadas,



pues hace surgir la necesidad de establecer marcos éticos y normativas que atañen a la configuración de caminos de la IA.

La IA sigue evolucionando y el compromiso de los seres humanos es aprovechar su potencial para el beneficio común, guiado por la equidad, la transparencia y la inclusión, hacia un futuro marcado por una sociedad más equitativa y sostenible.

## REFERENCIAS

ARRIETA, A.B., et al. “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI”. En: *Information Fusion*, 2020, vol. 58, pp. 82–115. doi: 10.1016/j.inffus.2019.12.012.

CAO, L. “AI in Finance: Challenges, Techniques, and Opportunities”. En: *ACM Computing Surveys*, 2022, vol. 55, no. 3, pp. 1–38. doi: 10.1145/3502289.

ESTEVA, A. et al. “Dermatologist-level classification of skin cancer with deep neural networks”. En: *Nature*, 2017, vol. 542, no. 7639, pp. 115–118. doi: 10.1038/nature21056.

BINNS, R. “Algorithmic Accountability and Public Reason”. En: *Philosophy and Technology*, 2018, vol. 31, no. 4, pp. 543–556. doi: 10.1007/s13347-017-0263-5.

BUOLAMWINI, J. y Gebru, T. “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification”. En: *Proceedings of Machine Learning Research*, 2018, vol. 81, pp. 1–15.

CASTRO-YAMA, S.Y. y Torres-Soler, L.C. Los algoritmos y la ética: puntos para reflexionar. En: CASTRO-YAMA, S., et al. *METATONIA*. pp. 181-217. Sao Paulo: Improving Performance Editora, 2026.

CASTRO-YAMA., S.Y. y TORRES-SOLER, L.C. *Algoritmos en la Administración pública*. Impretics, Cali, 2019.

FLORIDI, E. et al. “An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations”. En: *Philosophical Studies Series*, 2021, vol. 144, pp. 19–39. doi: 10.1007/978-3-030-81907-1 3.

FLORIDI, L. “Establishing the Rules for Building Trustworthy AI”. En: *Nature Machine Intelligence*, 2019, vol. 1, no. 6, pp. 261–262, Jun. doi: 10.1038/s42256-019-0055-y.

GERKE, S., Minssen, T. y COHEN, G. “Ethical and legal challenges of artificial intelligence-driven healthcare”. En: *Artificial intelligence in Healthcare*, Massachusetts, USA: Academic Press, pp. 295–336, 2020.

GOODFELLOW, I.J., et al. “Generative Adversarial Nets”. En: *Advances in Neural Information Processing Systems*, 2014, vol. 3, no. January, pp. 2672–2680, doi: 10.1007/978-3-658-40442-0 9.

GUNNING, D. “Explainable artificial intelligence (xai)”. En: *Defense Advanced Research Projects Agency (DARPA)*, 2017, vol. 2, no. 2, pp. 1–36.

JANIESCH, C., ZSCHECH, P. y HEINRICH, K. “Machine Learning and Deep Learning”. En: *Electronic Markets*, 2021, vol. 31, no. 3, pp. 685–695. doi: 10.1007/s12525-021-00475-2.

JOBIN, A., LENCA, M. y VAYENA, E. “The Global Landscape of AI Ethics Guidelines”. En: *Nature Machine Intelligence*, 2019, vol. 1, no. 9, pp.389–399. doi: 10.1038/s42256-019-0088-2.



JOUPPI, N.P., et al. “In-datacenter Performance Analysis of a Tensor Processing Unit”. En: *Proceedings of the 44th Annual International Symposium on Computer Architecture*, 2017, pp. 1–12. doi: 10.1145/3079856.3080246.

KRIZHEVSKY, A., SUTSKEVER, I. y HINTON, G. E. “ImageNet Classification with Deep Convolutional Neural Networks”. En: *Communications of the ACM*, 2017, vol. 60, no. 6, pp. 84–90. doi: 10.1145/3065386.

KSHETRI, N. “Can Blockchain Strengthen the Internet of Things?”. En: *IT Professional*, 2017, vol. 19, no. 4, pp. 68–72. doi: 10.1109/MITP.2017.3051335.

LECUN, Y., BENGIO, Y. y HINTON, G. “Deep Learning”. En: *Nature*, 2015, vol. 521, no. 7553, pp.436–444. doi: 10.1038/nature14539.

LEE, J., DAVARI, H., SINGH, J. y PANDHARE, V. “Industrial Artificial Intelligence for Industry 4.0-Based Manufacturing Systems”. En: *Manufacturing Letters*, 2018, vol. 18, pp. 20–23. doi: 10.1016/j.mfglet.2018.09.002.

LITJENS, G. et al. “A Survey on Deep Learning in Medical Image Analysis”. En: *Medical Image Analysis*, 2017, vol. 42, no. 42, pp. 60–88. doi: 10.1016/j.media.2017.07.005.

MCCARTHY, J., MINSKY, M. L., ROCHESTER, N. y SHANNON, C. E. “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence”. En: *AI Magazine*, 2006, vol. 27, no. 4, pp. 12–14. doi: 10.1609/aimag.v27i4.1904.

MITTELSTADT, B. D., ALLO, P., TADDEO, M., WACHTER, S. y FLORIDI, L. “The Ethics of Algorithms: Mapping the Debate”. En: *Big Data and Society*, 2016, vol. 3, no. 2, pp. 1–21. doi: 10.1177/2053951716679679.

NEWELL, A. y SIMON, H.A. *Human Problem Solving*. 3ra. ed., Prentice Hall, USA, 2012.

POZZOLO, A. D., CAELEN, O., LE BORGNE, Y. A., WATERSCHOOT, S. y BONTEMPI, G. “Learned Lessons in Credit card Fraud Detection from a Practitioner Perspective”. En: *Expert Systems with Applications*, 2014, vol. 41, no. 10, pp. 4915–4928. doi: 10.1016/j.eswa.2014.02.026.

PREM, E. “From Ethical AI Frameworks to Tools: a Review of Approaches”. En: *AI and Ethics*, 2023, vol. 3, no. 3, pp. 699–716, Aug. doi: 10.1007/s43681-023-00258-9.

RAJKOMAR, A., DEAN, J. y KOHANE, I. “Machine Learning in Medicine”. En: *Journal of Medicine*, 2019, vol. 380, no. 14, pp.1347–1358. doi: 10.1056/NEJMra1814259.

RUMELHART, D. E., HINTON, G. E. y WILLIAMS, R. J. “Learning Representations by Back-Propagating Errors.” En: *Nature*, 1986, vol. 323, no.6088, pp. 533–536. doi: 10.1038/323533a0.

SUTSKEVER, I., VINYALS, O. y LE, Q. V. “Sequence to Sequence Learning with Neural Networks”. En: *Proceedings of the 27th International Conference on Neural Information Processing Systems*, pp. 3104–3112, 2014.

TAN, P., CHEN, X., ZHANG, H., WEI, Q. y LUO, K. “Artificial Intelligence Aids in Development of Nanomedicines for Cancer Management”. En: *Seminars in Cancer Biology*, 2023, vol. 89, pp. 61–75. doi: 10.1016/j.semcancer.2023.01.005.

TORRES-SOLER, L. C. *Inteligencia Artificial: Qué es ...? Cómo ...?* Fundación Universitaria Manuela Beltrán, Bogotá, 1996.



WEBER, P., CARL, K. V. y HINZ, O. “Applications of Explainable Artificial Intelligence in Finance— a Systematic Review of Finance, Information Systems, and Computer Science literature”. En: *Management Review Quarterly*, 2023, pp. 1–41. doi: 10.1007/s11301-023-00320-0.

WEBER, P., PINSKI, M. y BAUM, L. “Toward an Objective Measurement of AI Literacy”. En: *Pacific-Asia Conference on Information Systems*, 2018, pp. 1–17.

*Submetido em abril de 2026*

*Aprovado em julho de 2026*

*Informações do (a) (s) autor(a)(es)*

*Nome:* Yama Sonia Yamilet Castro

*Afiliação Institucional:* Estudante de PhD Universidad de Barcelona

*Grau:* MsC. Economía, Especialista en Analítica de Datos

*e-mail:* [sonia.castroyama@hotmail.com](mailto:sonia.castroyama@hotmail.com)

*ORCID:* <https://orcid.org/0000-0003-1239-5268>

*Nome:* Luis Carlos Torres Soler

*Afiliação Institucional:* Universidade Autônoma da Colômbia

*Grau:* PhD. Pensamiento complejo, MsC. Ing. Sistemas, MA. Ciencias de la Educación, Matemático.

*Docente investigador*

*e-mail:* [lctorress@gmail.com](mailto:lctorress@gmail.com)

*ORCID:* <https://orcid.org/0000-0001-6756-4984>