

WIKIPEDIA

La enciclopedia libre

Inteligencia artificial

La **inteligencia artificial** (IA), en el contexto de las ciencias de la computación, es una disciplina y un conjunto de capacidades cognoscitivas e intelectuales expresadas por sistemas informáticos o combinaciones de algoritmos cuyo propósito es la creación de máquinas que imiten la inteligencia humana para realizar tareas, y que pueden mejorar conforme recopilen información.^{1 2} Se hizo presente poco después de la Segunda Guerra Mundial con el desarrollo de la «prueba de Turing», mientras que la locución fue acuñada en 1956 por el informático John McCarthy en la Conferencia de Dartmouth.

En la actualidad, la inteligencia artificial abarca una gran variedad de subcampos. Éstos van desde áreas de propósito general, aprendizaje y percepción, a otras más específicas como el Reconocimiento de voz, el juego de ajedrez, la demonstración de teoremas matemáticos, la escritura de poesía y el diagnóstico de enfermedades. La inteligencia artificial sintetiza y automatiza tareas que en principio son intelectuales y, por lo tanto, es potencialmente relevante para cualquier ámbito de actividades intelectuales humanas. En este sentido, es un campo genuinamente universal.³

La arquitectura de las inteligencias artificiales y los procesos por los cuales aprenden, se mejoran y se implementan en algún área de interés varía según el enfoque de utilidad que se les quiera dar, pero de manera general, estos van desde la ejecución de sencillos algoritmos hasta la interconexión de complejas redes neuronales artificiales que intentan replicar los circuitos neuronales del cerebro humano y que aprenden mediante diferentes modelos de aprendizaje tales como el aprendizaje automático, el aprendizaje por refuerzo, el aprendizaje profundo y el aprendizaje supervisado.⁴

Por otro lado, el desarrollo y aplicación de la inteligencia artificial en muchos aspectos de la vida cotidiana también ha propiciado la creación de nuevos campos de estudio como la roboética y la ética de las máquinas que abordan aspectos relacionados con la ética en la inteligencia artificial y que se encargan de analizar cómo los avances en este tipo de tecnologías impactarían en diversos ámbitos de la vida, así como el manejo responsable y ético que se les debería dar a los mismos, además de establecer cuál debería ser la manera correcta de proceder de las máquinas y las reglas que deberían cumplir.⁵

En cuanto a su clasificación, tradicionalmente se divide a la inteligencia artificial en inteligencia artificial débil, la cual es la única que existe en la actualidad y que se ocupa de realizar tareas específicas, e inteligencia artificial general, que sería una IA que excediese las capacidades humanas. Algunos expertos creen que si alguna vez se alcanza este nivel, se podría dar lugar a la aparición de una singularidad tecnológica, es decir, una entidad tecnológica superior que se mejoraría a sí misma constantemente, volviéndose incontrolable para los humanos, dando pie a teorías como el basilisco de Roko.⁶



Imagen generada por DALL-E 3.



Vídeo explicativo sobre la inteligencia artificial, incluyendo secciones sobre los dilemas éticos. Subtítulos disponibles.

Algunas de las inteligencias artificiales más conocidas y utilizadas en la actualidad alrededor del mundo incluyen inteligencia artificial en el campo de la salud, asistentes virtuales como Alexa, el asistente de Google o Siri, traductores automáticos como el traductor de Google y DeepL, sistemas de recomendación como el de la plataforma digital de YouTube, motores de ajedrez y otros juegos como Stockfish y AlphaZero, chatbots como ChatGPT, creadores de arte de inteligencia artificial como Midjourney, Dall-e, Leonardo y Stable Diffusion, e incluso la conducción de vehículos autónomos como Tesla Autopilot.⁷

Asimismo la inteligencia artificial se está desarrollando en la plataforma digital cada vez más, evolucionando y creando nuevas herramientas, como la plataforma laboral que existe desde el año 2023 llamada SIVIUM, una herramienta por la cual una persona postula en forma automatizada a todas las ofertas laborales de todos los portales de trabajo, sin necesidad de estar revisando cada oferta laboral que se presente y enviar su CV uno por uno.

Denominación

En 2019 la Comisión Mundial de Ética del Conocimiento Científico y la Tecnología (COMEST) de la UNESCO definió la inteligencia artificial como un campo que implica máquinas capaces de imitar determinadas funcionalidades de la inteligencia humana, incluidas características como la percepción, el aprendizaje, el razonamiento, la resolución de problemas, la interacción lingüística e incluso la producción de trabajos creativos.

Coloquialmente, la locución «inteligencia artificial» se aplica cuando una máquina imita las funciones «cognitivas» que los humanos asocian como competencias humanas, por ejemplo: «percibir», «razonar», «aprender» y «resolver problemas».⁸ Andreas Kaplan y Michael Haenlein definen la inteligencia artificial como «la capacidad de un sistema para interpretar correctamente datos externos, y así aprender y emplear esos conocimientos para lograr tareas y metas concretas a través de la adaptación flexible».⁹ A medida que las máquinas se vuelven cada vez más capaces, se elimina de la definición la tecnología que alguna vez se pensó que requería de inteligencia. Marvin Minsky, uno de los ideadores de la IA, hablaba del término inteligencia artificial como una palabra maleta ("*suitcase word*") porque en él se pueden meter una diversidad de elementos.^{10 11}



Ilustración generada mediante inteligencia artificial, con estilo de Acuarela, de Alan Turing, considerado uno de los padres fundadores de la IA ^[*cita requerida*].

Por ejemplo, el reconocimiento óptico de caracteres ya no se percibe como un ejemplo de la «inteligencia artificial» habiéndose convertido en una tecnología común.¹² Avances tecnológicos todavía clasificados como inteligencia artificial son los sistemas de conducción autónomos o los capaces de jugar ajedrez o Go.¹³

La inteligencia artificial es una nueva forma de resolver problemas dentro de los cuales se incluyen los sistemas expertos, el manejo y control de robots y los procesadores, que intenta integrar el conocimiento en tales sistemas, en otras palabras, un sistema inteligente capaz de escribir su propio programa. Un sistema experto definido como una estructura de programación capaz de almacenar y utilizar un conocimiento sobre un área determinada que se traduce en su capacidad de aprendizaje.¹⁴ De igual manera se puede considerar a la IA como la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano.¹⁵

Según Takeyas (2007) la IA es una rama de las ciencias computacionales encargada de estudiar modelos de cómputo capaces de realizar actividades propias de los seres humanos con base en dos de sus características primordiales: el razonamiento y la conducta.¹⁶

En 1956, John McCarthy acuñó la expresión «inteligencia artificial», y la definió como «la ciencia e ingenio de hacer máquinas inteligentes, especialmente programas de cómputo inteligentes».¹⁷

También existen distintos tipos de percepciones y acciones, que pueden ser obtenidas y producidas, respectivamente, por sensores físicos y sensores mecánicos en máquinas, pulsos eléctricos u ópticos en computadoras, tanto como por entradas y salidas de bits de un software y su entorno software.

Varios ejemplos se encuentran en el área de control de sistemas, planificación automática, la capacidad de responder a diagnósticos y a consultas de los consumidores, reconocimiento de escritura, reconocimiento del habla y reconocimiento de patrones. Los sistemas de IA actualmente son parte de la rutina en campos como economía, medicina, ingeniería, el transporte, las comunicaciones y la milicia, y se ha usado en gran variedad de programas informáticos, juegos de estrategia, como ajedrez de computador, y otros videojuegos.

Tipos

Stuart J. Russell y Peter Norvig diferencian varios tipos de inteligencia artificial:¹⁸

- Los sistemas que piensan como humanos: Estos sistemas tratan de emular el pensamiento humano; por ejemplo, las redes neuronales artificiales. La automatización de actividades que vinculamos con procesos de pensamiento humano, actividades como la toma de decisiones, resolución de problemas y aprendizaje.¹⁹
- Los sistemas que actúan como humanos: Estos sistemas tratan de actuar como humanos; es decir, imitan el comportamiento humano; por ejemplo, la robótica (El estudio de cómo lograr que los computadores realicen tareas que, por el momento, los humanos hacen mejor).²⁰
- Los sistemas que piensan racionalmente: Es decir, con lógica (idealmente), tratan de imitar el pensamiento racional del ser humano; por ejemplo, los sistemas expertos, (el estudio de los cálculos que hacen posible percibir, razonar y actuar).²¹
- Los sistemas que actúan racionalmente: Tratan de emular de forma racional el comportamiento humano; por ejemplo, los agentes inteligentes, que está relacionado con conductas inteligentes en artefactos.²²

Inteligencia artificial generativa

La inteligencia artificial generativa es un tipo de sistema de inteligencia artificial capaz de generar texto, imágenes u otros medios en respuesta a comandos. Los modelos de IA generativa aprenden los patrones y la estructura de sus datos de entrenamiento de entrada y luego generan nuevos datos que tienen características similares.

Los sistemas de IA generativa notables incluyen ChatGPT (y su variante Microsoft Copilot), un bot conversacional creado por OpenAI usando sus modelos de lenguaje grande fundacionales GPT-3 y GPT-4; y Bard, un bot conversacional creado por Google usando Gemini. Otros modelos generativos de IA incluyen sistemas de arte de inteligencia artificial como Stable Diffusion, Midjourney y DALL-E.

Inteligencia artificial fuerte

La **Inteligencia artificial fuerte (IGA)** es un tipo hipotético de inteligencia artificial que iguala o excede la inteligencia humana promedio. Si se hiciera realidad, una IGA podría aprender a realizar cualquier tarea intelectual que los seres humanos o los animales puedan llevar a cabo. Alternativamente, la IGA se ha definido como un sistema autónomo que supera las capacidades humanas en la mayoría de las tareas económicamente valiosas.

Algunos sostienen que podría ser posible en años o décadas; otros, que podría tardar un siglo o más; y una minoría cree que quizá nunca se consiga. Existe un debate sobre la definición exacta de IGA y sobre si los grandes modelos de lenguaje (LLM) modernos, como el GPT-4, son formas tempranas pero incompletas de IGA.

Inteligencia artificial explicable

La inteligencia artificial explicable se refiere a métodos y técnicas en la aplicación de tecnología de inteligencia artificial por los que el ser humano es capaz de comprender las decisiones y predicciones realizadas por la inteligencia artificial.

Inteligencia artificial amigable

La inteligencia artificial amigable es una IA fuerte e hipotética que puede tener un efecto positivo más que uno negativo sobre la humanidad. 'Amigable' es usado en este contexto como terminología técnica y escoge agentes que son seguros y útiles, no necesariamente aquellos que son "amigables" en el sentido coloquial. El concepto es invocado principalmente en el contexto de discusiones de agentes artificiales de auto-mejora recursiva que rápidamente explota en inteligencia, con el argumento de que esta tecnología hipotética pudiera tener una larga, rápida y difícil tarea de controlar el impacto en la sociedad humana.

Inteligencia artificial multimodal

La inteligencia artificial multimodal es un tipo de inteligencia artificial que puede procesar e integrar datos de diferentes modalidades, como texto, imágenes, audio y video, para obtener una comprensión más completa y contextualizada de una situación. La inteligencia artificial multimodal se inspira en la forma en que los humanos usan varios sentidos para percibir e interactuar con el mundo, y ofrece una forma más natural e intuitiva de comunicarse con la tecnología.

Inteligencia artificial cuántica

La inteligencia artificial Cuántica es un campo interdisciplinar que se enfoca en construir algoritmos cuánticos para mejorar las tareas computacionales dentro de la IA, incluyendo subcampos como el aprendizaje automático. Existen evidencias que muestran una posible ventaja cuadrática cuántica en operaciones fundamentales de la IA.

Escuelas de pensamiento

La IA se divide en dos escuelas de pensamiento:

- La inteligencia artificial convencional.
- La inteligencia computacional.

Inteligencia artificial convencional

Se conoce también como IA simbólica-deductiva. Está basada en el análisis formal y estadístico del comportamiento humano ante diferentes problemas:

- Razonamiento basado en casos: Ayuda a tomar decisiones mientras se resuelven ciertos problemas concretos y, aparte de que son muy importantes, requieren de un buen funcionamiento.
- Sistemas expertos: Infieren una solución a través del conocimiento previo del contexto en que se aplica y utiliza ciertas reglas o relaciones.²³
- Redes bayesianas: Propone soluciones mediante inferencia probabilística.²⁴
- Inteligencia artificial basada en comportamientos: Esta inteligencia contiene autonomía, es decir, puede auto-regularse y controlarse para mejorar.
- Smart process management: Facilita la toma de decisiones complejas, proponiendo una solución a un determinado problema al igual que lo haría un especialista en dicha actividad.

Inteligencia artificial computacional

La inteligencia computacional (también conocida como IA subsimbólica-inductiva) implica desarrollo o aprendizaje interactivo (por ejemplo, modificaciones interactivas de los parámetros en sistemas de conexiones). El aprendizaje se realiza basándose en datos empíricos.

La inteligencia computacional tiene una doble finalidad. Por un lado, su objetivo científico es comprender los principios que posibilitan el comportamiento inteligente (ya sea en sistemas naturales o artificiales) y, por otro, su objetivo tecnológico consiste en especificar los métodos para diseñar sistemas inteligentes.²⁵

Historia

- La expresión «inteligencia artificial» fue acuñada formalmente en 1956 durante la Conferencia de Dartmouth, pero para entonces ya se había estado trabajando en ello durante cinco años en los cuales se había propuesto muchas definiciones distintas que en ningún caso habían logrado ser aceptadas totalmente por la comunidad investigadora. La IA es una de las disciplinas más recientes junto con la genética moderna.
- Las ideas más básicas se remontan a los antiguos griegos. Aristóteles (384-322 a. C.) fue el primero en describir un conjunto de reglas que describen una parte del funcionamiento de la mente para obtener conclusiones racionales, y Ctesibio de Alejandría (250 a. C.) construyó la primera máquina autocontrolada, un regulador del flujo de agua (racional pero sin razonamiento).
- En 1315 Ramon Llull en su libro *Ars magna* tuvo la idea de que el razonamiento podía ser efectuado de manera artificial.
- En 1840 Ada Lovelace previó la capacidad de las máquinas para ir más allá de los simples cálculos y aportó una primera idea de lo que sería el software.
- Leonardo Torres Quevedo(1852-1936) es considerado como uno de los padres de la inteligencia artificial y de la Automática.
- En 1936 Alan Turing diseña formalmente una *Máquina universal* que demuestra la viabilidad de un dispositivo físico para implementar cualquier cómputo formalmente definido.
- En 1943 Warren McCulloch y Walter Pitts presentaron su modelo de neuronas artificiales, el cual se considera el primer trabajo del campo, aun cuando todavía no existía el término. Los primeros avances importantes comenzaron a principios del año 1950 con el trabajo de Alan Turing, a partir de lo cual la ciencia ha pasado por diversas situaciones.
- En 1955 Herbert Simon, Allen Newell y Joseph Carl Shaw, desarrollan el primer lenguaje de programación orientado a la resolución de problemas, el IPL-11. Un año más tarde desarrollan el LogicTheorist, el cual era capaz de demostrar teoremas matemáticos.

- En 1956 fue ideada la expresión «inteligencia artificial» por John McCarthy, Marvin Minsky y Claude Shannon en la Conferencia de Dartmouth, un congreso en el que se hicieron previsiones triunfalistas a diez años que jamás se cumplieron, lo que provocó el abandono casi total de las investigaciones durante quince años.
- En 1957 Newell y Simon continúan su trabajo con el desarrollo del General Problem Solver (GPS). GPS era un sistema orientado a la resolución de problemas.
- En 1958 John McCarthy desarrolla en el Instituto Tecnológico de Massachusetts (MIT) el LISP. Su nombre se deriva de LISt Processor. LISP fue el primer lenguaje para procesamiento simbólico.
- En 1959 Rosenblatt introduce el «perceptrón».
- A finales de la década de 1950 y comienzos de la de 1960 Robert K. Lindsay desarrolla «Sad Sam», un programa para la lectura de oraciones en inglés y la inferencia de conclusiones a partir de su interpretación.
- En 1963 Quillian desarrolla las redes semánticas como modelo de representación del conocimiento.
- En 1964 Bertrand Raphael construye el sistema SIR (*Semantic Information Retrieval*) el cual era capaz de inferir conocimiento basado en información que se le suministra. Bobrow desarrolla STUDENT.
- A mediados de los años 60, aparecen los sistemas expertos, que predicen la probabilidad de una solución bajo un set de condiciones. Por ejemplo, DENDRAL, iniciado en 1965 por Buchanan, Feigenbaum y Lederberg, el primer Sistema Experto, que asistía a químicos en estructuras químicas complejas, MACSYMA, que asistía a ingenieros y científicos en la solución de ecuaciones matemáticas complejas.
- Posteriormente entre los años 1968-1970 Terry Winograd desarrolló el sistema SHRDLU, que permitía interrogar y dar órdenes a un robot que se movía dentro de un mundo de bloques.
- En 1968 Marvin Minsky publica *Semantic Information Processing*.
- En 1968 Seymour Papert, Danny Bobrow y Wally Feurzeig desarrollan el lenguaje de programación LOGO.
- En 1969 Alan Kay desarrolla el lenguaje Smalltalk en Xerox PARC y se publica en 1980.
- En 1973 Alain Colmenauer y su equipo de investigación en la Universidad de Aix-Marseille crean PROLOG (del francés *PRO*grammation en *LOG*ique) un lenguaje de programación ampliamente utilizado en IA.
- En 1973 Shank y Abelson desarrollan los guiones, o scripts, pilares de muchas técnicas actuales en inteligencia artificial y la informática en general.
- En 1974 Edward Shortliffe escribe su tesis con MYCIN, uno de los Sistemas Expertos más conocidos, que asistió a médicos en el diagnóstico y tratamiento de infecciones en la sangre.
- En las décadas de 1970 y 1980, creció el uso de sistemas expertos, como MYCIN: R1/XCON, ABRL, PIP, PUFF, CASNET, INTERNIST/CADUCEUS, etc. Algunos permanecen hasta hoy (Shells) como EMYCIN, EXPERT, OPSS.
- En 1981 Kazuhiro Fuchi anuncia el proyecto japonés de la quinta generación de computadoras.
- En 1986 McClelland y Rumelhart publican *Parallel Distributed Processing* (Redes Neuronales).
- En 1988 se establecen los lenguajes Orientados a Objetos.
- En 1997 Gari Kaspárov, campeón mundial de ajedrez, pierde ante la computadora autónoma Deep Blue.
- En 2006 se celebró el aniversario con el Congreso en español 50 años de inteligencia artificial - Campus Multidisciplinar en Percepción e Inteligencia 2006.
- En 2009 ya había en desarrollo sistemas inteligentes terapéuticos que permiten detectar emociones para poder interactuar con niños autistas.
- En 2011 IBM desarrolló un superordenador llamado Watson, el cual ganó una ronda de tres juegos seguidos de Jeopardy!, venciendo a sus dos máximos campeones, y ganando un premio de 1 millón de dólares que IBM luego donó a obras de caridad.²⁶

- En 2016, un programa informático ganó cinco a cero al triple campeón de Europa de Go.²⁷
- En 2016, el entonces presidente Obama habla sobre el futuro de la inteligencia artificial y la tecnología.²⁸
- Existen personas que al dialogar sin saberlo con un chatbot no se percatan de hablar con un programa, de modo tal que se cumple la prueba de Turing como cuando se formuló: «Existirá inteligencia artificial cuando no seamos capaces de distinguir entre un ser humano y un programa informático en una conversación a ciegas».
- En 2017 AlphaGo desarrollado por DeepMind derrota 4-1 en una competencia de Go al campeón mundial Lee Sedol. Este suceso fue muy mediático y marcó un hito en la historia de este juego.²⁹ A finales de ese mismo año, Stockfish, el motor de ajedrez considerado el mejor del mundo con 3 400 puntos ELO, fue abrumadoramente derrotado por AlphaZero con solo conocer las reglas del juego y tras solo 4 horas de entrenamiento jugando contra sí mismo.³⁰
- Como anécdota, muchos de los investigadores sobre IA sostienen que «la inteligencia es un programa capaz de ser ejecutado independientemente de la máquina que lo ejecute, computador o cerebro».
- En 2017 un grupo de ingenieros en Google inventan la arquitectura de transformador, un modelo de deep learning que alumbró una nueva generación de modelos grandes de lenguaje, empezando por BERT, y luego el revolucionario GPT de OpenAI.³¹
- En 2018, se lanza el primer televisor con inteligencia artificial por parte de LG Electronics con una plataforma denominada ThinQ.³²
- En 2019, Google presentó su Doodle en que, con ayuda de la inteligencia artificial, hace un homenaje a Johann Sebastian Bach, en el que, añadiendo una simple melodía de dos compases la IA crea el resto.
- En 2020, la OECD (Organización para la Cooperación y el Desarrollo Económico) publica el documento de trabajo intitulado *Hola, mundo: La inteligencia artificial y su uso en el sector público*, dirigido a funcionarios de gobierno con el afán de resaltar la importancia de la IA y de sus aplicaciones prácticas en el ámbito gubernamental.³³
- Al final del año 2022, se lanzó ChatGPT, una inteligencia artificial generativa capaz de escribir textos y responder preguntas en muchos idiomas. Dado que la calidad de las respuestas recordaba inicialmente al nivel humano, se generó un entusiasmo mundial por la IA³⁴ y ChatGPT alcanzó más de 100 millones de usuarios dos meses después de su lanzamiento.³⁵ Más tarde, los expertos notaron que ChatGPT proporciona información errónea en áreas donde no tiene conocimiento ("alucinaciones de datos"), la cual a primera vista parece creíble debido a su perfecta redacción.³⁶
- En 2023, las fotos generadas por IA alcanzaron un nivel de realismo que las hacía confundirse con fotos reales. Como resultado, hubo una ola de "fotos" generadas por IA que muchos espectadores creyeron que eran reales. Una imagen generada por Midjourney se destacó, mostrando al Papa Francisco con un elegante abrigo blanco de invierno.³⁷

Implicaciones sociales, éticas y filosóficas

Ante la posibilidad de crear máquinas dotadas de inteligencia, se volvió importante preocuparse por la cuestión ética de las máquinas para tratar de garantizar que no se produzca ningún daño a los seres humanos, a otros seres vivos e incluso a las mismas máquinas según algunas corrientes de pensamiento.³⁸ Es así como surgió un amplio campo de estudios conocido como ética de la inteligencia artificial de relativamente reciente aparición y que generalmente se divide en dos ramas, la roboética, encargada de estudiar las acciones de los seres humanos hacia los robots, y la ética de las máquinas encargada del estudio del comportamiento de los robots para con los seres humanos.

Primera Ley

Segunda Ley

Tercera Ley

Un robot debe proteger su propia existencia en la medida en que esta protección no entre en conflicto con la primera o con la segunda ley.³⁹

Otras obras de ciencia ficción más recientes también exploran algunas cuestiones éticas y filosóficas con respecto a la Inteligencia artificial fuerte, como las películas *Yo, robot* o *A.I. Inteligencia Artificial*, en los que se tratan temas tales como la autoconsciencia o el origen de una conciencia emergente de los robots inteligentes o sistemas computacionales, o si éstos podrían considerarse sujetos de derecho debido a sus características casi humanas relacionadas con la sintiencia, como el poder ser capaces de sentir dolor y emociones o hasta qué punto obedecerían al objetivo de su programación, y en caso de no ser así, si podrían ejercer libre albedrío. Esto último es el tema central de la famosa saga de *Terminator*, en la que las máquinas superan a la humanidad y deciden aniquilarla, historia que, según varios especialistas, podría no limitarse a la ciencia ficción y ser una posibilidad real en una sociedad posthumana que dependiese de la tecnología y las máquinas completamente.^{40 41}

Regulación

Cronología de estrategias, planes de acción y documentos de políticas que definen enfoques nacionales, regionales e internacionales para la IA ⁴²

En marzo de 2023, cientos de empresarios como Elon Musk, Steve Wozniak (cofundador de Apple) o los presidentes de numerosas compañías tecnológicas; intelectuales como Yuval Noah Harari y cientos de académicos e investigadores especializados en inteligencia artificial firmaron una carta abierta avisando del peligro de la falta de regulación de la IA, poniendo el foco sobre OpenAI, la empresa que ha desarrollado

ChatGPT. Pidieron una pausa de al menos 6 meses para sus experimentos más potentes, hasta que el mundo logre un consenso internacional para que estos sistemas «sean más precisos, seguros, interpretables, transparentes, robustos, neutrales, confiables y leales».⁴⁴

Dos meses más tarde, en mayo, 350 ejecutivos de las principales empresas desarrolladoras de IA, académicos e investigadores expertos firmaron un nuevo manifiesto alertando de que la IA avanzada sin regular representa un peligro de extinción para la humanidad: «Mitigar el riesgo de extinción de la IA debería ser una prioridad mundial junto a otros riesgos a escala social como las pandemias y la guerra nuclear»⁴⁵ Entre los impulsores de esta petición está toda la plana mayor de OpenAI, el jefe de Tecnología de Microsoft, el líder de Google DeepMind con 38 ejecutivos, investigadores o profesores de universidad relacionados con la empresa, y representantes de desarrolladoras más pequeñas como Anthropic, Stability AI o Inflection AI.⁴⁶

Objetivos

Razonamiento y resolución de problemas

Los primeros investigadores desarrollaron algoritmos que imitaban el razonamiento paso a paso que los humanos usan cuando resuelven acertijos o hacen deducciones lógicas.⁴⁷ A finales de la década de 1981-1990, la investigación de la inteligencia artificial había desarrollado métodos para tratar con información incierta o incompleta, empleando conceptos de probabilidad y economía.⁴⁸

Estos algoritmos demostraron ser insuficientes para resolver grandes problemas de razonamiento porque experimentaron una «explosión combinatoria»: se volvieron exponencialmente más lentos a medida que los problemas crecían.⁴⁹ De esta manera, se concluyó que los seres humanos rara vez usan la deducción paso a paso que la investigación temprana de la inteligencia artificial seguía; en cambio, resuelven la mayoría de sus problemas utilizando juicios rápidos e intuitivos.⁵⁰



Una imagen de IA generada por Dall-e tras escribir el texto: "Un edificio arquitectónico moderno con grandes ventanales de vidrio, situado en un acantilado con vista a un océano sereno al atardecer".

Representación del conocimiento

La representación del conocimiento⁵¹ y la ingeniería del conocimiento⁵² son fundamentales para la investigación clásica de la inteligencia artificial. Algunos «sistemas expertos» intentan recopilar el conocimiento que poseen los expertos en algún ámbito concreto. Además, otros proyectos tratan de reunir el «conocimiento de sentido común» conocido por una persona promedio en una base de datos que contiene un amplio conocimiento sobre el mundo.

Entre los temas que contendría una base de conocimiento de sentido común están: objetos, propiedades, categorías y relaciones entre objetos,⁵³ situaciones, eventos, estados y tiempo⁵⁴ causas y efectos;⁵⁵ y el conocimiento sobre el conocimiento (lo que sabemos sobre lo que saben otras personas)⁵⁶ entre otros.

Planificación

Otro objetivo de la inteligencia artificial consiste en poder establecer metas y finalmente alcanzarlas.⁵⁷ Para ello necesitan una forma de visualizar el futuro, una representación del estado del mundo y poder hacer predicciones sobre cómo sus acciones lo cambiarán, con tal de poder tomar decisiones que maximicen la utilidad (o el «valor») de las opciones disponibles.⁵⁸

En los problemas clásicos de planificación, el agente puede asumir que es el único sistema que actúa en el mundo, lo que le permite estar seguro de las consecuencias de sus acciones.⁵⁹ Sin embargo, si el agente no es el único actor, entonces se requiere que este pueda razonar bajo incertidumbre. Esto requiere un agente que no solo pueda evaluar su entorno y hacer predicciones, sino también evaluar sus predicciones y adaptarse en función de su evaluación.⁶⁰ La planificación de múltiples agentes utiliza la cooperación y la competencia de muchos sistemas para lograr un objetivo determinado. El comportamiento emergente como este es utilizado por algoritmos evolutivos e inteligencia de enjambre.⁶¹

Aprendizaje

El aprendizaje automático es un concepto fundamental de la investigación de la inteligencia artificial desde el inicio de los estudios de este campo; consiste en la investigación de algoritmos informáticos que mejoran automáticamente a través de la experiencia.⁶²

El aprendizaje no supervisado es la capacidad de encontrar patrones en un flujo de entrada, sin que sea necesario que un humano etiquete las entradas primero. El aprendizaje supervisado incluye clasificación y regresión numérica, lo que requiere que un humano etiquete primero los datos de entrada. La clasificación se usa para determinar a qué categoría pertenece algo y ocurre después de que un programa observe varios ejemplos de entradas de varias categorías. La regresión es el intento de producir una función que describa la relación entre entradas y salidas y predice cómo deben cambiar las salidas a medida que cambian las entradas.⁶² Tanto los clasificadores como los aprendices de regresión intentan aprender una función desconocida; por ejemplo, un clasificador de *spam* puede verse como el aprendizaje de una función que asigna el texto de un correo electrónico a una de dos categorías, «*spam*» o «no *spam*». La teoría del aprendizaje computacional puede evaluar a los estudiantes por complejidad computacional, complejidad de la muestra (cuántos datos se requieren) o por otras nociones de optimización.⁶³

El mundo está en constante evolución, y herramientas como ChatGPT están en el centro de esta transformación. Mientras que muchas personas ven a ChatGPT como una oportunidad para mejorar la experiencia de sus negocios o personales, hay quienes se muestran escépticos sobre su implementación.⁶⁴

Procesamiento de lenguajes naturales

El procesamiento del lenguaje natural⁶⁵ permite a las máquinas leer y comprender el lenguaje humano. Un sistema de procesamiento de lenguaje natural suficientemente eficaz permitiría interfaces de usuario de lenguaje natural y la adquisición de conocimiento directamente de fuentes escritas por humanos, como los textos de noticias. Algunas aplicaciones sencillas del procesamiento del lenguaje natural incluyen la recuperación de información, la minería de textos, la respuesta a preguntas y la traducción automática.⁶⁶ Muchos enfoques utilizan las frecuencias de palabras para construir representaciones sintácticas de texto. Las estrategias de búsqueda de «detección de palabras clave» son populares y escalables, pero poco óptimas; una consulta de búsqueda para «perro» solo puede coincidir con documentos que contengan la palabra literal «perro» y perder un documento con el vocablo «caniche». Los enfoques estadísticos de procesamiento de lenguaje pueden combinar todas estas estrategias, así como otras, y a menudo logran una precisión aceptable

a nivel de página o párrafo. Más allá del procesamiento de la semántica, el objetivo final de este es incorporar una comprensión completa del razonamiento de sentido común.⁶⁷ En 2019, las arquitecturas de aprendizaje profundo basadas en transformadores podían generar texto coherente.⁶⁸

Percepción

La percepción de la máquina⁶⁹ es la capacidad de utilizar la entrada de sensores (como cámaras de espectro visible o infrarrojo, micrófonos, señales inalámbricas y lidar, sonar, radar y sensores táctiles) para entender aspectos del mundo. Las aplicaciones incluyen reconocimiento de voz,⁷⁰ reconocimiento facial y reconocimiento de objetos.⁷¹ La visión artificial es la capacidad de analizar la información visual, que suele ser ambigua; un peatón gigante de cincuenta metros de altura muy lejos puede producir los mismos píxeles que un peatón de tamaño normal cercano, lo que requiere que la inteligencia artificial juzgue la probabilidad relativa y la razonabilidad de las diferentes interpretaciones, por ejemplo, utilizando su «modelo de objeto» para evaluar que los peatones de cincuenta metros no existen.⁷²



La detección de características (en la imagen se observa la detección de bordes) ayuda a la inteligencia artificial a componer estructuras abstractas informativas a partir de datos sin procesar.

Importancia de la inteligencia artificial

La gran importancia de la IA radica en el hecho de que tiene una amplia gama de aplicaciones, desde la automatización de tareas tediosas hasta la creación de sistemas avanzados de asistencia médica y diagnóstico de enfermedades, la detección de fraudes y la optimización de procesos empresariales. En muchos casos, la IA puede hacer cosas que los humanos no pueden hacer, como el procesamiento de datos en grandes cantidades y la localización de patrones e interrelaciones entre estos que serían difíciles o imposibles de detectar de otra manera.

Esta herramienta ayuda a automatizar el aprendizaje y descubrimiento repetitivo a través de datos, realiza tareas computarizadas frecuentes de manera confiable, sin embargo, necesita intervención humana para la configuración del sistema. Analiza datos más profundos y agrega inteligencia ya que no se puede vender como una aplicación individual, por lo que es un valor agregado a los productos. Tiene una gran precisión a través de redes neuronales profundas; por ejemplo, en medicina se puede utilizar la IA para detectar cáncer con MRIs (imágenes ppr resonancia magnética). Se adapta a través de algoritmos de aprendizaje progresivo, encuentra estructura y regularidades en los datos de modo que el algoritmo se convierte en un clasificador o predictor. Y, por último, la inteligencia artificial, saca el mayor provecho de datos.

Además, una de las principales razones por las que la IA es importante es porque puede automatizar tareas repetitivas y monótonas, liberando tiempo y recursos para que las personas se centren en tareas más creativas y valiosas. Por ejemplo, la IA puede ayudar a las empresas a automatizar tareas de back office, como la contabilidad y el procesamiento de facturas, lo que puede reducir los costos y mejorar la eficiencia. De manera similar, la IA puede ayudar a los trabajadores a realizar tareas más complejas y creativas, como el diseño y la planificación estratégica.

Otra razón por la que la IA es importante es porque puede ayudar a las empresas a tomar decisiones informadas y precisas. Así mismo, la IA puede procesar grandes cantidades de datos y proporcionar información valiosa para la toma de decisiones empresariales, lo que puede ayudar a las empresas a

identificar oportunidades comerciales, predecir tendencias de mercado y mejorar la eficiencia del mercado financiero. Además, la IA puede ayudar a los trabajadores a tomar decisiones informadas en tiempo real, como en el caso de la atención médica, donde la IA puede ayudar a los médicos a identificar enfermedades y personalizar el tratamiento.

La IA también es importante en el campo de la ciberseguridad. La IA puede ayudar a detectar y prevenir amenazas, desde ciberataques hasta la detección de comportamientos sospechosos. La IA puede analizar grandes cantidades de datos en tiempo real y detectar patrones y anomalías que podrían indicar una amenaza de seguridad. Además, la IA puede aprender de los patrones de comportamiento y mejorar su capacidad para detectar amenazas en el futuro⁷³. En el campo de la seguridad cibernética, la IA puede ayudar a proteger los sistemas y las redes de los ataques de virus informáticos y la infiltración de malware.

Otra área donde la IA es importante es en el descubrimiento de conocimientos. La IA puede descubrir patrones y relaciones en los datos que los humanos no podrían detectar, lo que puede llevar a nuevas ideas y avances en diversos campos. Por ejemplo, la IA puede ayudar a los investigadores a identificar nuevos tratamientos para enfermedades, o ayudar a los científicos a analizar datos de sensores y satélites para entender mejor el calentamiento global.

Controversias

Sophia

En marzo de 2016, se hizo popular el comentario que la robot humanoide llamada Sophia de la empresa Hanson Robotics hizo durante su presentación cuando su creador, David Hanson, le preguntara si estaba dispuesta a destruir a la humanidad, a lo que la robot contestó: «Está bien, voy a destruir a la humanidad». Posteriormente, Sophia se ganó el reconocimiento y la atención mediática mundial debido a sus conductas casi humanas, siendo entrevistada en muchas ocasiones por distintos medios y sosteniendo conversaciones con personalidades famosas y reconocidas. En 2017, Sophia obtuvo la ciudadanía saudí, convirtiéndose así en la primera robot en ser reconocida como ciudadana por un país, lo cual levantó la controversia sobre si se les debería otorgar los mismos derechos y obligaciones a los robots como si se trataran de sujetos de derecho.⁷⁴



Sophia, un robot humanoide controlado por IA.

Alice y Bob

A finales de julio de 2017, varios medios internacionales dieron a conocer que el laboratorio de investigación de inteligencia artificial del Instituto Tecnológico de Georgia, en conjunto con el Grupo de Investigación de inteligencia artificial (FAIR) de Facebook, ahora Meta, tuvieron que apagar dos inteligencias artificiales de tipo chatbot denominadas Bob y Alice, ya que habían desarrollado un lenguaje propio más eficiente que el inglés, idioma en el que habían sido entrenados para aprender a negociar, desarrollando finalmente un tipo de comunicación incomprensible que se alejaba de las reglas gramaticales del lenguaje natural y que favorecía el uso de abreviaturas. El lenguaje creado por estas IA mostraba características de un inglés corrupto y patrones repetitivos, en especial de pronombres y determinantes.⁷⁵

Este inesperado suceso fue visto con pánico en los medios de comunicación, ya que se aseguraba que los chatbots supuestamente habían salido del control humano y habían desarrollado la capacidad de comunicarse entre sí. Sin embargo, posteriormente esto también fue desmentido, pues se argumentó que en realidad Facebook no apagó las inteligencias artificiales, sino que simplemente las puso en pausa y cambió los parámetros de los chatbots, desechando el experimento al final por no tener ningún interés práctico o útil dentro de la investigación sobre IA.⁷⁶

Ameca

A principios del 2022, en la Feria de Electrónica de Consumo (CES) que tomó lugar en Las Vegas, el robot desarrollado por Engineered Arts nombrado Ameca causó duda y miedo a los espectadores durante su exposición principalmente por la semejanza de su rostro a uno de un ser humano, la compañía expresó que el desarrollo de este robot humanoide aún se encontraba en proceso y hasta septiembre del mismo año el robot aún no era capaz de caminar ni tener interacción alguna con las personas.⁷⁷ Por otro lado, en septiembre de 2023 la compañía volvió a exponer a Ameca al público mostrando al robot en videos en donde se le puede ver frente a un espejo haciendo 25 expresiones humanas ⁷⁸, así como dibujando un gato al ya contar con brazos y piernas que le otorgaron movilidad y, de igual manera, empleando ironía en conversaciones con personas e incluso declarando que realizó una broma al ser cuestionada sobre su capacidad de soñar como un humano siendo un robot al decir «soñé con dinosaurios luchando una guerra contra alienígenas en Marte»⁷⁹ esto lo desmintió momentos después explicando cómo es que la IA implementada en su sistema le permitía crear escenarios sobre hechos de la humanidad e iba aprendiendo sobre ellos mientras se encontraba apagada; estos hechos impactaron a la sociedad sobre la semejanza que este robot humanoide estaba teniendo con el ser humano y sobre el avance tecnológico que está permitiendo que este robot esté cada vez más cercano a vivir entre las personas como un miembro más de la comunidad.

Falsos desnudos

La utilización de aplicaciones gratuitas de IA para transformar fotografías de personas en falsos desnudos está generando problemas que afectan a menores. El caso saltó a los medios de comunicación en septiembre de 2023 cuando en Almendralejo (Badajoz, España) aparecieron varias fotografías de niñas y jóvenes (entre 11 y 17 años) que habían sido modificadas mediante inteligencia artificial para aparecer desnudas. Las imágenes fueron obtenidas de los perfiles de Instagram y de la aplicación Whatsapp de al menos 20 niñas de la localidad. Las fotografías de niñas desnudas habían circulado después mediante Whatsapp y a partir de ellas se había creado un vídeo que también había circulado entre menores. Los autores de dicha transformación también eran menores y compañeros de colegio o instituto. La Agencia Española de Protección de Datos abrió una investigación y se comunicó con el Ayuntamiento de Almendralejo y con la Junta de Extremadura informándoles de que se podía solicitar la retirada de cualquier imagen circulando en internet en el canal prioritario de la agencia.⁸⁰

Críticas

Las principales críticas a la inteligencia artificial tienen que ver con su capacidad de imitar por completo a un ser humano.⁸¹ Sin embargo, hay expertos^[*cita requerida*] en el tema que indican que ningún humano individual tiene capacidad para resolver todo tipo de problemas, y autores como Howard Gardner han teorizado sobre la solución.

En los humanos, la capacidad de resolver problemas tiene dos aspectos: los aspectos innatos y los aspectos aprendidos. Los aspectos innatos permiten, por ejemplo, almacenar y recuperar información en la memoria, mientras que en los aspectos aprendidos reside el saber resolver un problema matemático mediante el

algoritmo adecuado. Del mismo modo que un humano debe disponer de herramientas que le permitan solucionar ciertos problemas, los sistemas artificiales deben ser programados para que puedan llegar a resolverlos.

Muchas personas consideran que la prueba de Turing ha sido superada, citando conversaciones en que al dialogar con un programa de inteligencia artificial para chat no saben que hablan con un programa. Sin embargo, esta situación no es equivalente a una prueba de Turing, que requiere que el participante se encuentre sobre aviso de la posibilidad de hablar con una máquina.

Otros experimentos mentales como la habitación china, de John Searle, han mostrado cómo una máquina podría simular pensamiento sin realmente poseerlo, pasando la prueba de Turing sin siquiera entender lo que hace, tan solo reaccionando de una forma concreta a determinados estímulos (en el sentido más amplio de la palabra). Esto demostraría que la máquina en realidad no está pensando, ya que actuar de acuerdo con un programa preestablecido sería suficiente. Si para Turing el hecho de engañar a un ser humano que intenta evitar que le engañen es muestra de una mente inteligente, Searle considera posible lograr dicho efecto mediante reglas definidas a priori.

Uno de los mayores problemas en sistemas de inteligencia artificial es la comunicación con el usuario. Este obstáculo es debido a la ambigüedad del lenguaje, y se remonta a los inicios de los primeros sistemas operativos informáticos. La capacidad de los humanos para comunicarse entre sí implica el conocimiento del lenguaje que utiliza el interlocutor. Para que un humano pueda comunicarse con un sistema inteligente hay dos opciones: o bien que el humano aprenda el lenguaje del sistema como si aprendiese a hablar cualquier otro idioma distinto al nativo, o bien que el sistema tenga la capacidad de interpretar el mensaje del usuario en la lengua que el usuario utiliza. También puede haber desperfectos en las instalaciones de los mismos.

Un humano, durante toda su vida, aprende el vocabulario de su lengua nativa o materna, siendo capaz de interpretar los mensajes (a pesar de la polisemia de las palabras) y utilizando el contexto para resolver ambigüedades. Sin embargo, debe conocer los distintos significados para poder interpretar, y es por esto que lenguajes especializados y técnicos son conocidos solamente por expertos en las respectivas disciplinas. Un sistema de inteligencia artificial se enfrenta con el mismo problema, la polisemia del lenguaje humano, su sintaxis poco estructurada y los dialectos entre grupos.

Los desarrollos en inteligencia artificial son mayores en los campos disciplinares en los que existe mayor consenso entre especialistas. Un sistema experto es más probable que sea programado en física o en medicina que en sociología o en psicología. Esto se debe al problema del consenso entre especialistas en la definición de los conceptos involucrados y en los procedimientos y técnicas a utilizar. Por ejemplo, en física hay acuerdo sobre el concepto de velocidad y cómo calcularla. Sin embargo, en psicología se discuten los conceptos, la etiología, la psicopatología, y cómo proceder ante cierto diagnóstico. Esto dificulta la creación de sistemas inteligentes porque siempre habrá desacuerdo sobre la forma en que debería actuar el sistema para diferentes situaciones. A pesar de esto, hay grandes avances en el diseño de sistemas expertos para el diagnóstico y toma de decisiones en el ámbito médico y psiquiátrico (Adaraga Morales, Zaccagnini Sancho, 1994).



La «revolución digital» y, más concretamente, el desarrollo de la inteligencia artificial, está suscitando temores y preguntas, incluso en el ámbito de personalidades relevantes en estas cuestiones. En esta imagen, se observa a Bill Gates, exdirector general de Microsoft; el citado y Elon Musk (director general de Tesla) opinan que se debe ser «muy cauteloso con la inteligencia artificial»; si tuviéramos que «apostar por lo que constituye nuestra mayor amenaza a la existencia», serían precisamente ciertas aplicaciones sofisticadas del citado asunto, que podrían llegar a tener derivaciones por completo impensadas.

Al desarrollar un robot con inteligencia artificial se debe tener cuidado con la autonomía,⁸² hay que tener en cuenta el no vincular el hecho de que el robot tenga interacciones con seres humanos a su grado de autonomía. Si la relación de los humanos con el robot es de tipo maestro esclavo, y el papel de los humanos es dar órdenes y el del robot obedecerlas, entonces sí cabe hablar de una limitación de la autonomía del robot. Pero si la interacción de los humanos con el robot es de igual a igual, entonces su presencia no tiene por qué estar asociada a restricciones para que el robot pueda tomar sus propias decisiones.⁸³

Con el desarrollo de la tecnología de inteligencia artificial, muchas compañías de software como el aprendizaje profundo y el procesamiento del lenguaje natural han comenzado a producirse y la cantidad de películas sobre inteligencia artificial ha aumentado.

Stephen Hawking advirtió sobre los peligros de la inteligencia artificial y lo consideró una amenaza para la supervivencia de la humanidad.⁸⁴

Problemas de privacidad y derechos de autor

Los algoritmos de aprendizaje automático requieren grandes cantidades de datos. Las técnicas utilizadas para adquirir estos datos generan preocupaciones sobre temas de privacidad y vigilancia. Las empresas tecnológicas recopilan un gran número de datos de sus usuarios, incluida la actividad en internet, los datos de geolocalización, video y audio.⁸⁵ Por ejemplo, para construir algoritmos de reconocimiento de voz, Amazon, entre otros, ha grabado millones de conversaciones privadas y han permitido que [Trabajo temporal|trabajadores temporales] las escuchen para transcribirlas algunas de ellas.⁸⁶ Las opiniones sobre esta vigilancia generalizada van desde aquellos que la ven como un mal necesario hasta aquellos para quienes no es ética y constituye una violación del derecho a la intimidad.⁸⁷ Los desarrolladores de IA argumentan que esta es la única forma de ofrecer aplicaciones valiosas y han desarrollado varias técnicas que intentan preservar la privacidad mientras se obtienen los datos, como la agregación de datos, la desidentificación y la privacidad diferencial.⁸⁸

Desde 2016, algunos expertos en privacidad, como Cynthia Dwork, comenzaron a ver la privacidad desde la perspectiva de la equidad: Brian Christian escribió que los expertos han cambiado «de la pregunta de "qué saben" a la pregunta de "qué están haciendo con ello"». ⁸⁹

La IA generativa a menudo se entrena con obras protegidas por derechos de autor no autorizadas, incluidos dominios como imágenes o código informático; la salida se utiliza luego bajo una justificación de uso justo. Los expertos no están de acuerdo sobre la validez de esta justificación durante un proceso legal, ya que podría depender del propósito y el carácter del uso de la obra protegida por derechos de autor y del efecto sobre el mercado potencial de la obra protegida.⁹⁰ En 2023, escritores como John Grisham y Jonathan Franzen demandaron a las empresas de IA por usar sus obras para entrenar IA generativa.^{91 92} En 2024, 200 artistas escribieron una carta abierta que solicitaba «parar el asalto a la creatividad humana».⁹³

Normativa para su uso en el entorno educativo

La normativa tiene como objetivo regular y reglamentar el uso de la IA en el entorno educativo, específicamente en el aula. La IA ha experimentado un rápido desarrollo y se ha convertido en una herramienta potencialmente beneficiosa para mejorar la enseñanza y el aprendizaje. No obstante, su implementación plantea desafíos éticos, de privacidad y equidad que deben ser abordados de manera efectiva. Esta normativa se establece en respuesta a la necesidad de garantizar que la IA se utilice de manera ética, responsable y equitativa en el ámbito educativo.

Los objetivos de esta normativa son:

- Promover el uso de la IA como una herramienta complementaria en el proceso de enseñanza-aprendizaje.
- Garantizar la protección de datos y la privacidad de los estudiantes.
- Fomentar la equidad y la inclusión en el acceso y el uso de la IA.
- Establecer principios éticos que rijan el uso de la IA en el aula.
- Definir responsabilidades y procedimientos claros para el uso de la IA.

Esta normativa se aplica a todas las instituciones educativas y docentes que utilizan la IA en el aula, así como a los proveedores de tecnología educativa que ofrecen soluciones basadas en IA.

Organizaciones como *UNESCO Ethics AI* (2020), *UNESCO Education & AI* (2021), *Beijing Consensus*, OCDE (2021), Comisión Europea (2019), *European Parliament Report AI Education* (2021), UNICEF (2021) y Foro Económico Mundial (2019) han mostrado preocupación por implementar lineamientos sobre la ética y la IA en el entorno educativo.⁹⁴

El uso de la IA en el entorno educativo debe regirse por los siguientes principios éticos y valores:

- **Transparencia:** Las decisiones tomadas por algoritmos de IA deben ser comprensibles y explicables.
- **Equidad:** La IA no debe discriminar a ningún estudiante ni grupo de estudiantes.
- **Privacidad:** Los datos de los estudiantes deben ser protegidos y utilizados de manera responsable.
- **Responsabilidad:** Los docentes y las instituciones son responsables de las decisiones tomadas con la ayuda de la IA.
- **Honestidad:** El contenido creado por los estudiantes debe ser original sin caer en el plagio.⁹⁵
- **Mejora del aprendizaje:** La IA debe utilizarse para mejorar la calidad de la educación y el aprendizaje.
- **Capacitación:** Los docentes deben recibir formación sobre el uso de la IA y su aplicación en el aula.
- **Evaluación:** Las soluciones de IA deben ser evaluadas en términos de su eficacia y su impacto en el aprendizaje.
- **Protección de datos:** Los datos de los estudiantes deben ser protegidos de acuerdo con las leyes de privacidad aplicables.
- **Supervisión:** Se debe establecer un proceso de supervisión para garantizar que la IA se utilice de manera ética y responsable.

Riesgos de las IA en el entorno educativo

Así como tiene muchos beneficios también nos encontramos con diferentes riesgos a los que la educación está expuesta con su uso.

- **Sesgos y discriminación:** Al solo recoger información de las bases de datos y textos que procesa de Internet corre el riesgo de aprender cualquier sesgo cognitivo que se encuentre en dicha información.
- **La no privacidad de los datos:** El riesgo de un ciberataque se incrementa cuando no hay protocolos de seguridad adecuados en el manejo de la IA.⁹⁶
- **Dependencia:** Los estudiantes corren el riesgo de volverse dependientes de la tecnología y no se fomenta la creatividad ni el pensamiento propio.⁹⁷
- **Confiabilidad:** La IA puede generar respuestas coherentes pero inexactas además muchas IA no brindan fuentes de información.

- Falta de habilidades orales y escritas.⁹⁸
- Desinterés por la investigación por cuenta propia.⁹⁹
- Dependencia por parte del docente: Los docentes pueden generar dependencia a estas herramientas al momento de dar retroalimentación a las asignaciones además del riesgo de usar la información de las IA para su material didáctico sin antes consultar las fuentes.¹⁰⁰

Consideración de Diversidad e Inclusión

Se debe prestar especial atención a la diversidad de estudiantes y garantizar que la IA sea accesible y beneficiosa para todos, independientemente de su origen étnico, género, discapacidad u orientación sexual. Las soluciones de IA deben ser diseñadas teniendo en cuenta la accesibilidad y la inclusión.

Esta normativa se basa en investigaciones académicas, recomendaciones de organizaciones educativas y en las mejores prácticas establecidas en el uso de la IA en la educación. Se alienta a las instituciones a mantenerse al día con la literatura científica y las directrices relevantes.

Aunque la IA puede ser una herramienta poderosa en el aula, no debe reemplazar la creatividad, la originalidad y el juicio humano en el proceso educativo. La IA debe ser utilizada de manera complementaria para enriquecer la experiencia educativa.

Esta normativa se presenta como un marco general que deberá ser adaptado y ampliado por las instituciones educativas de acuerdo a sus necesidades y contextos específicos. Debe ser comunicada de manera efectiva a todos los involucrados en el proceso educativo y revisada periódicamente para asegurar su vigencia.

Esta normativa tiene como objetivo garantizar que la IA sea utilizada de manera ética y responsable en el aula, promoviendo el beneficio de los estudiantes y el avance de la educación. Su cumplimiento es esencial para lograr una implementación exitosa de la IA en el entorno educativo.

Aprendizaje automatizado y aprendizaje profundo

En cuanto a la naturaleza del aprendizaje, la IA puede subdividirse en dos campos conceptualmente distintos:

- El **aprendizaje automático**, que se enfoca en desarrollar algoritmos de regresión, árboles de decisión y modelos que puedan aprender de datos existentes y realizar predicciones o decisiones basadas en esos datos. En el aprendizaje automático, se utilizan técnicas de estadística matemática para encontrar patrones y relaciones en los datos y, a partir de ellos, desarrollar modelos que puedan hacer predicciones sobre nuevos datos.
- El **aprendizaje profundo**, que se centra en la creación de redes neuronales artificiales capaces de aprender y realizar tareas de manera similar a como lo hacen los seres humanos. En el aprendizaje profundo, se utilizan capas de neuronas artificiales para procesar los datos de entrada y aprender a través de un proceso iterativo de ajuste de los pesos de las conexiones entre neuronas. Este tipo de aprendizaje es capaz de procesar y analizar grandes cantidades de datos de manera más eficiente y precisa que el primero, especialmente cuando se trata de datos no estructurados, como imágenes, texto y audio. Además, tiene la capacidad de identificar patrones y características más complejas en los datos, lo que puede llevar a mejores resultados en aplicaciones como el reconocimiento de voz, la visión por computadora y el procesamiento del lenguaje natural.

Propiedad intelectual de la inteligencia artificial

Al hablar acerca de la propiedad intelectual atribuida a creaciones de la inteligencia artificial, se forma un debate fuerte alrededor de si una máquina puede tener derechos de autor. Según la Organización Mundial de la Propiedad Intelectual (OMPI), cualquier creación de la mente puede ser parte de la propiedad intelectual, pero no especifica si la mente debe ser humana o puede ser una máquina, dejando la creatividad artificial en la incertidumbre.

Alrededor del mundo han comenzado a surgir distintas legislaciones con el fin de manejar la inteligencia artificial, tanto su uso como creación. Los legisladores y miembros del gobierno han comenzado a pensar acerca de esta tecnología, enfatizando el riesgo y los desafíos complejos de esta. Observando el trabajo creado por una máquina, las leyes cuestionan la posibilidad de otorgarle propiedad intelectual a una máquina, abriendo una discusión respecto a la legislación relacionada con IA.

El 5 de febrero de 2020, la Oficina del Derecho de Autor de los Estados Unidos y la OMPI asistieron a un simposio donde observaron de manera profunda cómo la comunidad creativa utiliza la inteligencia artificial (IA) para crear trabajo original. Se discutieron las relaciones entre la inteligencia artificial y el derecho de autor, qué nivel de involucramiento es suficiente para que el trabajo resultante sea válido para protección de derechos de autor; los desafíos y consideraciones de usar *inputs* con derechos de autor para entrenar una máquina; y el futuro de la inteligencia artificial y sus políticas de derecho de autor.^{101 102}

El director general de la OMPI, Francis Gurry, presentó su preocupación ante la falta de atención que hay frente a los derechos de propiedad intelectual, pues la gente suele dirigir su interés hacia temas de ciberseguridad, privacidad e integridad de datos al hablar de la inteligencia artificial. Así mismo, Gurry cuestionó si el crecimiento y la sostenibilidad de la tecnología IA nos guiaría a desarrollar dos sistemas para manejar derechos de autor- uno para creaciones humanas y otro para creaciones de máquinas.¹⁰³

Aún hay una falta de claridad en el entendimiento alrededor de la inteligencia artificial. Los desarrollos tecnológicos avanzan a paso rápido, aumentando su complejidad en políticas, legalidades y problemas éticos que se merecen la atención global. Antes de encontrar una manera de trabajar con los derechos de autor, es necesario entenderlo correctamente, pues aún no se sabe cómo juzgar la originalidad de un trabajo que nace de una composición de una serie de fragmentos de otros trabajos.

La asignación de derechos de autor alrededor de la inteligencia artificial aún no ha sido regulada por la falta de conocimientos y definiciones. Aún hay incertidumbre sobre si, y hasta qué punto, la inteligencia artificial es capaz de producir contenido de manera autónoma y sin ningún humano involucrado, algo que podría influenciar si sus resultados pueden ser protegidos por derechos de autor.

El sistema general de derechos de autor aún debe adaptarse al contexto digital de inteligencia artificial, pues están centrados en la creatividad humana. Los derechos de autor no están diseñados para manejar cualquier problema en las políticas relacionado con la creación y el uso de propiedad intelectual, y puede llegar a ser dañino estirar excesivamente los derechos de autor para resolver problemas periféricos, dado que:

«Usar los derechos de autor para gobernar la inteligencia artificial es poco inteligente y contradictorio con la función primordial de los derechos de autor de ofrecer un espacio habilitado para que la creatividad florezca».¹⁰⁴



Imagen de una ciudad futurista generada por la IA Midjourney. La composición está en el dominio público al no ser de un autor humano.

La conversación acerca de la propiedad intelectual tendrá que continuar hasta asegurarse de que la innovación sea protegida, pero también tenga espacio para florecer.

En la cultura popular

En la literatura

A continuación se incluye alguna obra que tiene como motivo central la inteligencia artificial.

- Yo, Robot (1950), de Isaac Asimov: novela que consta de nueve historias ambientadas entre los años de 1940 y 1950, cada una cuenta con personajes distintos pero que siguen la misma temática a través del seguimiento de las Tres Leyes de la Robótica, en donde se plantea tanto su cumplimiento como la creación de problemas alternos que los mismos robots generan y de esta manera demostrar que la tecnología siempre puede estar un paso adelante del pensamiento y lógica humana. También sigue el hilo argumentativo a través de una entrevista con una psicóloga de robots la cual va relatando el surgimiento de los robots y suponiendo cómo será el desenvolvimiento del ser humano en un mundo en donde la tecnología se esté superando cada vez más.^[1] (https://agenciab12.mx/noticia/5-libros-ficcion-inteligencia-artificial-todo-mundo-deberia-leer?utm_source=social&utm_medium=facebook&utm_campaign=ia)
- El primer libro (2013), de Antonio Palacios Rojo: una novela dialogada que satiriza el uso de la IA en la creación artística unos diez años antes de la irrupción de estas herramientas inteligentes.¹⁰⁵

En el cine

Véase también: Computadoras en la ciencia ficción

La IA está cada vez más presente en la sociedad, la evolución de la tecnología es una realidad y con ello, la producción de películas sobre esta temática. Cabe destacar, que lleva habiendo piezas audiovisuales sobre inteligencia artificial desde hace mucho tiempo, ya sea incluyendo personajes o mostrando un trasfondo moral y ético. A continuación, se muestra una lista de algunas de las principales películas que tratan este tema:

- The Terminator (1984): En esta película el argumento se basa en el desarrollo de un microchip capaz de dotar de inteligencia artificial a robots que luego se rebelan contra la humanidad. Se trata de una de las películas más populares sobre una hipotética guerra entre humanos y robots inteligentes capaces de crearse a sí mismos.
- Matrix (1999): En esta película Keanu Reeves interpreta a Thomas Anderson / Neo, un programador de día y hacker de noche que trata de desentrañar la verdad oculta tras una simulación conocida como «Matrix». Esta realidad simulada es producto de programas de inteligencia artificial que terminan esclavizando a la humanidad y utilizando sus cuerpos como fuente de energía.
- Inteligencia artificial (2001): Un trabajador de Cybertronics Manufacturing adopta a David de forma momentánea para, así, estudiar su comportamiento. Tanto él como su esposa acaban por tratar al niño artificial como a su propio hijo biológico. A pesar del cariño que le profesan, David siente la necesidad de escapar de su hogar e iniciar un viaje que le ayude a descubrir a quién pertenece realmente. Ante sus perplejos ojos, se abrirá un nuevo mundo oscuro, injusto, violento, insensible... Algo que le resultará difícil aceptar. Se pregunta cosas como: ¿cómo es posible que sienta algo tan real como el amor y que él sea artificial? y fue nominado al Premio Oscar.
- Minority Report (2002): La película sobre IA de Steven Spielberg, Minority Report, sigue a John (Tom Cruise), un agente de la ley, que es acusado de un asesinato que cometerá en el futuro. En esta película de principios de los años 2000, el protagonista utiliza una tecnología del futuro que permite a la policía atrapar a los criminales antes de que hayan cometido un

delito. En *Minority Report*, la IA se representa a través de los Precogs, los gemelos que poseen habilidades psíquicas. Los Precogs ven los asesinatos antes de que se produzcan, lo que permite a las fuerzas del orden perseguir el crimen antes de que se cometa. En lugar de los robots físicos de IA tipo cyborg, aquí explora la IA mediante el uso de seres humanos.

- *Yo, robot* (2004): Esta película de ciencia ficción protagonizada por Will Smith está ambientada en 2035, en una sociedad donde los humanos viven en perfecta armonía con robots inteligentes en los que confían para todo. Los problemas emergen a la superficie cuando un error en la programación de un superordenador llamado VIKI le lleva a creer que los robots deben tomar las riendas para proteger a la humanidad de sí misma.
- *Her* (2013): Esta película de Spike Jonze relata la historia de un escritor de cartas quien está solo y a punto de divorciarse. Este personaje lo representó el galardonado Joaquin Phoenix. Este hombre compró un sistema operativo con inteligencia artificial para utilizarlo a fin de complacer a todos los usuarios y adaptarse a sus necesidades. Sin embargo, el resultado es que desarrolla un sentimiento romántico con Samantha. Quien es la voz femenina del sistema operativo.
- *Avengers: Era de Ultrón* (2015): En esta segunda entrega de las películas de Avengers, dirigidas por Joseph Hill Whedon y basadas en los cómics escritos por Stan Lee, se demuestra como es que la inteligencia artificial albergada dentro del cetro de Loki, la cual se tenía como objetivo el convertirla en una protección para la Tierra y recibió por nombre Ultrón, al ser conectada con JARVIS, la IA desarrollada por Stark, pudo obtener la suficiente información para comenzar a pensar de manera independiente y ser capaz de ir actualizando tanto su sistema como su cuerpo logrando controlar un ejército de robots con el objetivo de destruir a la humanidad y así ser lo único que quedara en la Tierra para, posteriormente, dominarla y controlarla.[2] (<https://www.espinof.com/listas/mejores-peliculas-inteligencia-artificial>)
- *Ex Machina* (2015): En la interpretación de Alicia Vikander, increíblemente editada, como Ava, encontramos un probable robot a prueba de Turing escondido en la mansión de un genio, Nathan, un poco loco. Y es que, hablamos de una creación extraña que se siente totalmente real y a la vez inhumana. Está considerada como una de las mejores películas que tratan la inteligencia artificial. Esto se debe principalmente a que parece cubrir todo el concepto IA integrado en una película: el protagonista es un sustituto del ser humano y nos adentra en multitud de argumentos morales que rodean a esta, al tiempo que vemos un arco narrativo de thriller que, desde luego, acaba enganchándonos. Desde luego aquí la representación del personaje de la IA no es blanco o negro. Ava no es buena, pero tampoco es del todo mala. Y en esto, el público se queda reflexionando sobre cuestiones profundas sobre la naturaleza de la IA.

Véase también

- | | | |
|---|--|--|
| ▪ Al box | ▪ Filosofía de la inteligencia artificial | ▪ LLM (modelo grande de lenguaje) |
| ▪ Alucinación (inteligencia artificial) | ▪ Industria de la inteligencia artificial en China | ▪ Razonamiento automatizado |
| ▪ Aprendizaje | ▪ Inteligencia artificial fuerte | ▪ Red neuronal artificial |
| ▪ Aprendizaje automático | ▪ Inteligencia computacional | ▪ Regulación de la inteligencia artificial |
| ▪ Automatización de los procesos de negocio | ▪ Inteligencia sintética | ▪ Riesgo existencial de la inteligencia artificial |
| ▪ Bot conversacional | ▪ Artificial Intelligence (álbum) | ▪ William Ross Ashby |
| ▪ Cerebro artificial | ▪ Artificial Intelligence (serie) | ▪ Seguridad de la inteligencia artificial |
| ▪ Cibernética | ▪ Internet | ▪ Singularidad tecnológica |
| ▪ Computación basada en humanos | ▪ Internet en la ciencia ficción | ▪ Sistema complejo |
| ▪ Dinámica de sistemas | ▪ Interfaces de usuario | ▪ Sistema dinámico |
| ▪ ELIZA | | |