

La Antropomorfización de las Máquinas: ¿Existe la Inteligencia Artificial?¹

Carlos José Pereira²

Pereira-carlos-jose@outlook.com

ORCID: <https://orcid.org/0009-0002-7915-5779>

Artículo presentado para su publicación el día: 14 de mayo de 2025

Artículo aceptado para su publicación el día: 10 de julio de 2025

ARK/CAICYT: <https://id.caicyt.gov.ar/ark:/s25915266/s0uvj9ucr>

RESUMEN

Este artículo explora la antropomorfización de la Inteligencia Artificial (IA), destacando cómo las analogías entre máquinas y humanos, aunque útiles, pueden ser engañosas. Términos como “aprendizaje” y “memoria” aplicados a la IA oscurecen las diferencias fundamentales entre el aprendizaje humano y el aprendizaje automático. La IA, a pesar de sus avances, todavía enfrenta desafíos para replicar la complejidad de la cognición humana, especialmente en áreas como la toma de decisiones y la comprensión de los matices. La dependencia de la IA de grandes volúmenes de datos y potencia computacional contrasta con

¹ Artículo aceptado para su publicación el día 10 DE JULIO DE 2025.

² Doctorando en Ciencias Empresariales y Sociales por la Universidad de Ciencias Empresariales y Sociales (UCES). Lic. en Ciencias Contables por la Fundación Universitaria de la Región de Blumenau (FURB).

la capacidad humana de aprender de unos pocos sucesos. Además, la ausencia de emociones en las máquinas plantea interrogantes sobre su capacidad para alcanzar la verdadera inteligencia. El exceso de confianza en los sistemas automatizados también es problemático, una situación evidenciada por estudios que muestran cómo la presencia de robots, compartiendo tareas, puede conducir a errores humanos. Si bien la IA es una herramienta poderosa, es crucial reconocer sus limitaciones y evitar atribuirle indiscriminadamente características humanas. Comprender las distinciones entre humanos y máquinas es esencial para el desarrollo y el uso ético de la IA.

RESUMO

Este artigo explora a antropomorfização da Inteligência Artificial (IA), destacando como as analogias entre máquinas e humanos, embora úteis, podem ser enganosas. Termos como “aprendizado” e “memória”, aplicados à IA, obscurecem as diferenças fundamentais entre o aprendizado humano e o aprendizado de máquina. A IA, apesar de seus avanços, ainda enfrenta desafios para replicar a complexidade da cognição humana, especialmente em áreas como a tomada de decisões e a compreensão de nuances. A dependência da IA de grandes volumes de dados e de poder computacional contrasta com a capacidade humana de aprender a partir de poucos eventos. Além disso, a ausência de emoções nas máquinas levanta questionamentos sobre sua capacidade de alcançar uma inteligência genuína. O excesso de confiança em sistemas automatizados também é problemático, situação evidenciada por estudos que mostram como a presença de robôs, ao compartilhar tarefas, pode levar a erros humanos. Embora a IA seja uma ferramenta poderosa, é crucial reconhecer suas limitações e

evitar atribuir-lhe indiscriminadamente características humanas. Compreender as distinções entre humanos e máquinas é essencial para o desenvolvimento e o uso ético da IA.

ABSTRACT

This article explores the anthropomorphization of Artificial Intelligence (AI), highlighting how analogies between machines and humans, while useful, can be misleading. Terms such as “learning” and “memory” applied to AI obscure the fundamental differences between human learning and machine learning. Despite its advances, AI still faces challenges in replicating the complexity of human cognition, especially in areas such as decision-making and understanding nuances. AI’s dependence on large volumes of data and computational power contrasts with the human ability to learn from just a few events. Moreover, the absence of emotions in machines raises questions about their capacity to achieve true intelligence. Overconfidence in automated systems is also problematic, a situation evidenced by studies showing how the presence of robots sharing tasks can lead to human errors. While AI is a powerful tool, it is crucial to recognize its limitations and avoid indiscriminately attributing human characteristics to it. Understanding the distinctions between humans and machines is essential for the ethical development and use of AI.

Introducción

Históricamente, los dispositivos tecnológicos buscaban superar las limitaciones físicas de los seres humanos sin replicar sus capacidades mentales. No obstante, en las últimas décadas se ha iniciado un proceso de intensa investigación y desarrollo centrado en la creación de sistemas informáticos que

imitan la estructura y la función del cerebro. Así, la tecnología, según Sadin, comenzó a adoptar un fuerte enfoque *antropomórfico* (Sadin, 2020, pp. 17–20).

Es a partir de este enfoque que deriva el significado morfológico del término antropomorfizar, es decir, atribuir a algo no humano formas o características humanas, tal como lo definen Houaiss y Villar (2001, p. 240). Con respecto al término máquina, se utiliza para efectos de este trabajo y según la cibernética, en referencia a fórmulas matemáticas, cálculos y reglas de transformación, y no se limita necesariamente a dispositivos electrónicos o mecánicos (Luhmann, 2011, pp. 108–109).

En este campo conceptual, un algoritmo es una máquina y la ejecución de un programa consiste en proporcionar un acoplamiento estructural entre dos sistemas, hardware y software, cuyos elementos son de naturaleza maquina. Tavares-Pereira afirma que un algoritmo es una máquina virtual que depende de una máquina física para manifestarse y que quien realmente aprende es el algoritmo y no la máquina física, en el contexto de la inteligencia artificial (Tavares-Pereira, 2021, p. 70). A mediados del siglo pasado, los pioneros de la inteligencia artificial se fijaron una misión con un objetivo ambicioso y claro: replicar la inteligencia humana en una máquina. Esta combinación de claridad de propósito y complejidad de la tarea atrajo a algunas de las mentes más brillantes en el entonces incipiente campo de la informática. Desde su aparición, la IA ha experimentado varios ciclos de crecimiento y declive. A las fases de gran optimismo les siguieron fases de estancamiento señaladas por Lee como “inviernos de IA” (Lee, 2019, p. 19).

El objetivo de la IA se refuerza con la afirmación de Domingos (2017, p. 30) de que el propósito es enseñar a los ordenadores a hacer lo que los humanos hacen bien, es decir, principalmente a aprender. Si no aprende, una computadora no puede comportarse como un ser humano, lo que

conduce a la antropomorfización. El aprendizaje automático es un subcampo de la IA en rápido crecimiento que ha eclipsado el objetivo original de la disciplina (Domingos, 2017, pp. 30-31). **Analogías Útiles, Pero También Engañosas**

El proceso de antropomorfización de las máquinas conduce a analogías que pueden ser útiles y también presentarse como funcionalmente engañosas. Esto ocurre cuando uno se pregunta ¿qué significa realmente *imitar al ser humano* en sus capacidades, comportamientos y características? En una dimensión acelerada del tiempo no lineal, términos como *redes neuronales, memoria, pensamientos, alucinaciones* e incluso *inteligencia, que son expresiones atribuidas al ser humano*, no pueden migrar al área de la tecnología de la información, como si existiera una correspondencia operacional simétrica entre la forma en que aprenden las máquinas y la forma en que aprenden los sistemas psíquicos.

De acuerdo con Maturana y Varela (2001), las palabras, al designar objetos o situaciones en el mundo, no reflejan el funcionamiento del sistema nervioso, ya que este no opera como una representación del mundo (p. 230).

Kenneth Burke, como se cita en Luhmann (2011), afirma que, en la búsqueda de un lenguaje que refleje fielmente la realidad, la humanidad se enfrenta a un reto: la necesidad de seleccionar y, en consecuencia, distorsionar esta misma realidad. El vocabulario, como herramienta de representación, debe encontrar un equilibrio entre reflejar el mundo y, al mismo tiempo, reducirlo a términos comprensibles. Esta selección implica inevitablemente una deformación: se vuelve problemática cuando el lenguaje o los cálculos utilizados no se ajustan al objeto de estudio (pp. 179–180).

Se debe tener cuidado con los términos o palabras ampliamente utilizados en un dominio determinado, con significados definidos y establecidos. Estos se transponen, principalmente por

analogía, a otras áreas del conocimiento. La distinción entre los dos tipos de aprendizaje, el *machine learning* y el aprendizaje humano, tal como afirma Tavares-Pereira (2021), es evidente e implícita en el propio término *machine learning*, pues al destacar solo “aprendizaje” se oscurece la calificación de “máquina”, que es claramente limitante. Esto da lugar a una “falsa equivalencia”, pero a menudo se ignora (Tavares-Pereira, 2021, pp. 110–112).

1. Máquinas que Sustituyen a los Seres Humanos, Pero con Métodos y Técnicas Diferentes

A pesar de que las técnicas y algunos métodos utilizados para el desarrollo de la inteligencia artificial son nombrados con términos relacionados con los sistemas psíquicos, principalmente humanos y permiten que las computadoras los reemplacen en diversas actividades - incluidas las actividades dependientes del cerebro-, no se puede decir que sus operaciones internas sean iguales a la forma en que operan los sistemas psíquicos, que es mediante el pensamiento, tampoco son iguales a los sistemas sociales que operan a través de la comunicación y a los organismos que operan a través de la vida.

1.1. Cerebro y mente

La palabra Cerebro, en una concepción antropocéntrica de la palabra, se usa actualmente en dos sentidos, a saber, la parte física, representada por la forma de un kilogramo y trescientos gramos de materia gris, y como Mente. Y a menudo indistintamente, como una unidad, donde la mente y la parte física se confunden.

Minsky (*The Society of Mind*, 1989, p. 288) se pregunta por qué una mente parece tan diferente de cualquier otro tipo de cosa y afirma que las mentes no son cosas, al menos no comparten ninguna de las propiedades habituales de las cosas, como colores, tamaños, formas o pesos. Las mentes están más

allá del alcance de los sentidos del oído, la vista, el olfato y el gusto. Aunque no son cosas en absoluto, ciertamente poseen vínculos vitales con los cerebros.

Aunque el sistema de la ciencia ha evolucionado mucho sobre el funcionamiento del cerebro humano e incluso no humano (de algunos animales), los científicos saben mucho sobre el cerebro humano, pero aún no ha sido posible comprender su funcionamiento en su conjunto. Michio Kaku (*Physics of the Future*, 2014, p. 18) corrobora esta observación, afirmando que cuando observamos otros órganos del cuerpo, como músculos, huesos y pulmones, parece haber una relación obvia y razones que podemos ver de inmediato; sin embargo, la estructura del cerebro parece un poco caótica, como si las partes se golpearan entre sí. De hecho, un mapa del funcionamiento del cerebro a menudo ha sido denominado “cartografía para tontos”. Damasio define la mente como un flujo continuo de patrones mentales, y muchos de ellos parecen estar lógicamente interrelacionados. El flujo avanza en el tiempo, rápido o lento, ordenadamente o a saltos, y ocasionalmente se mueve no a lo largo de una sino varias secuencias. A veces las secuencias son concurrentes, otras veces convergentes y divergentes, o incluso superpuestas (2015, p. 272).

Recientemente, en su *newsletter*, OpenAI, la creadora del software ChatGPT, anunció que su modelo ‘o1’ es capaz de razonar. La compañía presenta este modelo como un cambio radical en el avance del desarrollo de la inteligencia artificial generativa. Según la compañía, este nuevo modelo está diseñado para razonar a través de las decisiones, como lo hacen los humanos (Stokel-Walker, 2024). Aunque así, se entiende que las máquinas no piensan, porque sus operaciones maquímicas no constituyen pensamientos. Existen diferencias muy relevantes y notables en la etapa actual del desarrollo

tecnológico entre el aprendizaje humano y el aprendizaje automático, que aún impiden la sustitución de agentes humanos por agentes automatizados en una amplia gama de tareas.

Tavares-Pereira (2021) presenta muy acertadamente las dificultades casi insuperables, en el estado actual de la tecnología, de utilizar agentes automatizados que actúan como jueces, por ejemplo. Esto no significa que las máquinas no puedan utilizar diferentes métodos y técnicas que les permitan superar a los humanos en la resolución de problemas complejos. Kai-Fu Lee (2019, p. 20) afirma que se subestima el poder de la IA para producir inteligencia sobrehumana en áreas muy especializadas. Tampoco se dice que no se deba buscar una aproximación con los métodos y técnicas de los sistemas psíquicos. Sin embargo, una cuestión fundamental es que dotar a los computadores de inteligencia no necesita equivaler al camino tortuoso adoptado por la naturaleza durante el proceso evolutivo que originó los cerebros, los sistemas psíquicos y la revolución cognitiva de la que habla Yuval Noah Harari (2017, pp. 32–33).

1.2. Sistemas expertos frente a redes neuronales

A mediados del siglo pasado, inicio del proceso técnico para dotar de inteligencia a los ordenadores, algunos informáticos adoptaron un enfoque basado en reglas, creando “sistemas simbólicos” o “sistemas expertos”, enseñando al ordenador a pensar codificando una serie de reglas lógicas. A partir de entrevistas con expertos en ciertos temas, codificaron su sabiduría para la toma de decisiones en computadoras. Este enfoque funcionó incluso para la ejecución de tareas simples y bien definidas, como juegos simples, pero se desmoronó cuando el universo de opciones era demasiado complejo.

Otros científicos de la computación, B. Widrow, G. Carpenter, S. Grossberg, P.J. Werbos, D. Parker, D. Rumelhart y otros, han adoptado otro enfoque, basado en “redes neuronales”. El campo de las redes neuronales siguió un camino diferente. En lugar de enseñar a la computadora las reglas que el cerebro humano ya domina, estos investigadores buscaron replicar el cerebro humano en sí. Al observar que las complejas redes de neuronas en el cerebro de los animales eran las únicas capaces de generar inteligencia tal y como la conocemos, decidieron ir directamente a la fuente. Este enfoque imita la estructura del cerebro, creando capas de neuronas artificiales que pueden recibir y transmitir información de manera similar a las redes neuronales biológicas. A diferencia del enfoque basado en reglas, los desarrolladores de redes neuronales a menudo no proporcionan reglas específicas para la toma de decisiones. En su lugar, alimentan las redes con numerosos ejemplos de datos de un fenómeno específico y determinado - como imágenes, partidas de ajedrez o sonidos - lo que permite a las propias redes identificar patrones en los datos. Cuanta menos intervención humana, mejor (Lee, 2019, pp. 20-21).

2. ¿Existe Inteligencia Artificial?

La inteligencia es un término amplio y difícil de definir. Kevin Kelly afirma que nuestra propia inteligencia tiene una comprensión muy pobre de lo que es la inteligencia (2020). Durante mucho tiempo, la gente ha estado estudiando y tratando de definir qué *es realmente la inteligencia* o ser *inteligente*.

El término *inteligencia* ayuda a entender la etapa actual de desarrollo de la inteligencia artificial con énfasis en el aprendizaje automático, como se verá en el siguiente enfoque reflexivo.

2.1. El término Inteligencia

A pesar de que existe una imposibilidad básica de analizar ciertos fenómenos sistémicos, sus detalles, sus características y la forma en que hacen sus selecciones internas para obtener un resultado, un observador externo utiliza términos para designar estos fenómenos. Incluso si no entendemos exactamente cómo un sistema psíquico autorreferencial elige la solución específica a un problema, este fenómeno sistémico se llama *inteligencia*. La imposibilidad de observar cómo la información desencadena cambios estructurales en un sistema sin interrumpir su autoidentificación se denomina *aprendizaje* (Luhmann, 1998, pp. 119-120).

Luhmann afirma que estas atribuciones de términos son artificios de los observadores, a través de los cuales se interpreta lo inobservable. Uno se acostumbra a este proceso y se cree que una persona tiene inteligencia y es capaz de aprender. Y esto difícilmente puede ser contradicho, porque nadie puede observar estos fenómenos nombrados con más precisión de lo que lo permiten los términos que se les atribuyen. Hay muchos ejemplos que demuestran la inutilidad de la búsqueda de un sustrato psíquico o incluso orgánico de algo como la inteligencia y el aprendizaje. Al final, se concluye que estos términos son herramientas interpretativas utilizadas por los observadores para dar sentido a procesos que no pueden ser observados directamente (Luhmann, 1998, pp. 119-120).

Las cualidades abstractas de estos términos permiten a un observador aplicar estas mismas herramientas a su propia autoobservación, creyendo que las tiene y también comienza a atribuirles, por ejemplo, a las máquinas. Es muy difícil, y hay que tener buenos argumentos para refutar estas atribuciones, porque la observación está limitada por los propios términos que definen estos conceptos.

2.2. Inteligencia Artificial

El área de conocimiento de la IA tiene una larga historia que comienza en los primeros días de las computadoras. El término *inteligencia artificial* se registró por primera vez en 1956, en la Conferencia del Dartmouth College en New Hampshire (EE.UU.) y se refería a un nuevo campo del conocimiento, aunque las ideas sobre esta área son anteriores a 1956, ya que se remontan a la Segunda Guerra Mundial. En 1943, Warren McCulloch y Walter Pitts escribieron un artículo sobre las estructuras de razonamiento artificial en forma de modelo matemático que imitan el sistema nervioso humano (De Castro Barbosa y Bezerra, 2020, p. 93).

Obviamente, en ese momento, también estaban trabajando con datos que todavía eran muy incipientes. Hoy en día, en cualquier cafetería de Silicon Valley, la gente debate si estos algoritmos son solo código como cualquier otro, con personas a cargo, o si los que piensan de esa manera no están entendiendo la profundidad de esta nueva tecnología. Los argumentos no son del todo racionales. Muchos individuos se impresionan con los modelos generativos y se mantienen vigilantes a la espera de una mayor inteligencia. La posición más pragmática sería pensar en la IA como una herramienta y no mitificar la tecnología, sino asumir que no existe la inteligencia artificial, escapando un poco del enfoque antropomórfico, para que se pueda trabajar mejor y poder usarla bien como medio para diversos fines. Gestionar las nuevas tecnologías de forma inteligente requiere pensar de esta manera, argumenta Lanier (2023).

2.3. Inteligencia artificial versus inteligencia humana

Jean Piaget (2014) afirma que la inteligencia es un término genérico que designa formas superiores de organización o equilibrio de las estructuras cognitivas. La inteligencia no consiste en una categoría aislada y discontinua de procesos cognitivos, como si fuera una estructurante entre otras, sino en una forma de equilibrio hacia la que tienden todas las estructuras (pp. 33-34). Kelly (2020) argumenta que tendemos a pensar en la inteligencia como una sola dimensión; como si se pudiera evaluar con una sola nota que se hace cada vez más grande, cuanto más inteligente es el evaluado. La medición del coeficiente intelectual es un ejemplo, es decir, se puede tener una escala creciente empezando por un ratón, pasando por un orangután, luego un individuo aparentemente estúpido y culminando con los genios con puntuaciones muy altas. La inteligencia humana, sin embargo, es como una sinfonía con diferentes notas.

En la década de 1970 (Davis & Christodoulou & Seider & Gardner) presentaron la Teoría de las Inteligencias Múltiples (IM) de Howard Gardner, que desafía la concepción tradicional de la inteligencia como una entidad única y general, medible por un tipo específico de prueba, al clasificar la inteligencia en siete categorías: lingüística, lógico-matemática, espacial, musical, corporal-cinestésico, interpersonal e intrapersonal (2011, pp. 485-503). Obviamente, tal clasificación es esclarecedora, pero no explica claramente los matices estructurales de las capacidades humanas para resolver problemas. Hay individuos que se diferencian de la mayoría en varias de estas categorías señaladas por Gardner, siendo muy buenos en varios o excepcionalmente buenos en uno y muy por debajo de las expectativas en otros.

A su vez, ¿podrían los animales, y no los humanos, tener diferentes tipos de inteligencia? ¿Y las máquinas con modelos generativos tendrían un tipo especial de inteligencia, una lógico-matemática, ya

que se basan en resolver problemas matemáticos, realizar operaciones matemáticas, identificar patrones y permitir diálogos con humanos que la mayoría de las veces presenta argumentos coherentes?

Pronto, los algoritmos basados en técnicas de IA conducirán los coches mejor que los humanos, ya que no se distraen; si están conduciendo, se enfocan solo en esa actividad y procesan la información mucho más rápido, además de analizar contextos mucho más amplios para la toma de decisiones. Además, según Kai-Fu Lee, no cabe duda de que los algoritmos de IA acabarán superando a los médicos humanos en su capacidad para diagnosticar enfermedades y recomendar tratamientos (2019, p. 249). ¿Significa esto que estos algoritmos son más inteligentes que los humanos? En algunos aspectos, sin embargo, estamos lejos de tener computadoras con las habilidades de los niños de 3 años, como lo demuestra muy bien Fei-Fei Li. Un niño de tres años tiene mucho que aprender sobre el mundo, pero ya es un experto en dar sentido a lo que observa (2017), mientras que los ordenadores más avanzados aún no realizan esta tarea. Un ser humano aprende de pocos datos, mientras que una computadora necesita muchos datos y ocurrencias para presentar estadísticamente casos similares. Aamodt y Wang dicen que un bebé, después de ver una serie de imágenes de gatos, obtiene una mirada más larga y presta más atención a una sola imagen de un perro que se le presenta. Esto significa que los bebés pueden distinguir a los gatos de los perros con una sola ocurrencia diferenciada (2013, p. 33).

Por lo tanto, el cerebro humano realiza la compleja función de procesar un flujo constante de información sensorial. En medio de un gran volumen de datos irrelevantes, necesita identificar y extraer información relevante, actuar sobre ella y, posiblemente, almacenarla en la memoria a largo plazo. La comprensión de la memoria humana presenta un desafío importante: ¿cómo se pueden formar nuevos

recuerdos sin que los recuerdos existentes se corrompan o se pierdan? Este desafío se conoce como *el dilema estabilidad-plasticidad* (Loesch y Sari, 2011, pp. 111-112).

Stephen Grossberg, investigador en el campo de la inteligencia artificial y la neurociencia computacional, conocido por sus importantes contribuciones al desarrollo de modelos teóricos que explican cómo el cerebro procesa la información (citado en Loesch y Sari, 2011), describe el problema de la *estabilidad-plasticidad* afirmando que el desafío central radica en cómo se puede diseñar un sistema de aprendizaje para adaptarse a la información relevante permaneciendo estable frente a información irrelevante. Entre las preguntas que se plantean para hacer frente a este reto se encuentran: ¿cómo puede este sistema cambiar entre los modos de estabilidad y plasticidad, garantizando la flexibilidad sin comprometer la estabilidad y evitando el caos durante la adaptación? Más precisamente, ¿cómo puede el sistema retener el conocimiento previamente adquirido mientras continúa aprendiendo nueva información? ¿Cómo evitar que la adquisición de nueva información interfiera con las memorias ya consolidadas? (p. 112). Con esto en mente, Grossberg desarrolló la Teoría de la *Resonancia Adaptativa* (ART), para su uso en redes neuronales artificiales, en 1976.

ART busca demostrar cómo integrar un modelo de aprendizaje competitivo en una estructura de control autoorganizada, de modo que el reconocimiento de patrones y el aprendizaje autónomo permanezcan estables incluso cuando se exponen a una secuencia aleatoria de patrones de entrada. La autonomía, en este contexto, implica la ausencia de control o supervisión externa. Los sistemas de aprendizaje autónomos deben aprender de sus propios errores, observaciones e interacciones con el entorno. Deben ser capaces de generalizar, creando categorías completas a partir de datos específicos. Para ello, estos sistemas necesitan organizar los conocimientos adquiridos, identificar la información

relevante y reorganizarse cuando sea necesario. Los detalles inicialmente irrelevantes pueden volverse importantes más tarde, lo que requiere que el sistema se ajuste automáticamente (Loesch & Sari, 2011, p. 112).

Si bien es posible analizar las regularidades de las relaciones externas de un sistema psíquico, no es posible conocer en detalle la parte interna de este sistema, debido a su complejidad (Luhmann, 2011, p. 65). Así, en la medida en que se le da autonomía a una red neuronal artificial para organizar el conocimiento adquirido a partir de los datos y reorganizarse cuando sea necesario, estableciendo sus propios parámetros, es imposible observar exactamente la parte interna de este sistema debido a su complejidad, y por lo tanto no se puede decir que la forma de autoorganización de una red neuronal artificial sea la misma que la de un sistema psíquico, aunque tenemos acceso a las regularidades de sus manifestaciones.

2.4. Inteligencia humana, máquinas y emociones

Según Capra, existen fuertes evidencias de que la inteligencia humana, la memoria y las decisiones humanas, nunca son completamente racionales, sino que siempre se manifiestan coloreadas por las emociones. Nuestro pensamiento siempre va acompañado de sensaciones y procesos somáticos (2010, p. 68). Según la teoría luhmanniana de los sistemas, el sistema psíquico no está aislado del cuerpo. El cuerpo, con sus diversos sistemas, constituye el entorno del sistema psíquico (Luhmann, 1990, pp. 103-14). El entorno es constitutivo del sistema, porque no hay sistema sin entorno y viceversa.

Según Maturana, nuestra cultura devalúa las emociones, porque estamos permanentemente haciendo una demanda de comportamiento racional, pero no se pueden ignorar las emociones;

Maturana ejemplifica el hecho de que cuando aprendió a leer en una semana, lo movió la envidia (2001, pp. 108-109).

Damásio concluyó, con estudios en su laboratorio, que la emoción es parte de los procesos de razonamiento y decisión, ya sea buena o mala. Los hallazgos tuvieron en cuenta a individuos que eran completamente racionales en la forma en que llevaban sus vidas hasta el momento en que, como resultado del daño neurofisiológico en lugares específicos del cerebro, perdieron cierta clase de emociones y, como resultado, también perdieron la capacidad de tomar decisiones racionales (2015, pp. 43-44).

La pregunta es si las máquinas pueden ser verdaderamente inteligentes sin emociones. El término “mecanizado” significa, por un lado, la ausencia de sentimientos, emociones o intereses; por otro lado, por el contrario, representa un compromiso obsesivo con una sola causa. Ambos extremos sugieren una falta de humanidad e incluso estupidez, ya sea debido a la incapacidad de realizar múltiples tareas o a la falta de propósito. La IA necesita cierto nivel de emoción o experiencia para alcanzar su máximo potencial y evitar caer en extremos dañinos (Minsky, 1989, p. 163). Giskard, el robot de Isaac Asimov, afirmó que “dado que las emociones son pocas y las razones son diversas [...], podemos predecir más fácilmente el comportamiento de una multitud que el de una persona” (Isaac Asimov, citado en Minsky, 1989, p. 172).

La tecnología actual no está dotando a los ordenadores de emociones para sus decisiones y esto puede descalificarlos para tareas que requieren esta cualidad humana. Pero no se puede decir que en el futuro no se puedan configurar algoritmos con hiperparámetros que les obliguen a simular el proceso emocional, a tener más o menos “emociones” a la hora de crear sus parámetros internos como

resultado de sus elecciones, o que trabajen con métodos y técnicas que sean independientes de las emociones y que las máquinas puedan hacer elecciones iguales o mejores que las selecciones que hacen los seres humanos para resolver problemas.

Siempre se enfatiza el aspecto positivo de las emociones, especialmente el comportamiento amoroso de los seres humanos, y se olvidan las barbaridades cometidas por estos mismos seres humanos emocionales, individual o colectivamente. Los seres humanos, paradójicamente, cometen inhumanidades. Quizás las máquinas inteligentes puedan ser superiores a los humanos en la toma de ciertas decisiones racionales no influenciadas por las emociones o ser un complemento fundamental de los humanos ante decisiones complejas, presentando nuevos escenarios y contextos que les ayuden en sus decisiones involucradas por las emociones. Las máquinas pueden, quizás, ser una posibilidad de salvación para el mantenimiento y el aumento de la *humanidad* y no al revés como algunos defienden.

3. Las máquinas: de la alta fiabilidad a las herramientas que pueden cometer errores

Un estudio de Pinker (2023) sugiere que las personas prestan menos atención a los detalles cuando creen que una máquina les está ayudando. El fenómeno, conocido como “merodeo social”, se produce cuando los individuos de un equipo reducen su esfuerzo al depender de otros para hacer el trabajo. La investigación demostró que esta dinámica también se aplica a la relación entre humanos y robots. Los participantes que creían que estaban siendo asistidos por un robot para inspeccionar componentes electrónicos tuvieron un desempeño un 20 por ciento peor que aquellos que trabajaron solos, incluso si el robot no estaba realmente presente. Pinker dice que este hallazgo plantea preocupaciones sobre la colaboración entre humanos e IA en áreas críticas como la medicina y la aviación, donde la falta de atención puede tener graves consecuencias, y que la dependencia excesiva

de los sistemas automatizados puede llevar a las personas a “mirar, pero no ver”, comprometiendo la seguridad y la eficiencia.

El 22 de junio de 2023, un juez federal de Manhattan (EE.UU.) multó a dos abogados por citar una investigación jurídica falsa generada por ChatGPT en un caso de lesiones personales. El juez dictaminó que el caso no tenía precedentes y sancionó la firma de abogados Levidow, Levidow & Oberman y a dos de sus abogados con una multa de 5.000 dólares por utilizar material falso creado por inteligencia artificial generativa en un caso judicial. El uso de encuestas falsas se hizo para apoyar a un cliente que afirmó haberse lesionado en un vuelo comercial. Los abogados castigados, Steven A. Schwartz y Peter LoDuca, utilizaron ChatGPT para respaldar el argumento de que la demanda presentada por su cliente era válida. Schwartz afirmó desconocer la capacidad de la herramienta para “inventar casos”, y LoDuca dijo que confiaba en la investigación de su colega. El juez, sin embargo, señaló que los abogados inicialmente insistieron en la veracidad de las encuestas, incluso después de ser interrogados. La firma Levidow, Levidow & Oberman dijo que acatará la decisión, pero niega la mala fe (Mulvaney, 2023).

Gerrit De Vynck dice que el término *alucinaciones* se ha popularizado en el área de la IA para identificar los resultados inexactos de los nuevos algoritmos generativos. Los técnicos, investigadores y escépticos de la IA están enfocados en resolver este problema, aunque no se sabe con certeza cuándo sucederá y si sucederá. Ahora que millones de personas acceden a esta tecnología y la integran en campos críticos y sensibles como la medicina y el derecho, es crucial comprender este proceso alucinógeno y encontrar formas de mitigarlo (2023).

Gran parte del trabajo con el *aprendizaje automático* se lleva a cabo en segundo plano. Los algoritmos se utilizan para predecir la demanda, organizar los resultados de búsqueda, recomendar productos, detectar fraudes, realizar traducciones y otras tareas. Aunque menos evidente, el impacto futuro del *machine learning* será de este tipo: discreto, pero con mejoras sustanciales en nuestras operaciones principales (Balakrishnan citado en Mussa, 2020, p. 161).

La transformación de máquinas precisas en máquinas imprecisas tuvo lugar de forma gradual e imperceptible con la llegada de la IA. Los modelos generativos más recientes demuestran que seleccionar y fusionar datos para componer información rica y útil, sobre cualquier tema, con cierta creatividad, también abre las puertas a resultados cuestionables y a menudo erróneos.

4. Potencia computacional, algoritmos, datos y modelos de inteligencia artificial

Hay algunos componentes esenciales en la estructuración de lo que se denomina inteligencia artificial, sin los cuales la IA tal y como la conocemos hoy en día, con sus avances más recientes, los modelos generativos, no sería posible. La delimitación de este estudio se centra en los componentes técnicos, dejando de lado el componente humano, aunque es el componente más importante, ya que es el que crea los algoritmos básicos; genera datos brutos o trabajos de preparación y etiquetado que serán utilizados por las máquinas; utiliza los resultados de la IA y justifica su existencia.

4.1. componentes esenciales en la estructuración de la IA

Desde el punto de vista técnico, destacan algunos componentes, sin una secuencia jerárquica, considerando la interdependencia entre ellos.

En primer lugar, hay que considerar la máquina desde el punto de vista del *hardware* y de la gran capacidad informática. En segundo lugar, se enfatiza la máquina desde el punto de vista del algoritmo (*software*) creado por humanos y configurado por ellos. En tercer lugar, los millones de datos disponibles en formato digital. En cuarto lugar, el *modelo*, resultado de la interacción entre el algoritmo creado por un agente humano y los datos. El modelo es un algoritmo inicialmente rudimentario, con hiperparámetros aplicados por humanos y seleccionados arbitrariamente a partir de pruebas con masas iniciales de datos. Sin embargo, es al mismo tiempo, un nuevo algoritmo creado durante la operación en el proceso de aprendizaje, transformándose en una base de conocimiento estructurada con miles de millones e incluso billones de parámetros internos creados por el propio algoritmo. Esta base puede ser utilizada por otros algoritmos, agentes automatizados, que se comunican con agentes humanos. En este sentido, el algoritmo crea sus propias estructuras y sus propios elementos durante el proceso de aprendizaje.

4.1.1. Potencia de cómputo

Es notable que tanta potencia de cómputo, anclada en computadoras de alto consumo energético que deben estar en ambientes refrigerados, para simular un pedazo de tejido humano que pesa una libra y media, cabe dentro de un cráneo, aumenta la temperatura corporal solo unos pocos grados, usa solo veinte vatios de energía y solo necesita unas pocas hamburguesas para mantenerse a flote (Kaku, 2023, p. 255).

Las herramientas de IA generativa requieren servidores extremadamente potentes con chips diferenciados y caros que consumen mucha energía (Dotan & Seetharaman, 2023).

Algunos expertos proyectan que el consumo de electricidad para las potentes máquinas necesarias para la IA pronto podría requerir la adición equivalente a la generación de energía de un pequeño país. Desde 2010, el consumo de energía para los centros de datos se ha mantenido casi estable, en alrededor del 1 por ciento de la proporción de la producción mundial de electricidad, según la Agencia Internacional de Energía. Sin embargo, la rápida adopción de la IA puede representar un aumento significativo en el consumo de electricidad (Mims, 2023). Los algoritmos de IA, basados en *el aprendizaje profundo*, utilizan procedimientos matemáticos complejos y requieren un *hardware* grande y muy rápido para procesar los datos (Mussa, 2020, p. 90-92).

4.1.2. Algoritmos desarrollados por humanos

Según Knuth, la noción de algoritmo es básica para toda programación informática y se refiere a un conjunto finito de reglas que proporciona una secuencia de operaciones para resolver un tipo específico de problema. Knuth ejemplifica con los términos *receta*, *proceso*, *método*, *técnica*, *procedimiento* y *rutina* (Donald E. Knuth citado en Tavares-Pereira, 2021, pp. 261-264). Los algoritmos, por lo tanto, son conjuntos de instrucciones paso a paso que definen cómo una computadora debe procesar datos para lograr un objetivo específico. Esto presupone una estructura muy rígida.

Las analogías de Knuth, incluida una *receta*, pueden no expresar exactamente las dificultades que surgen para la escritura de algoritmos por parte de agentes humanos.

Una de las cosas más difíciles de entender para una persona no experta, cuando se habla de programar una computadora, es que no se le puede decir: “tres huevos medianos”, o “sal al gusto”; no hay interpretación. Una computadora requiere una disciplina que enfatice el uso de la lógica

matemática. Las oraciones deben tener solidez sintáctica y semántica y no pueden ser ambiguas (Pereira, 2022).

En un ordenador, ¿es útil la parte física, el *hardware* (tecnología dura), sin programas? ¿Es útil un programa, *software* (tecnología blanda), sin hardware? No hay que pensar mucho en ello para decir que hay una interdependencia entre *el hardware* y el *software*. La computadora es la unidad de diferencia entre estos dos tipos de tecnología. En otras palabras, la computadora no es solo *hardware* o solo *software*, no es ni una cosa ni la otra. Es algo diferente. Cuando un programa se está ejecutando, hay una simbiosis entre las tecnologías blandas y duras; estas tecnologías requieren de científicos especializados en cada una de las áreas y su avance se produce por separado, pero existe una coevolución entre ellas. Cuanto más *hardware*, más *software* complejo se podía desarrollar, y el *software* complejo requería el desarrollo de hardware más potente en términos de velocidad y capacidad de almacenamiento; esto se ve claramente con el progreso en el área de la IA. Algunos algoritmos que se conocen desde hace algún tiempo solo se han vuelto efectivos con *hardware* nuevo y mucho más rápido. Pero lo que hace que el ordenador sea tan especial es precisamente esta unión, es decir, un mismo *hardware* puede dar vida a muchos algoritmos diferentes.

Pedro Domingos dice que, si retrocediéramos en el tiempo, a principios del siglo XX, y le dijéramos a la gente que pronto se inventaría una máquina y que resolvería problemas en todas las áreas del conocimiento, la misma máquina para todos los problemas, nadie lo creería. El sentido común de la época decía que cada máquina sólo podía hacer una cosa: las máquinas de coser no mecanografiaban y las máquinas de escribir no cosen (pp. 57-58).

Los algoritmos desarrollados por humanos tienen sus estructuras muy bien conocidas y dominadas por quienes los desarrollan. Muchas pruebas a lo largo del tiempo garantizan una cierta fiabilidad a la hora de introducir datos estructurados con significados esperados, normalmente de una zona concreta. Las reglas de las transformaciones sobre los datos se explican en el código y todos los que tienen acceso al código pueden conocer estas reglas. En los algoritmos tradicionales, entran los datos, que pasan por reglas de transformación asignadas paso a paso por los programadores para obtener los resultados deseados e iguales, si los datos son siempre los mismos.

Esta visión algorítmica se mantiene parcialmente en el caso del desarrollo de modelos generativos de IA, ya que la estructuración de los modelos ya no es responsabilidad exclusiva de los programadores, sino de procesos de autoorganización proporcionados por algoritmos que tratan los datos y los patrones existentes en ellos, creando sus parámetros internos, su conocimiento.

4.1.3. Datos

El éxito de los algoritmos de IA también depende de grandes volúmenes de datos. En la fase de implementación, la cantidad y la calidad de los datos se convierten en el factor determinante. En el contexto del aprendizaje profundo, se aplica la máxima ‘cuantos más datos, mejor’. La disponibilidad de abundantes datos permite que los algoritmos diseñados por equipos promedio superen a los desarrollados por expertos de élite. Si bien los investigadores de vanguardia siguen desempeñando un papel crucial en la evolución del campo de la IA, la creciente disponibilidad de datos es el principal impulsor de la transformación industrial en curso, liderada por el aprendizaje profundo (Lee, 2019, pp. 27-28).

4.1.4. El modelo como base de conocimiento estructurado

La IA ha liberado a los humanos de la imposible especificación y creación de algoritmos altamente complejos para resolver ciertos problemas relacionados con una determinada área del conocimiento. En muchos casos, era imposible para los programadores establecer en algoritmos los miles de reglas y sus relaciones para procesar datos y obtener información.

Un algoritmo de IA, desarrollado por un agente humano, es factible de desarrollar, porque, aunque complejo con sus códigos, técnicas, fórmulas e hiperparámetros, puede obtener patrones a partir de millones de datos y almacenarlos en una representación codificada binariamente, llamada *modelo*. Por lo tanto, un modelo es una representación codificada de un aprendizaje específico basado en datos. Un modelo no es generado solo por el algoritmo creado por el agente humano, ni es exactamente un espejo de los datos en bruto. El modelo es el resultado de la aplicación de reglas de transformación genéricas presentes en el algoritmo, que también crea parámetros internos desconocidos y complejos en tiempo de ejecución, dependiendo de los datos de entrada. Este proceso funciona como una caja negra, ya que el agente humano no ha especificado exactamente cuál debe ser la estructura del modelo codificado. La estructuración del modelo se establece mediante la operación del algoritmo utilizando los datos.

Para la cibernética, y desde un punto de vista más mecanicista, *la caja negra* y *la caja blanca* son términos convenientes y figurativos de uso bien definido, que se refieren a los procesos internos de un sistema. El término caja negra indica que un observador no necesariamente necesita tener información sobre la estructura mediante la cual se realiza la operación, ya sea por ignorancia o porque el fenómeno no puede ser observable directamente. Por otro lado, la caja blanca es algo con un plan estructural

definido para asegurar una relación *input-output* previamente determinada. (Wiener, 1970, p. 13). Los algoritmos tradicionales pueden considerarse cajas blancas, y los algoritmos de IA que se convierten en modelos creando sus propios parámetros internos son cajas negras. Cada modelo está vinculado al algoritmo y a los datos, y sus estructuras cambian si los agentes humanos modifican el algoritmo y sus hiperparámetros o la base de datos de entrada.

Las empresas de tecnología gastan sumas de dinero probando sus productos. Pero debido a la forma en que se diseñan los sistemas de IA, que construyen modelos a partir de características de millones de puntos de datos que pueden generar miles de millones e incluso billones de parámetros y relaciones, no se pueden desmontar y analizar en busca de *errores* como lo hace el *software tradicional* (McMillan, 2023). Dada la complejidad de las estructuras de los modelos generativos, Sven Cattell prefiere el término “caos” en lugar de caja negra (Citado en McMillan, 2023). Pero en este caso, el caos significa sólo una complejidad estructurada que no puede ser comprendida adecuadamente por los agentes humanos.

Los modelos son representaciones matemáticas y computacionales establecidas a partir de datos, a través de algoritmos ideados por humanos. Al final, son máquinas que se modifican a sí mismas, que aprenden. Los modelos encapsulan el conocimiento adquirido durante el entrenamiento y permiten a la IA hacer predicciones, clasificaciones o generar nuevos datos.

Las salidas de los algoritmos tradicionales reflejan claramente lo determinado por el agente humano responsable de su programación y el tratamiento de los datos está explícito en el código. Si llega un dato imprevisto, el algoritmo generalmente no sabe cómo aplicar las reglas de transformación y puede lanzar una excepción que apunte a un error.

Sin embargo, en los algoritmos de IA, la salida, el modelo generado, refleja la unidad de diferencia entre el algoritmo base y sus hiperparámetros y los datos; en otras palabras, el modelo generado no es explícito por ningún agente humano en el algoritmo base, ni su estructura ni sus elementos, ni los elementos y la estructura están presentes en los datos. No hay similitud en el algoritmo base ni en los datos. Desde el punto de vista de la Teoría de Sistemas de Niklas Luhmann, el algoritmo base con sus hiperparámetros, así como el *hardware* son máquinas alrededor de los modelos y el mundo está representado por los datos. Lo que no está en los datos no está en el mundo.

El modelo se genera durante el proceso dinámico de la operación de formación sistémica; un sistema es dinámico y no estático. El algoritmo trata diferentes tipos de datos en busca de patrones no predichos explícitamente por los agentes humanos, creando su propia información. En otras palabras, la información creada por el algoritmo no está presente en los datos, no está estructurada como una selección lista para ser importada, se crea en el proceso dinámico de formación sistémica.

Diferentes algoritmos de IA pueden encontrar información diferente en los mismos datos, ya que la información se crea en el proceso de operación, lo que implica diferencias estructurales y de elementos en el modelo generado. Hay que tener en cuenta que el proceso que se forma entre los datos y el algoritmo base está operacionalmente cerrado y ni el algoritmo ni los datos, ni ningún agente externo, puede determinar por su “voluntad” cómo se deben crear las estructuras y elementos de este modelo ni aportar conocimiento directamente.

Se afirma que existe una relación importante entre los datos y el algoritmo base para la autogeneración del modelo, sin interferencias externas. El modelo utiliza la complejidad del algoritmo base para autoconstruir sus estructuras y elementos. El modelo en construcción, de esta manera, se

acopla al algoritmo base para la percepción y el tratamiento de los datos. En otras palabras, el modelo utiliza la complejidad del algoritmo base para construir sus propias estructuras y sus propios elementos. Este acoplamiento del modelo al algoritmo base y la percepción de los datos no es contradictorio con la estructuración de los datos y la estructuración del algoritmo, que permanecen intactas. De este modo, el algoritmo y los datos son base y ganan valor estructural para la construcción del sistema emergente que es el modelo, y con ello, para la construcción de un tipo específico de realidad, que es la conexión entre algoritmo y datos. Esta irritación constante³ del algoritmo base y de los datos es lo que estimula la creación del modelo.

Al final, el modelo generado es lo que se podría llamar aprendizaje, en este trabajo un aprendizaje automático, no igual, pero sí similar a lo que ocurre con los sistemas psíquicos humanos y animales. El aprendizaje, en este contexto maquínico, así como en el contexto de un sistema psíquico, es la designación de no poder observar cómo la información, extraída de los datos, desencadena cambios estructurales parciales en un sistema, sin que el sistema pierda su característica. Sin embargo, todavía hay máquinas. *La inteligencia* es precisamente la denominación para no poder observar cómo, en este sistema autorreferencial, en contacto con uno mismo, se produce la elección que uno hace para la solución del problema.

Esto es precisamente lo que afirma Luhmann, que la búsqueda de una base para conceptos como la *memoria* y la *inteligencia* es infructuosa. Estos conceptos son en realidad constructos que

³ Esta interacción constante entre el algoritmo base y los datos es fundamental para la existencia del modelo. Según Luhmann, la línea que separa las operaciones del sistema de su entorno también distingue lo que estimula el sistema y lo que no. Sólo de esta manera un sistema puede convertir las "irritaciones" en causalidades. El sistema no puede hacer frente a todas las irritaciones del entorno (Luhmann, 2011, p. 132-134).

utilizamos para interpretar el comportamiento de sistemas complejos. Es importante destacar que esto no quita la existencia del algoritmo base y los datos y la necesidad de ellos para que el modelo se autoconstruya, porque, al final, las características de este entorno están presentes en el modelo, en este sistema emergente, es decir, pertenecen a esta realidad que se forma, pero ya no se puede buscar un sustrato, ni en el algoritmo ni en los datos. Eso coincide exactamente con el modelo emergente. Esta falta de un sustrato de esta nueva estructura adquirida y emergente, el modelo, se compensa justamente con la libertad de elección para crear esta estructura y sus elementos, lo que indica que podría ser de otra manera, refiriéndose al concepto de *contingencia*. La elección y formación de la estructura y los elementos del modelo se fijan durante el funcionamiento del algoritmo y su relación con los datos, y no pueden existir de ninguna otra manera.

El progreso de la IA se ha producido de forma gradual y se ha producido una disrupción con los llamados *modelos generativos*. ¿En qué se diferencian estos modelos en relación con los que se han difundido hasta ahora denominados *discriminativos*?

4.1.4.1. Modelos discriminativos.

Los modelos, comúnmente utilizados hasta ahora, que han dominado el mundo de la IA en los últimos tiempos, están siendo empujados a una categoría llamada “discriminativa”. Esta categoría está siendo eclipsada por la categoría de lo “generativo”. Hay muchas formas de crear programas que aprenden, pero en general, se utilizan tres técnicas básicas: el aprendizaje no supervisado, el aprendizaje supervisado y el aprendizaje por refuerzo. Todos se basan en datos históricos. Para los modelos discriminativos, el más común es el uso del aprendizaje supervisado, en el que se proporcionan dos conjuntos de datos al ordenador, como son los textos y sus clasificaciones realizadas por agentes

humanos. La gran característica de una computadora que se basa en un “modelo discriminativo” es la capacidad de comparar datos y encontrar otros similares. En el caso de las imágenes, un modelo discriminativo puede diferenciar a un perro de un gato, ya que esto se enseñó a partir de miles de imágenes puestas a disposición (Pereira, 2023).

4.1.4.2. Modelos generativos

Usando el ejemplo de imágenes dado anteriormente, un modelo generativo puede generar nuevas fotos de animales que parecen animales reales.

Mientras que un modelo discriminativo depende de etiquetas asociadas a los datos: esto es un gato, esto es un perro, esto es un tigre, es decir, depende del aprendizaje supervisado, un modelo generativo puede o no utilizar este recurso.

El modelo generativo puede, en este caso, aprender de miles de millones de puntos de datos no supervisados. Es decir, sin que los humanos tengan que etiquetar los datos de forma explícita. En términos de ejemplo, se puede describir que existen modelos generativos simples que sugieren la siguiente palabra en una secuencia de palabras para formar una oración y, en este caso, no hay necesidad de supervisión por parte de agentes humanos.

Los algoritmos aprenden “percibiendo” cómo se usa el lenguaje. El “conocimiento”, por lo tanto, se extrajo de la observación de miles de millones de textos que están disponibles en internet, siendo esta la forma en que el lenguaje ha sido dominado por estos nuevos algoritmos generativos que no tienen idea de lo que significa “buenos días”, pero sugieren “día” después de “bueno”, porque estadísticamente es lo que más ocurre. Esto es lo que experimentamos todos los días en nuestros

teléfonos celulares al escribir un texto en una de las aplicaciones. Los modelos generativos pueden realizar varias correlaciones con los datos que manejan. Por ejemplo, los ojos no pueden aparecer en la barbilla de las personas, generalmente están entre la frente y la nariz. Para la generación de una nueva imagen, se sigue esta regla. Si alguien le pide a la IA generativa que genere una imagen de un gato patinando, la patineta debe estar debajo de los pies del gato. Las imágenes no son simples copias, sino generaciones de nuevas imágenes. La distribución de datos en un contexto forma parte del enfoque de la IA que utiliza modelos generativos (Pereira, 2023).

5. Consideraciones finales

La Inteligencia Artificial, impulsada por los avances en el aprendizaje automático y los modelos generativos, está redefiniendo rápidamente varios sectores de la sociedad. Sin embargo, la antropomorfización de la IA, aunque comprensible, puede dar lugar a expectativas poco realistas y riesgos potenciales.

Hay algunos componentes esenciales en la estructuración de lo que se denomina inteligencia artificial, sin los cuales la IA tal y como la conocemos hoy en día, con sus avances más recientes, los modelos generativos, no sería posible. La delimitación de este estudio en los componentes técnicos, sin tener en cuenta el componente humano, demostró la dependencia de la IA de grandes volúmenes de datos y potencia computacional, contrastando con la capacidad humana de aprender con pocas ocurrencias.

Es fundamental reconocer que la IA, a pesar de su capacidad para procesar grandes volúmenes de datos y realizar tareas complejas, aún está lejos de replicar la inteligencia humana en su totalidad. La

comprensión de los matices del lenguaje, la toma de decisiones éticas y la capacidad de innovar creativamente siguen siendo desafíos importantes para la IA.

Además, la creciente integración de la IA en áreas críticas como la medicina y el derecho requiere un enfoque cauteloso. El exceso de confianza en los sistemas automatizados puede conducir a errores graves, y la falta de transparencia en los modelos de IA plantea preguntas sobre la responsabilidad y el sesgo de las soluciones dadas por la IA.

Por lo tanto, el futuro de la IA depende no solo de los avances tecnológicos, sino también de un debate crítico y racional sobre sus límites e implicaciones éticas. La claridad conceptual a la hora de distinguir entre humanos y máquinas es esencial para garantizar que la IA se utilice de forma responsable y beneficiosa para la sociedad.

La IA debe verse como una herramienta poderosa, pero no como un sustituto de la inteligencia humana. La colaboración entre humanos y máquinas, en la que cada uno desempeña su papel específico, es el camino más prometedor para el futuro próximo de la humanidad.

Referencias

Aaamodt, S., & Wang, S. (2013). Bem-vindo ao cérebro do seu filho: Como a mente se desenvolve desde a concepção até a faculdade. Cultrix.

Bousquette, I. (2023, junio 22). Companies want workers to trust AI—but not completely. The Wall Street Journal. <https://www.wsj.com/articles/companies-want-workers-to-trust-ai-but-not-completely-3ce09029>

Capra, F. (2010). A teia da vida. Cultrix.

Damásio, A. (2015). O mistério da consciência: Do corpo e das emoções ao conhecimento de si. Companhia das Letras.

Davis, K., Christodoulou, J., Seider, S., & Gardner, H. E. (2011). The theory of multiple intelligences. En R. J. Sternberg & S. B. Kaufman (Eds.), *Cambridge handbook of intelligence* (pp. 485–503). Cambridge University Press.

De Castro Barbosa, X., & Bezerra, R. F. (2020). Breve introdução à história da inteligência artificial. *Jamaxi*, 4(1).

De Vynck, G. (2023, mayo 30). ChatGPT 'hallucinates.' Some researchers worry it isn't fixable. *The Washington Post*.

<https://www.washingtonpost.com/technology/2023/05/30/ai-chatbots-chatgpt-bard-trustworthy/>

Domingos, P. (2017). O algoritmo mestre. Novatec.

Dotan, T., & Seetharaman, D. (2023). Big Tech struggles to turn AI hype into profits. *The Wall Street Journal*. <https://www.wsj.com/tech/ai/ais-costly-buildup-could-make-early-products-a-hard-sell-bdd29b9f>

Fei-Fei, L. (2017). How we're teaching computers to understand pictures [Video]. TED.

[https://www.ted.com/talks/fei fei li how we re teaching computers to understand
_pictures](https://www.ted.com/talks/fei_fei_li_how_we_re_teaching_computers_to_understand_pictures)

Gohar, K., & Utley, J. (2024). For conversations you dread, try a chatbot. The Wall Street Journal. <https://www.wsj.com/tech/ai/for-conversations-you-dread-try-a-chatbot-8eecb643>

Harari, Y. N. (2017). Uma breve história da humanidade. LPM.

Houaiss, A., & Villar, M. (2001). Dicionário Houaiss da língua portuguesa. Objetiva.

Japiassú, H. (1990). Dicionário básico de filosofia. Jorge Zahar.

Jin, B. (2024). The 27-year-old billionaire whose army does AI's dirty work. The Wall Street Journal. <https://www.wsj.com/tech/ai/alexandr-wang-scale-ai-d7c6efd7>

Kaku, M. (2014). The future of mind: The scientific quest to understand, enhance, and empower the mind. City University of New York.

Kelly, K. (2020). How AI can bring on a second industrial revolution [Video]. TED. [https://www.ted.com/talks/kevin_kelly_how_ai_can_bring_on_a_second_industrial_re
volution](https://www.ted.com/talks/kevin_kelly_how_ai_can_bring_on_a_second_industrial_revolution)

Kurzweil, R. (2014). Como criar uma mente: Os segredos do pensamento humano. Aleph.

Lanier, J. (2023). There is no A.I. The New Yorker.

<https://www.newyorker.com/science/annals-of-artificial-intelligence/there-is-no-ai>

Lee, K. (2019). Inteligência artificial: Como os robôs estão mudando o mundo, a forma como amamos, nos comunicamos e vivemos. Globo Livros.

Loesch, C., & Sari, S. T. (1996). Redes neurais artificiais: Fundamentos e modelos. Editora da Furb.

Luhmann, N. (1990). Sociedad y sistema: La ambición de la teoría. Espasa.

Luhmann, N. (1998). Sistemas sociales: Lineamientos para una teoría general. CEJA.

Luhmann, N. (2005). A realidade dos meios de comunicação. Paulus.

Luhmann, N. (2006). La sociedad de la sociedad. Herder.

Luhmann, N. (2011). Introdução à teoria dos sistemas. Vozes.

Maturana, H. (2001). Cognição, ciência e vida cotidiana. UFMG.

Maturana, H., & Varela, F. (2001). A árvore do conhecimento. Palas Athena.

Mims, C. (2023). AI is ravenous for energy. Can it be satisfied? The Wall Street Journal. <https://www.wsj.com/tech/ai/ai-energy-consumption-fc79d94f>

Minsky, M. (1989). A sociedade da mente. Francisco Alves.

Minsky, M. (2022). Introduction to “The Society of Mind” [Video]. YouTube.

<https://www.youtube.com/watch?v=-pb3z2w9gDg>

Mulvaney, E. (2023). Judge sanctions lawyers who filed fake ChatGPT legal research. The Wall Street Journal. <https://www.wsj.com/articles/judge-sanctions-lawyers-who-filed-fake-chatgpt-legal-research-9ebad8f9>

Mussa, A. (2020). Inteligência artificial: Mitos e verdades: As reais oportunidades de criação de valor nos negócios e os impactos no futuro do trabalho. Saint Paul.

Pereira, C. J. (2022). Sou jovem: É difícil aprender a programar um computador? O Analista de Blumenau. <https://www.oanalistadeblumenau.com.br/blog/4/sou-jovem-dif-cil-aprender-a-programar-um-computador>

Pereira, C. J. (2023). Unleashed the beast – A inteligência artificial generativa (IAG) pode ser domada? O Analista de Blumenau. <https://www.oanalistadeblumenau.com.br/blog/35/unleashed-the-beast-a-intelig-ncia-artificial-generativa-iag-pode-ser-domada>

Pinker, S. (2023). Does working with robots make humans slack off? The Wall Street Journal. <https://www.wsj.com/tech/ai/does-working-with-robots-make-humans-slack-off-b7b645c1>

Piaget, J. (2014). A psicologia da inteligência. Vozes.

Piaget, J. (2014a). Relações entre a afetividade e a inteligência no desenvolvimento mental da criança. Wak.

Sadin, É. (2020). La inteligencia artificial o el desafío del siglo: Anatomía de un antihumanismo radical. Caja Negra.

Stokel-Walker, C. (2024). OpenAI says the latest ChatGPT can ‘think’ – and I have thoughts. The Guardian.

Tavares-Pereira, S. (2021). O machine learning nas decisões: O uso jurídico dos algoritmos aprendizes. ArtSam.

Vygotsky, L. S. (1998). Pensamento e linguagem. Martins Fontes.

Wiener, N. (1970). Cibernética: Ou contrôle e comunicação no animal e na máquina. Polígono e Universidade de São Paulo.