



**FACULTAD DE HUMANIDADES Y EDUCACIÓN  
INSTITUTO DE FILOSOFÍA**

Tesis para Optar al Título Profesional de Profesor de Enseñanza Media en Filosofía y en los Grados Académicos de Licenciado en Filosofía y Licenciado en Educación

**INTELIGENCIA ARTIFICIAL Y EDUCACIÓN:  
UNA REFLEXIÓN NECESARIA**

JOAQUIN VALENTINO TAPIA ROMERO

NIVEL ALCANZADO DE LA CARRERA: CURSANDO MAYORES EXIGENCIAS

PROFESOR GUÍA: RODRIGO ROBERT ZEPEDA

PROFESORA INFORMANTE: JENNY DONOSO GAMBOA

2025

## **Agradecimientos**

Deseo expresar mi más sincero agradecimiento a mi familia y amigos por su constante apoyo y aliento a lo largo de este proceso. En especial, a mi padre, cuya presencia incondicional ha sido fundamental. Asimismo, agradezco profundamente a los profesores y al cuerpo académico de esta universidad por sus valiosas enseñanzas y orientación. Un especial agradecimiento a mi profesor guía, cuya invaluable orientación, conocimientos y experticia en el campo me permitieron desarrollar las competencias necesarias para llevar a cabo la Tesis, le estoy profundamente agradecido por su invaluable contribución.

Expreso mi más sincero agradecimiento a mi pareja, cuyo apoyo incondicional y aliento constante fueron fundamentales para superar los desafíos inherentes a este proceso. Su amor fue un bálsamo para mi alma en un momento de profunda vulnerabilidad, proporcionándome la fortaleza necesaria para superar un desafío personal que marcó un antes y un después en mi vida. Su cariño y comprensión me proporcionaron la fortaleza emocional necesaria para perseverar y alcanzar esta meta.

Con profunda emoción, agradezco a mi madre por su incansable apoyo y cariño. Su sabiduría ha enriquecido mi vida y su amor me ha dado la fuerza necesaria para superar cualquier obstáculo. Sus enseñanzas han sido mi brújula y su cariño, mi refugio. No hay palabras que puedan expresar la gratitud que siento por todo lo que ha hecho por mí. A ella dedico este logro como un humilde reconocimiento a todo lo que ha significado para mí.

# Índice

<b>Introducción.....</b>	<b>5</b>
<b>Capítulo 1 Bases teóricas.....</b>	<b>8</b>
<b>Capítulo 2 Perspectivas De La Ciencia Cognitiva.....</b>	<b>16</b>
<b>Capítulo 3 La Inteligencia Artificial.....</b>	<b>23</b>
<b>Capítulo 4 La Inteligencia Artificial Generativa y la Educación.....</b>	<b>33</b>
<b>Conclusión.....</b>	<b>42</b>
<b>Bibliografía.....</b>	<b>48</b>

“La Inteligencia Artificial no te reemplazará; lo hará una persona que sepa usarla”.

Carlos Álvarez Navas-Parejo

## Introducción

En las últimas décadas, hemos sido testigos de una transformación fascinante en el campo de la Inteligencia Artificial (IA), que ha sido paralela a nuestro progreso tecnológico y social. Durante este tiempo, la historia de la IA, su evolución, ha sido impulsada por diversos y significativos avances tecnológicos que han permitido el aumento en la capacidad de procesamiento de las computadoras, el desarrollo de microchips, así como también el desarrollo de lenguajes informáticos que han facilitado la emergencia de algoritmos más complejos capaces de integrar enormes cantidades de información. Como parte de esta creciente sofisticación, los sistemas de IA en la actualidad pueden realizar tareas que hasta hace unos pocos años atrás eran tema de la ciencia ficción, como el reconocimiento de voz y de rostros, la traducción automática, el diagnóstico médico, la creación artística y la conducción autónoma, por nombrar solo algunas. Hoy resulta evidente que la IA ha irrumpido en la vida cotidiana de las personas a un ritmo mucho más rápido de lo que se anticipaba hace unos pocos años atrás.

Desde los asistentes virtuales en los teléfonos inteligentes hasta los sistemas de recomendación en las plataformas de streaming, la IA se ha infiltrado en diversos aspectos de la vida diaria de las personas, pues dada su capacidad para aprender y colaborar en múltiples tareas se ha convertido en una herramienta indispensable en los requerimientos de la sociedad actual, siendo capaz de analizar grandes cantidades de datos, identificar patrones y realizar predicciones, lo que la hace ideal para facilitar la toma de decisiones en una amplia gama de aplicaciones. El uso de asistentes virtuales como Siri, el Asistente de Google o Alexa, se ha vuelto algo cotidiano, al igual que los sistemas de recomendación en plataformas como Netflix o Spotify, que nos sugieren contenido personalizado. La IA se ha ido constituyendo en una parte integral de nuestra experiencia diaria, pues facilita la forma en que nos comunicamos, trabajamos, nos entretenemos y nos informamos. Desde su introducción en la segunda mitad del siglo XX, la IA ha tenido una significativa evolución tanto conceptual como tecnológica, llegando en la actualidad a convertirse en una de las disciplinas más dinámicas que determinan tanto nuestra vida social como económica, condición que se advierte en el peso que tienen las empresas vinculadas a la IA en los principales índices financieros a nivel mundial.

Para comprender la evolución de la IA es fundamental situar su desarrollo en el contexto de la llamada ciencia cognitiva, la cual surge hacia mediados del siglo XX como un intento de integrar a las diversas disciplinas que abordan el tema de la cognición, entre las que se encuentran la psicología, la neurociencia, la lingüística, la informática, la antropología cognitiva y la filosofía de la mente. En el caso específico de la ciencia de la informática o la computación, a la que más tarde se agrega la robótica, ésta se propone comprender y replicar los procesos cognitivos humanos en sistemas artificiales, explorando nuevas teorías y enfoques para mejorar la capacidad de las máquinas para realizar tareas inteligentes. Allen Newell y Herbert Simon, fundadores del enfoque computacional, explicitan su visión de la IA al señalar que:

Los símbolos descansan en la raíz de la acción inteligente que es, desde luego, el tema principal de la inteligencia artificial. En cuanto a eso, es una cuestión central para toda la ciencia de la computación, ya que toda la información se procesa en computadoras en función de fines y medimos la inteligencia de un sistema por su capacidad para alcanzar objetivos establecidos frente a variaciones, dificultades y complejidades planteadas por el ámbito de la tarea que hay que realizar. (Newell y Simon, 1994, pág. 124).

El gran avance de la IA, así como también la sofisticación de las tareas que nos demanda la vida actual, ha llevado a que emerjan preguntas y desafíos éticos, como la privacidad de los datos, el sesgo algorítmico, el impacto en el empleo, el riesgo de plagio y el uso de robots en redes sociales, entre otros, que conllevan nuevos retos y demandas a los sistemas educativos. Gómez (2021) nos indica que la IA se está volviendo más avanzada y omnipresente, así como también las voces que advierten sobre los desafíos actuales y futuros se están volviendo más enérgicas.

En el ámbito educativo, la IA presenta enormes potenciales y desafíos. Benokraitis y Feagin (2019), plantean que además de facilitar el desempeño docente, especialmente en cuestiones administrativas como la determinación de las calificaciones finales en función del peso de las evaluaciones, la IA puede servir para mejorar la redacción de exámenes, automatizar las evaluaciones y disminuir así significativamente el tiempo que antes se les dedicaba a estos procesos. Esta tecnología puede ayudar a los profesores a optimizar múltiples tareas, tales como

la creación de recursos educativos, la generación de preguntas y la organización de actividades. Uno de los eventuales riesgos, que ya está presente con el uso de la información digital, es el de fomentar la práctica de *copiar y pegar* para elaborar trabajos o ensayos al utilizar aplicaciones tales como ChatGPT, Bing IA u otra aplicación de IA generativa.

En este contexto, esta tesis pretende explorar una dimensión del impacto de la IA en la educación, centrándose en el grado de conocimiento que hay sobre estas nuevas aplicaciones y de qué forma apoyar a los educadores ante la creciente irrupción de la IA. Este trabajo inicialmente abordará el concepto mismo de la IA, su devenir desde algunos planteamientos filosóficos hasta su integración en la ciencia cognitiva, para desde ahí derivar en su historia y en algunos de los efectos que puede tener en la educación, las nuevas exigencias que conlleva su masificación, donde cabe preguntarse si están las herramientas necesarias para enfrentar los desafíos que supone el uso y masificación de la IA. Se expondrán algunas políticas públicas educacionales en Chile y en otros países, así como el impacto que estas nuevas tecnologías tienen y podrán tener en el quehacer educativo.

## Capítulo 1 Bases teóricas

Con mayor o menor conciencia de nuestra parte, en los últimos años, la IA ha estado afectando diversos ámbitos de nuestras vidas, influyendo de manera directa o indirecta en nuestra manera habitual de vivir, por lo que parece necesario introducirse de manera más formal en el tema para preguntarse y reflexionar sobre los alcances y derivas que tiene y tendrá en las comunidades educacionales la convivencia con esta emergente tecnología. Dado el amplio uso y la popularidad de las aplicaciones de la IA, puede parecer evidente entender su significado y asociarla con tecnologías de reciente aparición, sin embargo, es importante destacar que el concepto de IA tiene una historia que se remonta a la primera mitad del siglo XX, a lo largo de la cual ha sido objeto de diversos debates y controversias. Un aspecto necesario de destacar, para contextualizar adecuadamente el tema de este trabajo, es que, si bien sus orígenes formales son relativamente recientes, sus raíces están conectadas con temas clásicos de la filosofía, la psicología y la educación, como lo son la naturaleza de la mente, la inteligencia, la cognición y el aprendizaje.

La inteligencia ha sido un tema que ha ocupado a los filósofos y pensadores durante miles de años, desde lo que podríamos distinguir como los orígenes de la filosofía occidental. En la Grecia clásica, Sócrates consideraba que el alma era la sede del pensamiento (inteligencia) y de los valores morales, fomentando el uso de la mayéutica, un método de reflexión y enseñanza basado en preguntas y análisis, para guiar a sus alumnos hacia la resolución de problemas y la construcción del conocimiento. La labor filosófica de Sócrates consistió en hacer conscientes a los individuos de su ignorancia, como un preámbulo para alcanzar, a través del diálogo mayéutico, un conocimiento que se pudiera considerar esencialmente verdadero. Si bien Sócrates no contribuyó directamente al campo de la inteligencia artificial, sus ideas y métodos, en tanto maestro de Platón y éste de Aristóteles, han tenido una influencia significativa en el desarrollo de la filosofía y de la visión de mundo que occidente ha ido construyendo hasta nuestros días. La búsqueda de definiciones universalmente válidas, a las que aspiraba Sócrates, no sólo marcaron la pauta de la tradición platónica y aristotélica (Bechtel, 1991), sino que también su método dialógico, la mayéutica, de carácter eminentemente circular o reflexivo, da cuenta de la forma o patrón que, actualmente, se considera característico todo proceso cognitivo y que, por tanto, está presente

también en los algoritmos de aprendizaje o de retropropagación (en inglés Backpropagation), donde los sistemas, artificialmente creados, a partir de datos simples, van ajustando sus respuestas mediante procesos recursivos de retroalimentación para mejorar la precisión de las mismas con el paso del tiempo, un proceso eminentemente circular que caracteriza a todo aprendizaje y actividad mental (Bilen, 2016).

Sócrates fue un firme defensor del pensamiento crítico y consideraba que cuestionar y examinar nuestras propias creencias y conocimientos era esencial para alcanzar la sabiduría. En el desarrollo de la IA, este principio se traduce en la necesidad de crear sistemas que no solo se limiten a ejecutar tareas, sino que también puedan analizar, evaluar y mejorar sus propios procesos lógicos, generando respuestas más eficaces y eficientes, operaciones que definen lo que se entiende por acción inteligente. Este enfoque crítico, reflexivo, recursivo y circular ha sido fundamental para el desarrollo de tecnologías avanzadas, tales como los sistemas de aprendizaje profundo (Deep learning) basados en modelos de redes neuronales (Holdsworth & Scapicchio, 2024).

En un intento por encontrar las definiciones universalmente verdaderas que buscaba Sócrates, las que eran muy difíciles de hallar en medio de la contingencia que supone el mundo físico, Platón postulará la existencia de un mundo abstracto, su famosa teoría de las Ideas o Formas (Bechtel, 1991). Para Platón, el mundo accesible a nuestros sentidos es solo una sombra de la verdadera realidad, que consiste en formas eternas e inmutables. Esta conceptualización ha influido en la manera clásica de entender la IA, en cuanto a la búsqueda de representaciones ideales y abstractas del conocimiento, así como en las diversas teorías que han sostenido el carácter innato de éste. Una aplicación de esta mirada se puede advertir en el desarrollo de los algoritmos de aprendizaje automático, donde a partir de un conjunto finito de instrucciones se intenta simular procesos o construir modelos, de carácter general o universal, que puedan replicar acciones específicas en un contexto determinado (Diez, 2023).

La noción de que el fundamento del conocimiento se encuentra en un mundo ideal, etéreo, exclusivamente teórico, y no en la experiencia del mundo sensible, en la contingencia, fue difícil de sostener incluso para Platón, quien se mostró crítico de esta postura en algunos de sus diálogos, siendo Aristóteles el encargado de ofrecer un modelo alternativo que volviera a poner la atención

en el mundo material al que accedemos a través de los sentidos (Bechtel, 1991). Para el estagirita, el *noûs*, concepto griego que se suele traducir como inteligencia, es la facultad del alma que permite la comprensión y el conocimiento de los primeros principios y las verdades universales. Es a través del *noûs* que somos capaces de entender y razonar sobre el mundo que nos rodea (Seggiaro, 2014). Para acceder a estos principios generales, universales, Aristóteles sostiene que los seres humanos mediante el entendimiento activo (*nous poietikós*) somos capaces de separar las formas universales de la materia particular de la que consta una sensación, es decir, conocemos lo universal gracias a un proceso de abstracción (*apháiresis*) (Aristóteles, 1995). De esta manera, la mente o el intelecto opera con representaciones del mundo sensorial, las cuales suponen la capacidad del alma de capturar la esencia de un objeto o fenómeno. Para el reconocido profesor de filosofía de la mente William Bechtel (1991), al igual como se ha dado en el seno de la ciencia cognitiva contemporánea, el interés de Aristóteles estaba en proponer un modelo acerca de la manera que tiene la mente para representar los objetos del mundo, para operar con conceptos, definiciones, universales, abstrayéndose de los casos particulares o singulares. En el ámbito de la IA, la representación sigue siendo un concepto clave, ya que se asume que los sistemas de inteligencia artificial deben procesar datos del mundo de manera efectiva. Así, por ejemplo, los algoritmos de aprendizaje automático utilizan representaciones simbólicas que permiten a los sistemas computacionales identificar patrones y hacer predicciones. Los sistemas de IA deben ser capaces de generalizar a partir de datos específicos, inferir patrones, secuencias, pautas, un proceso que se asemeja a la abstracción humana mediante el razonamiento inductivo. Cuando un programa de IA aprende a identificar, por ejemplo, gatos en imágenes, lo hace reconociendo características generales, abstractas, que definen visualmente a un gato, independientemente de las variaciones materiales contingentes, como su color, tamaño o posición.

Otro antecedente filosófico importante de destacar en el desarrollo de la ciencia cognitiva y que, ciertamente, influyó en la conceptualización de la IA, se sitúa hacia mediados del siglo XVII, cuando el filósofo, matemático y físico francés Rene Descartes plantea que el fundamento de la propia existencia se halla en la razón, pues a partir de sus reflexiones la única certeza inicial es que él es una cosa que piensa. En su texto *Meditaciones metafísicas*, publicado originalmente en 1642,

Descartes establece un procedimiento de duda radical donde cuestiona todas las creencias de las que no está seguro. Bechtel (1991), nos plantea que Descartes contempló la posibilidad de estar bajo la influencia de un genio maligno, un *burlador sumamente poderoso y astuto*, cuyo propósito sería engañarlo en la mayor medida posible. La razón para formular tales interrogantes era para aclarar su mente de todas las proposiciones dudosas que no habían sido completamente verificadas, atribuyendo la mayoría de sus equivocaciones a su tendencia de aceptar ideas sin un examen minucioso. Una vez eliminadas las nociones erróneas, el objetivo de Descartes era edificar un nuevo conjunto de verdades científicas fundamentadas en principios indudables. La primera de estas verdades que estableció fue el reconocimiento de su propia existencia, la cual consideró una consecuencia necesaria del hecho de que estaba pensando en el momento en que planteaba dichas dudas. Desde su perspectiva, no era posible que se diera una situación en la que él pensara algo y, simultáneamente, no existiera. De este razonamiento surge la célebre expresión de Descartes *cogito ergo sum, pienso, luego existo*.

Afirmó que la idea de su existencia era «clara y distinta». Para él una idea era clara cuando captábamos su esencia; era distinta cuando la percibíamos diferenciada de otras ideas. Descartes formó la hipótesis de que todas las ideas claras y distintas son verdaderas y se marcó a la tarea de justificarlo. (Bechtel, 1991, pág. 9)

El análisis racional que desarrolló lo llevó, a enfrentar la cuestión de cómo una entidad inmaterial (*res cogitans*), como la mente, puede interactuar con una entidad material (*res extensa*), como el cuerpo. De aquí surge la idea del dualismo cartesiano, que postula que la mente y el cuerpo son entidades separadas. La ciencia cognitiva subraya que el principal aporte del racionalismo no reside en su pretensión de alcanzar un conocimiento infalible, sino en el hecho de enfatizar la importancia que tiene el razonamiento para fundamentarla. Bechtel (1991) argumenta que una corriente importante dentro de la ciencia cognitiva concibe la cognición fundamentalmente como un proceso de razonamiento. Esta idea es particularmente influyente en la comunidad de la IA, donde el enfoque que se volvió clásico o hegemónico hacia mediados del siglo XX concibe que la cognición se puede modelar mediante el diseño de programas que incorporan bases de conocimiento y mecanismos de inferencia lógica (Varela, 2000). Esta propuesta de que un

ordenador pueda simular el razonamiento, se ve respaldada por su capacidad de realizar inferencias lógicas, lo que sugiere su utilidad como modelo computacional del pensamiento (Bechtel, 1991). Además, los sistemas de IA simulan procesos mentales humanos, desde el reconocimiento de patrones hasta la toma de decisiones complejas, basándose en la idea de que estos procesos pueden ser descompuestos, analizados, en componentes lógicos y mecánicos más simples, tal como sugiere el método cartesiano (Varela, 2000).

Otro pensador que suele citarse como antecedente en los desarrollos de la IA es Gottfried Leibniz, filósofo y matemático alemán, quien en 1714 escribió su famosa obra *Monadología*, en la cual señala que los cartesianos cometieron un error conceptual, al no tener en cuenta aquellas percepciones de las que las personas no somos conscientes. En su libro, introduce el concepto de monadas, sustancias simples e indivisibles, dotadas de vida, que constituyen el universo, y que siguen cada una un conjunto particular de reglas preestablecidas. Según Leibniz, el universo se regiría por un principio de armonía, de carácter innato, que aseguraría que las monadas actúen en concordancia entre ellas sin necesidad de interacción directa. Las monadas serían activas y conscientes, haciendo posibles fenómenos como la percepción, el pensamiento y la inteligencia.

Como matemático, Leibniz desarrolló el cálculo, al igual que Newton, siguiendo la noción de límite, concepto que extendió para aplicarlo al ámbito de la percepción, donde propuso la idea de umbrales o límites de percepción, convirtiéndose así en precursor de los procesos mentales inconscientes. A partir de la máquina de calcular diseñada por Pascal, también creó la llamada Rueda, Cilindro o Máquina de Leibniz, que durante varios siglos hizo posible realizar cálculos matemáticos en forma mecánica y que fue la antesala de las calculadoras y computadoras modernas. También sistematizó la notación binaria, la cual había sido planteada anteriormente, entre otros por Francis Bacon, la cual hacía posible representar el mundo a partir de una notación mínima.

El Doctor en Filosofía Javier Legris, en *Conocimiento simbólico. Un capítulo de la historia de la metodología científica* (2005), nos plantea que, en su libro *Meditaciones sobre el conocimiento, la verdad y las ideas*, escrito en 1684, Leibniz subraya la importancia de emplear símbolos en lugar de objetos concretos para realizar análisis, un método que él denomina *pensamiento ciego* o

*simbólico*. La idea central de este concepto es que los símbolos pueden facilitar la manipulación y el entendimiento de ideas que, de otro modo, resultarían inabarcables o incluso imposibles de visualizar, lo que ilustra con el ejemplo del *quiliógono*, que se refiere a un polígono de mil lados, una figura que sólo se puede representar mediante un símbolo. Para Leibniz, el conocimiento simbólico es fundamental para explorar dominios cognitivos que trascienden nuestra intuición, como el infinito. Mientras que el conocimiento intuitivo se basa en la contemplación directa de ideas simples, el conocimiento simbólico nos permite construir representaciones de conceptos abstractos a través de signos. Leibniz sostiene que, en última instancia, todo nuestro conocimiento es simbólico, pues incluso las ideas más simples son representadas mentalmente mediante símbolos (Legris, 2005). Esto implica que nuestra capacidad de comprender y comunicar ideas está intrínsecamente ligada al uso de un lenguaje simbólico, lo cual es crucial para el desarrollo del conocimiento en cualquier ámbito, ya sea en matemáticas, filosofía y/o ciencia.

Leibniz, al desarrollar esta idea, sentó las bases para la mecanización del pensamiento y se convirtió en un referente clave en el campo del simbolismo operativo. Esta perspectiva no solo es relevante en el contexto de la filosofía y la matemática, sino que también ha tenido un impacto duradero en campos como la informática y la lógica, donde los sistemas simbólicos son fundamentales para el desarrollo de algoritmos y procesos de computación, en los cuales se asume que el mundo puede ser adecuadamente representado mediante dichos sistemas.

Legris (2005) da a entender que estas ideas dieron lugar más tarde al concepto de sistema formal, el cual alude a un sistema simbólico construido en base a reglas precisas, que está compuesto de elementos básicos llamados términos y de enunciados o proposiciones que se construyen en base a éstos. Algunas de estas proposiciones son axiomas, que se aceptan como verdades iniciales y, a partir de ellos, utilizando reglas de inferencia, se pueden deducir nuevos enunciados llamados teoremas. La noción de sistema formal inspiró la hipótesis de que la IA se puede lograr mediante la manipulación de símbolos, una idea fundamental en la arquitectura o modelo simbólico, clásico o serial, que concibe la inteligencia como computación (*computare*) de símbolos, los que se pueden implementar físicamente en una máquina.

Allen Newell y Herbert Simon, fundadores en la década de los '50 del enfoque computacional al diseñar los primeros programas en IA, señalan en un clásico artículo publicado en 1976 que:

Los símbolos descansan en la raíz de la acción inteligente que es, desde luego, el tema principal de la inteligencia artificial. En cuanto a eso, es una cuestión central para toda la ciencia de la computación, ya que toda la información se procesa en computadoras en función de fines y medimos la inteligencia de un sistema por su capacidad para alcanzar objetivos establecidos frente a variaciones, dificultades y complejidades planteadas por el ámbito de la tarea que hay que realizar. (Newell y Simon, 1994, pág. 124)

En este mismo artículo, que selecciona Margaret Boden (1994), plantean su famosa hipótesis de los sistemas de símbolos físicos, según la cual:

Un sistema de símbolos físicos cuenta con los medios necesarios y suficientes para realizar actos de inteligencia general. Por 'necesario' queremos decir que cualquier sistema que exhiba inteligencia general probará con análisis que es un sistema de símbolos físicos. Por 'suficiente' queremos decir que cualquier sistema de símbolos físicos de tamaño suficiente puede organizarse además para exhibir inteligencia general. Por 'actos de inteligencia general' queremos indicar el mismo alcance de inteligencia que observamos en la actividad humana: que en cualquier situación real puede ocurrir un comportamiento adecuado para los fines del sistema y adaptado a las exigencias del medio ambiente dentro de ciertos límites de velocidad y complejidad. (Newell y Simon, 1994, pág. 128)

En relación con esta manera de concebir la cognición o la inteligencia, el neurobiólogo chileno Francisco Varela señala:

¿Qué significa exactamente decir que la cognición se puede definir como computación? Una computación es una operación que se lleva a cabo sobre símbolos, es decir, sobre elementos que representan aquello que designan. La noción clave es la de representación o 'intencionalidad', el término filosófico para aludir al 'acerca de'. El argumento cognitivista es que la conducta inteligente presupone la aptitud para representar el mundo como si fuera de ciertas maneras. Por lo tanto no podemos explicar la conducta cognitiva a menos que demos por sentado que un agente actúa representando rasgos relevantes de la

situación en que se halla. En la medida en que esta representación de una situación sea precisa, la conducta del agente tendrá éxito (siempre que todo lo demás permanezca igual). (Varela, 1997, pág. 64-65.)

## Capítulo 2 Perspectivas De La Ciencia Cognitiva

A principios del siglo XX, la inteligencia se convirtió en un tema de gran interés y estudio para la psicología, una ciencia emergente en ese momento, donde destacan los trabajos de Galton, Binet y Spearman, entre otros. Este periodo marcó un punto de inflexión significativo en la manera en que se comprendía la mente humana y sus capacidades, pues antes de esta época, el estudio de la mente y el comportamiento humano a menudo se mezclaba con la filosofía y la metafísica. Sin embargo, con la aparición de métodos experimentales y el uso de herramientas estadísticas, los psicólogos comenzaron a investigar de manera más rigurosa y sistemática los denominados procesos mentales, cambio que no sólo permitió un mejor entendimiento de la mente humana, sino que también permitió desarrollar un enfoque más científico para el estudio de la inteligencia humana.

Uno de los primeros en intentar medir el nivel intelectual de los seres humanos fue el polímata, antropólogo, estadístico y psicólogo inglés Francis Galton, quien propuso medir las funciones sensoriales y motoras como indicadores de la inteligencia, pues creía que los sentidos agudos y los reflejos rápidos estaban directamente relacionados con la capacidad intelectual de una persona. Con base en esta hipótesis, desarrolló diversos instrumentos y técnicas para medir características como la agudeza visual, la sensibilidad auditiva y los tiempos de reacción, aunque posteriormente se advirtió que estas mediciones no se correlacionaron significativamente con el rendimiento escolar u otros indicadores prácticos de la inteligencia (Davidoff, 1989). El psicólogo francés Alfred Binet, por otro lado, se enfocó en habilidades cognitivas como la atención, la memoria, el pensamiento lógico y la comprensión de las oraciones, aspectos fundamentales del funcionamiento mental que están más relacionados con el aprendizaje y la resolución de problemas. Basándose en sus investigaciones, desarrolló pruebas que permitían determinar la edad mental de un individuo, comparando su rendimiento con el de la media de personas de diferentes edades. Esta idea resultó revolucionaria, ya que permitía una evaluación más precisa de las capacidades cognitivas de un individuo en relación con su grupo de edad. Durante la década de 1920, el psicólogo inglés Charles Spearman propuso la teoría bifactorial de la inteligencia. Según esta teoría, la inteligencia se compone de un factor general único, conocido como *factor g*, y de factores específicos (libre de la influencia cultural). El factor *g* es una habilidad general que subyace a todas las capacidades

cognitivas de una persona, mientras que los factores específicos se refieren a habilidades o capacidades particulares. Sin embargo, años después, otros psicólogos cuestionaron la idea de un factor general único y propusieron que la inteligencia general está compuesta por una serie de capacidades específicas. Uno de los principales defensores de esta postura fue Thurstone, quien propuso el enfoque de múltiples factores, tales como la velocidad y precisión para resolver operaciones aritméticas básicas, la habilidad de comprensión de ideas, la memoria, el razonamiento y la percepción de relaciones espaciales. Desde una perspectiva diferente, Piaget, define la inteligencia como la forma de equilibrio hacia la cual tienden todas las estructuras cognitivas, condición fundamental para la supervivencia del organismo, ya que media las interacciones más complejas entre el individuo y su entorno. La concepción constructivista de Piaget, respecto a que la inteligencia se desarrolla a partir de coordinaciones sensoriomotoras recursivas, ha sido reafirmada en algunos enfoques más contemporáneos de la psicología y de la inteligencia artificial.

La profesora de filosofía y científica cognitiva inglesa Margaret Boden, plantea que la inteligencia no se puede reducir a una sola capacidad, sino que es un conjunto complejo de habilidades y capacidades que nos permiten procesar la información de diversas maneras. Así, la inteligencia artificial utiliza una amplia variedad de técnicas para resolver diferentes tipos de problemas (Boden, 2017), constituyendo un amplio espectro en lugar de una sola dimensión. Concebir la inteligencia como espectro supone que, en el caso de los seres humanos, encontramos diferentes habilidades y capacidades cognitivas, todas las cuales están interconectadas, conformando una red o sistema que se expresa de diferentes maneras en cada individuo. Por ejemplo, la capacidad de razonar no solo implica el pensamiento lógico y analítico, sino que también habilidades de razonamiento abstracto y conceptual. Este conjunto complejo de capacidades nos permite abordar y solucionar problemas desde múltiples perspectivas, haciendo uso de nuestra creatividad y pensamiento crítico. De igual manera, la habilidad de aprender no se limita a la adquisición de conocimientos académicos, sino que incluye la capacidad de aprender de la experiencia, de los errores y de las interacciones sociales. En el ámbito de la inteligencia artificial, estas mismas nociones se traducen en el desarrollo de algoritmos y modelos que puedan, de manera similar, abordar una amplia gama de problemas y situaciones. Desde el procesamiento del lenguaje natural

y la visión por computadora, hasta el aprendizaje automático y el razonamiento simbólico, cada técnica y enfoque tiene sus fortalezas y limitaciones, que al combinarse permiten la creación de sistemas más robustos y versátiles.

La inteligencia se constituyó en un tema central de la denominada ciencia cognitiva, donde surge la inteligencia artificial como una rama particular de la computación, tanto en su versión de la arquitectura computacional o simbólica, como en la de la arquitectura conexionista o de redes neuronales. El filósofo chileno e investigador en el ámbito de las neurociencias y las ciencias cognitivas Francisco Varela (2000) en su libro *El fenómeno de la vida*, nos plantea que la ciencia cognitiva, a pesar de que es una disciplina relativamente joven, surge a mediados del siglo XX, ha tenido un profundo impacto y se espera que continúe siendo así con el avance de la tecnología. Para Varela, el progreso en este campo se basa en apuestas conceptuales audaces, dado que la ciencia cognitiva incluye varias disciplinas que trabajan de manera integrada, como la neurociencia, la inteligencia artificial, la psicología cognitiva, la lingüística y la filosofía de la mente, las cuales procuran comprender los procesos cognitivos desde diferentes perspectivas. Los pioneros en este ámbito, inicialmente buscaron definir su campo con el término cibernética, sin embargo, este concepto se ha vuelto confuso y la disciplina ha evolucionado hacia una identidad más específica. Esta transición refleja la necesidad de la ciencia cognitiva de establecer un enfoque claro y delimitado para impulsar su desarrollo. La fase cibernética de la ciencia cognitiva, que se comenzó a desarrollar en la primera mitad del siglo XX, generó una amplia variedad de resultados bastante impresionantes y fecundos, entre los que destacan para Varela (2000), el uso de la lógica matemática para formalizar el funcionamiento del sistema nervioso, que permitió el desarrollo del modelo de las redes neuronales; el diseño e invención de los computadores digitales; la formulación de la teoría de sistemas y sus aplicaciones en los ámbitos de la ingeniería, la biología, la economía, la antropología y la terapia familiar; así como también la descripción y comprensión de los sistemas autoorganizativos. Durante esta etapa, se lograron avances significativos en el desarrollo de herramientas y conceptos que ahora son fundamentales en nuestra vida diaria, pero que antes de esa década formativa no existían. Estos avances fueron posibles gracias a la colaboración y el intercambio de conocimientos entre personas de diversos ámbitos que se reunieron en Estados

Unidos en la primera mitad del siglo pasado, especialmente en las denominadas Conferencias Macy.

Varela (2000) señala que los cibernéticos buscaban crear una ciencia objetiva de la mente, alejándose de las interpretaciones subjetivas que sostenían algunos psicólogos y filósofos de la época. Para ellos, la mente era un sistema de procesos mecánicos que podían ser descritos mediante fórmulas matemáticas. El trabajo de McCulloch y Pitts, *A logical calculus immanent in nervous activity*, del año 1943, donde postulan que la lógica es la clave para entender el funcionamiento del cerebro y su actividad mental, fue un hito en este campo, inspirando el desarrollo de numerosas investigaciones. McCulloch y Pitts, ven al cerebro como una computadora biológica donde las neuronas, al activarse o desactivarse, prenderse o apagarse, realizan operaciones binarias de tipo lógicas, que se pueden asimilar a los valores de verdad (V/F), de manera que el cerebro completo se puede concebir como una gran máquina de razonamiento, cuyo operar depende de unidades simples de procesamiento que corresponden a las neuronas. Este modelo o arquitectura se conoce como conexionismo, procesamiento distribuido en paralelo (PDP) o redes neuronales. A medida que la red opera recursivamente ajustando sus conexiones mediante mecanismos de retroalimentación, las interacciones se fortalecen o debilitan, modificando el comportamiento de la red, los resultados o respuestas que presenta. Los pesos de los nodos, el operar de la red como sistema, cambian dependiendo de las variables lógicas que subyacen a la dinámica que los regula. Así, la lógica se vuelve el fundamento del sopesar, del procesamiento que presenta la red neuronal o sistema conexionista.

Varela (2000) comenta que la computación ha recorrido un largo camino desde los tubos al vacío hasta los microchips, pero la arquitectura de von Neumann permanece como el pilar central. Este avance tecnológico sentó las bases para el cognitivismo, el paradigma dominante en el estudio de la mente a mediados del siglo XX. La segunda fase de la ciencia cognitiva se inició en 1956, con las emblemáticas reuniones de Cambridge y Dartmouth. En estos eventos, figuras clave como Simon, Chomsky, Minsky y McCarthy presentaron un nuevo paradigma: la mente como un sistema simbólico. Este enfoque, denominado computacional, serial o de procesamiento de la información, aunque se nutrió de investigaciones interdisciplinarias previas, estableció un sólido marco

conceptual o teórico que dominó durante los primeros años de desarrollo de la ciencia cognitiva. Al equiparar la cognición con el procesamiento de representaciones simbólicas, facilitó la conexión entre la psicología y las ciencias de la computación, concibiendo la cognición como manipulación de símbolos mediante ciertas normas o reglas, los cuales podían ser físicos o abstractos, representando aspectos del mundo real (Varela, 2000). La efectividad de un sistema cognitivo depende de su capacidad para utilizar estos símbolos para resolver problemas.

Las manifestaciones del cognitivismo nunca son más visibles que en la Inteligencia Artificial, que es la interpretación lógica de la hipótesis cognitivista. Con el correr de los años, se han logrado muchos avances teóricos y se han desarrollado aplicaciones tecnológicas interesantes dentro de esta orientación: sistemas expertos, robótica, procesamiento de imágenes. (Varela, 2000, pág. 188)

La hipótesis cognitiva encuentra su máxima expresión en la IA, donde los sistemas computacionales simulan procesos cognitivos humanos. Al mismo tiempo, esta hipótesis se aplica al estudio de la mente humana, buscando comprender cómo las representaciones mentales, subyacen a nuestras creencias, deseos y planes. La clave está en concebir estas representaciones como entidades formales manipuladas por procesos computacionales internos. Sin embargo, desde la filosofía, la psicología y la ciencia cognitiva, más recientemente, el concepto de representación ha sido bastante cuestionado, existiendo diversas versiones sobre el mismo, pues la noción clásica que nos remonta a Aristóteles y que se mantiene con diversos matices a lo largo de la historia, no resulta del todo coherente con los desarrollos epistemológicos que han tenido estas disciplinas, en cuanto no es posible establecer una relación directa y lineal entre el mundo y el sujeto o agente cognitivo, por lo que la idea de re-presentar el mundo, que éste es tal como el sujeto lo percibe, resulta bastante cuestionable. Esta mirada va a poner el foco en el significado que la experiencia tiene para el sujeto, surgiendo así críticas importantes al enfoque simbólico o computacional.

A diferencia del modelo serial o clásico, el conexionismo, que surgió en la misma época y que luego fue desacreditado, no opera con representaciones simbólicas, pues la respuesta adecuada, la conducta inteligente, emerge del operar recursivo de la red, mediante un algoritmo de aprendizaje basado en los planteamientos de Donald Hebb, el cual sugiere que la fuerza de conexión entre dos

neuronas aumenta si ambas se activan al mismo tiempo, interacciones que se van modificando constantemente mediante un proceso de retroalimentación que le permite al sistema ajustar su operar para alcanzar la respuesta inicialmente definida como objetivo. El enfoque computacional, en cuanto a su operar, es bastante estático y poco autónomo, pues requiere del programador para hacer correcciones a las reglas que lo rigen, en cambio, el modelo de redes neuronales se caracteriza por su carácter dinámico y autónomo para lograr su propósito. Vicente Raja, investigador en la filosofía de la ciencia cognitiva, señala que el “conexionismo, por tanto, sigue siendo un paradigma relevante dentro de las ciencias cognitivas y sigue proponiendo alternativas a los modelos clásicos tanto en psicología y neurociencia cognitiva como en inteligencia artificial.” (Raja, 2019).

Inspirado en la tradición filosófica existencialista, en los planteamientos de la neurociencia, el pragmatismo, la cibernética, la teoría de la cognición situada y distribuida, así como también en los trabajos de Maturana y Varela, el filósofo y teórico cognitivo Andy Clark, junto al filósofo analítico David Chalmers, proponen una tesis o teoría denominada *la mente extendida*, EMT (Extended Mind Thesis), según la cual la cognición no se limita al cerebro, sino que se distribuye en un sistema más amplio que incluye el cuerpo y el entorno. Esta visión desafía la concepción tradicional de la mente como un proceso puramente interno, proponiendo una cognición situada y corporizada (Cherubini 2011). En el artículo *The Extended Mind*, Clark y Chalmers (1998), siguiendo las ideas de James, Dewey y Vygotsky, sostienen que nuestro entorno juega un papel activo en nuestros procesos cognitivos, pues al realizar una tarea, el mundo exterior se integra de manera tan estrecha con nuestra mente que se convierte en un componente esencial de nuestro pensamiento, el entorno se convierte en un facilitador dinámico que influye y potencia nuestras capacidades mentales. El medio ofrece elementos que pueden ser herramientas, dispositivos y también personas, otros significativos que, mediante la interacción se convierten en componentes fundamentales del proceso cognitivo. Desde esta perspectiva, las mentes no están limitadas a lo que ocurre dentro de nuestro cerebro, sino que se extienden hacia el entorno que nos rodea. Un ejemplo claro de esta teoría se puede observar en el uso de dispositivos tecnológicos, como los teléfonos inteligentes. Cuando utilizamos un teléfono para realizar cálculos, recordar fechas o incluso planificar nuestra agenda, no solo estamos usando una herramienta, sino que estamos

incorporando ese dispositivo como parte de nuestra capacidad cognitiva. Sin el teléfono, sería mucho más difícil realizar esas tareas de manera eficiente, lo que demuestra cómo el entorno se convierte en una extensión de nuestra mente (Emilio et al., 2021).

En su obra monográfica de 2008, Andy Clark (Bietti, 2011) explora en profundidad las condiciones bajo las cuales los artefactos culturales y las tecnologías pueden integrarse de manera tan estrecha con la mente humana que se convierten en extensiones funcionales de nuestra cognición, por lo que sus ideas reafirman los planteamientos hechos por James, Dewey y Vygotsky a principios del siglo XX. Al igual como lo hiciera el antropólogo y teórico de la cibernética inglés Gregory Bateson en los años sesenta, Clark utiliza el ejemplo de una persona con discapacidad motriz que se apoya con un bastón para ilustrar cómo los artefactos pueden extender nuestros cuerpos de manera cognitiva. El bastón no es solo una herramienta, sino que se convierte en una parte integral de su cuerpo, permitiendo una nueva relación con el entorno. El cerebro se adapta a esta nueva configuración, creando un esquema corporal ampliado que incluye el bastón. Para Clark, la interacción constante entre nuestro cerebro, cuerpo y entorno nos permite desarrollar nuevas formas de pensar y resolver problemas, creando sistemas cognitivos que se adaptan a las situaciones reales.

## Capítulo 3 La Inteligencia Artificial

En el año 1950, el matemático, lógico, filósofo e informático teórico, Alan Turing, en su artículo *Maquinaria computacional e inteligencia*, publicado en la revista *Mind*, planteó la posibilidad de que las máquinas pudieran pensar, razonar o computar información de manera similar a un ser humano, condición que se podría evaluar permitiendo que una máquina participara en lo que denominó *juego de la imitación*, donde el computador tenía que simular un rol y darle pistas, verdaderas o falsas, a un interrogador. En este artículo, se refiere también al diseño de un computador digital que opera con variables discretas y al proceso que podría llevar a que la máquina compita intelectualmente con un ser humano, mediante un proceso de aprendizaje. Así, los desafíos de la computación quedaron planteados al comenzar la segunda mitad del siglo XX.

Unos años más tarde, en 1956, se considera que John McCarthy, un joven profesor de matemáticas e informático, acuña oficialmente el término *Inteligencia Artificial* en la Conferencia de Dartmouth, organizada por él mismo en colaboración con su amigo Marvin Minsky del MIT. El propósito de la actividad era explorar la idea de que cualquier aspecto del aprendizaje o de la inteligencia podría ser descrito de manera precisa para ser simulado por una máquina computacional. La conferencia es, principalmente, recordada no por sus logros, sino porque fue el primer encuentro de los cuatro hombres que se convertirían en los líderes de la Inteligencia Artificial en los Estados Unidos durante las siguientes dos décadas, McCarthy, Minsky, Newell y Simon (Dormido y Cruz García, 1989).

La evolución de la inteligencia artificial ha ido de la mano con los avances en robótica. Por lo mismo, la transición hacia la producción automatizada dio un paso significativo en 1938 con la creación del primer manipulador para pintura por parte de Devilviss. Este dispositivo, diseñado por H. Roselund y W. Pollard, representó una ruptura con los paradigmas productivos tradicionales, al incorporar la robótica a las líneas de ensamblaje, dominadas hasta entonces exclusivamente por la mano de obra humana. En la década de 1940, David A. Huffman realizó contribuciones fundamentales al desarrollar algoritmos de retroalimentación combinatoria. Estos algoritmos permitieron crear sistemas secuenciales con una rudimentaria capacidad de memoria, sentando las

bases para la construcción de funciones progresivas en sistemas automatizados. La formalización matemática de estos automatismos se sustentó en el álgebra de Boole y la teoría de autómatas finitos, proporcionando un marco teórico sólido para el desarrollo de sistemas inteligentes. Entre 1943 y 1946, un equipo liderado por Herman Goldstine, John Presper Eckert y John Mauchly, en el Laboratorio de Investigación Balística (BRL), logró un hito histórico al desarrollar la primera computadora digital electrónica: el ENIAC (Electronic Numerical Integrator and Computer). Este avance pionero sentó las bases de la computación moderna, transformando radicalmente los procesos industriales y empresariales. El concepto de robot, tal como lo conocemos hoy en día, debe gran parte de su desarrollo a la contribución de George Devol. Su patente de 1956, que describía un controlador magnético capaz de animar máquinas con una amplia gama de movimientos, transformó radicalmente la percepción de lo que una máquina podía hacer. A partir de este momento, el término *robot* se consolidó para referirse a dispositivos mecánicos programables y adaptables (Sánchez-Martín et al., 2007).

Durante la década siguiente, la IA experimentó un gran interés y desarrollo, creándose programas que pudieron resolver problemas matemáticos complejos, así como también implementar juegos de estrategia como el ajedrez. En los años siguientes, se crearon algoritmos más avanzados y se diseñaron técnicas de aprendizaje automático para que las máquinas puedan identificar patrones y tomar decisiones basadas en datos. Un momento clave en la evolución de la IA fue la creación de los sistemas expertos en la década de 1970, donde destacan las figuras de Newell y Simon. Estos sistemas empleaban reglas y conocimientos específicos para abordar problemas en áreas especializadas, como el diagnóstico médico o la asesoría legal. Aunque estos sistemas tuvieron cierto éxito, su principal limitación era que solo podían resolver problemas dentro de su campo de conocimiento, esto es, en un ámbito absolutamente abstracto, ajeno a la contingencia y a la praxis del vivir, dado que no tenían acceso directo al entorno y las complejidades que éste supone, por lo que adolecía de una efectiva y eficaz capacidad de adaptación, lo que se distingue como conocimiento de sentido común. Al no contar con mecanismos que le hicieran posible acoplarse fluidamente al ambiente, un sistema experto es incapaz de establecer coordinaciones sensoriales y motoras que son fundamentales para desenvolverse en un entorno real y complejo. En este sentido,

el insecto más vulgar tiene capacidades de adaptación que resultan mucho más adecuadas, inteligentes, que un sofisticado sistema experto.

Esta dificultad que presentaron los sistemas expertos, que se concebían como enormes archivadores digitales de información, llevó a cuestionar el modelo de inteligencia con el que estaba operando la arquitectura simbólica, serial o clásica de la IA, pues todo indicaba que no era suficiente con acumular y manipular datos explícitos a gran velocidad para poder responder de manera adecuada a las exigencias que supone la contingencia de la vida cotidiana. Los sistemas computacionales clásicos, al adolecer de cuerpo, no tienen la capacidad de sentir y actuar de manera rápida para cumplir con el objetivo básico de sobrevivir a las exigencias del entorno. Todo indicaba que la cognición natural, la conducta inteligente en el mundo real, no tiene tiempo para operar con archivadores, más aún cuando éstos suponen la utilización de costosos recursos que resultan ecológicamente imposibles.

Las críticas al modelo serial o clásico llevaron al resurgimiento del enfoque conexionista o de redes neuronales, cuya inspiración está en el operar del sistema nervioso y en el desarrollo que éste presenta como resultado de la interacción del organismo con su entorno, su capacidad para modificarse o ajustarse mediante mecanismos de retroalimentación, lo que se distingue como procesos de aprendizaje, cambios estructurales en la red que alteran sus conexiones y, por tanto, su funcionamiento. En estos sistemas, a diferencia del modelo computacional clásico, no hay unidades centrales de procesamiento (CPU) ni almacenes explícitos de memoria, sino que la *inteligencia* se encuentra distribuida en toda la red, la memoria consiste en patrones de acción recursivos, donde no hay un “cerebro” central. Esta concepción de la inteligencia supone que ésta emerge de las conexiones dinámicas que presenta el sistema como totalidad, de las coordinaciones sensoriomotoras que puede realizar de manera contingente a las entradas y salidas que se le presentan, por lo que se asemeja en su operar a la conducta que presentan los organismos en la vida cotidiana, siendo muchos más autónomos y flexibles que los sistemas expertos.

En los años siguientes, el enfoque de la inteligencia artificial viró hacia el aprendizaje automático, el reconocimiento de patrones y el procesamiento del lenguaje natural, el cual supone los procesos anteriores (Historia y evolución de la ia, s.f). En las últimas décadas con el crecimiento exponencial

de la cantidad de datos disponibles (Big Data), la IA experimentó cambios significativos con el desarrollo del aprendizaje profundo (Deep Learning). Con esta nueva tecnología, el desarrollo de nuevos chips, las redes neuronales profundas, más amplias y sofisticadas, pudieron procesar y analizar grandes volúmenes de datos de manera más eficiente y precisa. Estas redes están compuestas por múltiples capas de neuronas interconectadas, lo que les permite aprender y extraer características complejas de los datos. El aprendizaje profundo ha demostrado ser efectivo en áreas como el procesamiento del lenguaje natural, la visión por computadora y el reconocimiento de voz. Por ejemplo, en el procesamiento del lenguaje natural, las redes neuronales profundas pueden comprender y generar texto de manera más precisa y natural. Este enfoque ha revolucionado la IA al permitir que las máquinas aprendan y se adapten a través de grandes conjuntos de datos. Además, el crecimiento de la capacidad de cálculo y el acceso a la nube ha facilitado el entrenamiento y la implementación de modelos de aprendizaje profundo en diversos campos (Slegarraga, 2023).

En la última década, hemos sido testigos de un creciente uso, generalizado, de la IA en diversos ámbitos de la vida cotidiana. Poco a poco nos hemos ido habituando a la presencia de asistentes personales como Siri y Alexa, algoritmos que nos ofrecen recomendaciones en plataformas de streaming, como Netflix o Prime Video, así como también en aplicaciones de comercio electrónico, no parece una exageración señalar que la IA se ha vuelto parte importante de nuestras vidas, apoyándonos diariamente en la toma de decisiones. Uno de los avances más significativos en este último período ha sido el progreso en el procesamiento del lenguaje natural, que se ha manifestado en el desarrollo de modelos como BERT, GPT-2, GPT-3 y, más recientemente, GPT-4, que han demostrado una gran capacidad para comprender y generar texto de manera más precisa, coherente y sofisticada. Estos modelos de procesamiento del lenguaje natural han impulsado avances en áreas como la traducción automática, la generación de texto humanoide y el análisis de sentimientos, potenciando y ampliando los alcances de la IA en la vida cotidiana de las personas. (Sanabria, 2023).

Una de las aplicaciones actuales más conocidas de chatbot, bot o robot de charla o conversacional que surgieron en la década de los sesenta, es ChatGPT, lanzado originalmente en 2022 por la

empresa estadounidense OpenAI y que se encuentra ahora en su cuarta versión. Al consultarle ¿qué es la IA?, ChatGPT respondió lo siguiente:

La inteligencia artificial (IA) es un campo de la informática que busca crear sistemas y programas capaces de realizar tareas que normalmente requieren inteligencia humana. Estas tareas incluyen el reconocimiento de voz, la comprensión de lenguaje natural, la toma de decisiones, el análisis de datos, la visión por computadora y el aprendizaje. La IA se basa en algoritmos y modelos matemáticos que permiten a las máquinas aprender de datos, reconocer patrones y hacer predicciones o tomar decisiones basadas en esa información. En resumen, la IA busca simular procesos de pensamiento humano en las máquinas, pero aún tiene limitaciones y depende en gran medida de la calidad de los datos y los algoritmos que se utilizan. (OpenAI, 2024)

La clásica idea de que es posible crear máquinas que emulen el pensamiento y la inteligencia humana ha sido un tema históricamente debatible, tanto por las limitaciones propias de los enfoques propuestos, como por el concepto de pensamiento e inteligencia que subyace a éstos. En el ámbito de la filosofía, las críticas han estado presentes desde los inicios de la disciplina, siendo el filósofo estadounidense Hubert Lederer Dreyfus, uno de los más destacados. A partir de un enfoque fenomenológico-existencial, Dreyfus ha cuestionado las suposiciones fundamentales de la investigación en IA, cuestionando que ésta pueda, en algún momento, igualar o incluso superar a la inteligencia humana. En su libro *What Computers Can't Do* del año 1972, Dreyfus criticó la idea de que la mente humana pueda ser completamente simulada por una máquina. Según él (*Dreyfus E Os Limites Da I.A.*, 2015), el enfoque de la IA de buscar reglas formales para explicar el comportamiento humano es demasiado simplista. Los humanos, a diferencia de las máquinas, utilizamos el sentido común y hacemos juicios basados en el contexto, lo cual es difícil de formalizar. Esto hace que el enfoque inicial de la IA, basado en reglas precisas, se encontrara con un problema recurrente: para ejecutar cualquier tarea, se necesitaban reglas para aplicar esas reglas, y así sucesivamente. La falta de evidencia de que el cerebro funcione de manera estrictamente reglada sugiere que sería más acertado comparar el sistema nervioso con una computadora analógica que con una digital. La suposición de que la cognición humana puede reducirse a un

conjunto de reglas procedimentales ignora la naturaleza emergente y dinámica de los procesos mentales. Aunque la percepción sensorial proporciona los insumos básicos para la cognición, su interpretación e integración dependen de factores contextuales y experienciales que escapan a una formalización algorítmica. Esta limitación plantea un desafío fundamental para la IA, que debe trascender los modelos computacionales tradicionales para capturar la riqueza y complejidad de la mente humana. Una posible vía para resolver esta cuestión implicaría la construcción de una estructura jerárquica de reglas y situaciones, dotada de la suficiente flexibilidad para adaptarse a los cambios contextuales. No obstante, la naturaleza dinámica y contextualizada del conocimiento humano dificulta enormemente la tarea de aislar y representar de forma exhaustiva los elementos básicos que componen dicha estructura (*Dreyfus E Os Limites Da I.A.*, 2015). La capacidad humana para atribuir significado a la experiencia y construir un mundo simbólico es un fenómeno complejo y multifacético, influenciado por factores biológicos, psicológicos y culturales. La inteligencia humana, en contraste con la inteligencia artificial, se caracteriza por su flexibilidad, su capacidad para adaptarse a nuevas situaciones y su habilidad para generar ideas originales. Mientras que las máquinas procesan información de acuerdo a algoritmos preestablecidos, los seres humanos son capaces de trascender los límites de la lógica formal y crear nuevos significados. Sus ideas, que son conocidas por desafiar la corriente principal de la investigación en inteligencia artificial, han generado profundos debates y reflexiones en campos que van desde la filosofía hasta la ciencia cognitiva. A largo plazo, sus críticas plantean importantes cuestiones éticas y filosóficas sobre la finalidad y los límites del desarrollo de la inteligencia artificial, invitando a reflexionar sobre lo que realmente significa ser inteligente y consciente.

El profesor en filosofía John Searle, destacado por sus valiosas contribuciones a los campos de la filosofía del lenguaje, la mente y la conciencia, ilustra su crítica a la IA en su experimento mental conocido como la “*habitación china*”, el que es presentado por primera vez en su artículo *Minds, Prains, and programans* del año 1980. Searle, propone imaginar una habitación cerrada donde se encuentra una persona que no entiende nada de chino, dentro de la cual hay un conjunto de reglas, que sirven como un manual de instrucciones muy detalladas, escritas en un idioma que la persona sí entiende. Estas reglas explican cómo combinar símbolos chinos de manera específica, las cuales

le permiten elaborar frases y responder mensajes. La persona recibe una serie de símbolos chinos a través de una ranura en la puerta, y debe responder usando los mismos caracteres siguiendo las reglas del manual. Desde el exterior, las respuestas que salen de la habitación parecen ser las de una persona que habla chino con fluidez. Sin embargo, la persona dentro de la habitación no entiende nada de lo que está haciendo, simplemente sigue unas reglas sin comprender el significado de los símbolos, la dimensión semántica del mensaje. El conjunto de instrucciones o reglas opera como un *programa*, la ejecución de un algoritmo y los caracteres chinos se corresponden con la notación binaria usada por los computadores. Lo crucial del argumento es que la persona, a pesar de no comprender el idioma chino, es capaz de responder los mensajes. La habitación china es una metáfora muy ejemplificadora, que cuestiona la idea de que una máquina pueda llegar a comprender, a acceder a la semántica del lenguaje, tener una mente o conciencia similar a la humana simplemente procesando información. Searle sugiere que la inteligencia implica algo más que la capacidad de seguir reglas y manipular símbolos (Salcedo Albarán, 2004).

Más recientemente, el profesor de ciencias de la computación de la Universidad de California, Stuart Russell critica en su libro "Human Compatible: Artificial Intelligence and the Problem of Control" (2019), la falta de control sobre los sistemas de IA. Advierte que, si no se desarrollan sistemas de IA seguros y confiables, estos podrían tener consecuencias negativas para la sociedad.

Si se produce una explosión de inteligencia, y si todavía no hemos resuelto el problema de controlar máquinas con una inteligencia apenas superior a la humana (por ejemplo, si no podemos impedirles que realicen estas mejoras recursivas), entonces no nos quedaría tiempo para resolver el problema de control y el juego se terminaría. (Russell, 2019, pág. 144)

Russell (2019) destaca la importancia crucial de desarrollar sistemas de IA que sean seguros y confiables. Estos sistemas deben estar diseñados de tal manera que puedan ser monitorizados y controlados correctamente para evitar que funcionen de manera autónoma sin restricciones. Sin un control eficaz, las IA podrían actuar de formas impredecibles y potencialmente peligrosas. Este es un punto que genera mucha preocupación entre los expertos en el campo, ya que la inteligencia artificial tiene el potencial de afectar diversos aspectos de la vida cotidiana, desde la economía, la

salud, la educación e incluso hasta la seguridad nacional. Otra advertencia clave que Russell menciona en su libro es el riesgo asociado con una *explosión de inteligencia*, un concepto que se refiere al rápido y desmesurado aumento de la capacidad intelectual de las máquinas. En este escenario hipotético, las máquinas podrían empezar a mejorar sus propias capacidades de manera recursiva, sin intervención de un operador. Esta mejora recursiva podría escapar rápidamente de nuestro control, llevando eventualmente a resultados catastróficos. El autor señala que, si llegásemos a un punto en el que las máquinas superaran nuestras capacidades sin haber resuelto previamente el problema del control, nos enfrentaríamos a una situación extremadamente peligrosa, pues sería casi imposible detener esta progresión vertiginosa, y *el juego se terminaría* como menciona en su libro, aludiendo a un posible punto de no retorno donde el destino de la humanidad quedaría en manos de las máquinas. Si bien lo último suena muy a ciencia ficción, los sistemas de IA van mejorando sus capacidades más rápido de lo que los científicos han podido predecir o especular.

Nick Bostrom, un destacado filósofo de la prestigiosa Universidad de Oxford, ha explorado en profundidad el fascinante y complejo tema de la IA en su libro *Superinteligencia: Caminos, Peligros, Estrategias*. En esta obra, Bostrom (2016) presenta una serie de argumentos y fundamentos que sugieren que la IA tiene el potencial de eventualmente superar la inteligencia humana. Este posible avance podría llevarnos a un futuro incierto, aunque también lleno de oportunidades y beneficios potenciales. Bostrom, subraya que la superación de la inteligencia humana por parte de la IA no es un simple salto tecnológico, sino un evento de enorme trascendencia. En donde imagina un escenario donde la IA no solo iguale nuestra capacidad de pensamiento y razonamiento, sino que la supere en múltiples dimensiones. Por ejemplo, las mejoras en inteligencia biológica o en organización podrían aumentar nuestra capacidad para anticipar riesgos y diseñar una superinteligencia artificial segura y beneficiosa. Esta perspectiva nos invita a contemplar un panorama donde las máquinas puedan resolver problemas hasta ahora irresolubles, innovar de maneras inimaginables y transformar radicalmente todos los sectores de la sociedad, desde la medicina hasta la educación.

El cofundador de Microsoft y magnate empresarial Bill Gates, en su blog *GatesNotes* (2023) señala que es comprensible que la transformación digital genere inquietud, la IA nos plantea nuevos interrogantes. Gates señala que cada revolución tecnológica ha suscitado debates, en donde la experiencia nos enseña que, si bien los desafíos son reales, la humanidad ha demostrado una gran capacidad para adaptarse y aprovechar las oportunidades que surgen. Declara que la IA es la innovación más transformadora que veremos en nuestras vidas, en donde recalca la magnitud del cambio que traerá a diversos aspectos de nuestra existencia. Un aspecto crucial que Gates destacó es la necesidad de un debate público informado y saludable, lo cual implica una responsabilidad compartida por parte de educadores, gobiernos, empresas y medios de comunicación para garantizar que la sociedad en su conjunto tenga una comprensión clara de qué es la IA, cómo funciona y qué implicaciones tiene. Para lograr un debate constructivo y productivo, es indispensable fomentar la educación y la concientización en torno a estas nuevas tecnologías. Esto no solo ayudará a disipar temores infundados, sino que también permitirá a las personas tomar decisiones informadas y participar activamente en la configuración del futuro tecnológico. Gates destaca que en los próximos cinco a diez años, el software impulsado por la IA cumplirá finalmente la promesa de revolucionar la forma en que las personas enseñan y aprenden. La misma será capaz de conocer los intereses y el estilo de aprendizaje de cada individuo para adaptar el contenido y mantenerlo interesado. También medirá la comprensión, detectará cuando se pierda el interés y comprenderá a qué tipo de motivación responde, proporcionando retroalimentación inmediata.

Para el investigador de las ciencias de la computación Philip Agre, la filosofía tiene un papel fundamental al momento de hablar de IA, pues para él “la IA es filosofía desde sus fundamentos” (2005, p. 155). El autor se pregunta (Masís, 2023) si es posible encontrar una implementación tecnológica que esté completamente libre de los debates filosóficos sobre la mente, el mundo y el conocimiento, argumentando que tal desafío no es posible y criticando a quienes desarrollan la IA por ignorar estos debates fundamentales. Al intentar *purificar* los programas de IA de cualquier influencia lingüística, se pierde la conexión con la historia y la cultura, creando una falsa ilusión de objetividad. Los profesionales de la computación suelen pensar que al utilizar fórmulas matemáticas exactas, pueden eliminar las ambigüedades del lenguaje natural. Agre, en cambio,

aboga por una práctica más reflexiva, la *práctica técnica crítica*. Esta implica examinar constantemente cómo nuestras acciones están moldeadas por factores históricos y sociales, y buscar formas de mejorar nuestra labor a partir de estas reflexiones. La propuesta de Agre representa un llamado a trascender la mera ejecución de técnicas y procedimientos, invitando a los profesionales a convertirse en agentes activos de transformación, que no solo dominen herramientas, sino que comprenden el contexto en el que estas se aplican. Agre aboga por una práctica técnica que reconozca y explore el papel fundamental de las metáforas en la construcción y comprensión de los sistemas artificiales. Al considerar tanto la realidad física como la construcción discursiva, su enfoque busca comprender cómo las metáforas que motivan la implementación tecnológica pueden convertirse en objetos de investigación científica. De esta manera, la IA no solo se limita a resolver problemas prácticos, sino que se convierte en una herramienta para reflexionar sobre la naturaleza del conocimiento y la cognición humana. Al expandir la comprensión de cómo funcionan los sistemas artificiales, Agre sugiere que la IA puede convertirse en una forma de escuchar y aprender de la realidad (Masís, 2023).

## Capítulo 4 La Inteligencia Artificial Generativa y la Educación

La inteligencia artificial generativa (IAG) representa una rama emergente y sumamente prometedora dentro del amplio campo de la inteligencia artificial. Este tipo de tecnología se enfoca en la creación de nuevos contenidos utilizando datos existentes como base, pero a diferencia de los sistemas tradicionales de IA, que han sido diseñados principalmente para llevar a cabo tareas específicas como la clasificación de datos o la predicción de resultados, la IAG busca ir mucho más allá. En lugar de simplemente analizar información y categorizarla, la IAG tiene la habilidad de generar respuestas originales y creativas, aprendiendo de los usuarios que interactúan con ella. Esto puede abarcar una variedad de formatos, tales como texto, imágenes, música e incluso código de programación. Por ejemplo, en el ámbito de la escritura, los modelos de IAG pueden redactar artículos, cuentos o ensayos completos, imitando estilos literarios diversos (Polvorosa, s. f.).

La idea de IAG no fue algo novedoso a los tiempos actuales, tuvo sus primeros indicios en el año 1966 con la aparición de *Eliza*, el cual fue un hito fundamental en la historia de la computación. Este era un célebre programa de IA desarrollado por Joseph Weizenbaum en el MIT. Eliza fue un chatbot pionero, diseñado para simular conversaciones psicoterapéuticas, demostró cómo un programa relativamente simple podía generar la ilusión de comprensión y empatía, influyendo significativamente en el desarrollo de la IA conversacional y en la percepción pública sobre las capacidades y limitaciones de las máquinas (Serrano & Moreno-García, 2024). Posteriormente, surge ALICE, el cual fue un precursor de los chatbots modernos, demostrando que era posible crear sistemas capaces de mantener conversaciones coherentes con los usuarios. Sin embargo, su enfoque basado en reglas contrasta con los modelos de lenguaje de gran escala utilizados en la actualidad, los cuales han demostrado una capacidad superior para comprender y generar lenguaje natural. ALICE, a pesar de su simplicidad, sirvió como un banco de pruebas para explorar las posibilidades y los desafíos de la interacción humano-máquina. Los chatbots han experimentado un notable avance en las últimas décadas, gracias a la mejora constante de los algoritmos y el aumento de la potencia computacional, en donde la aparición de ChatGPT marca un antes y un después. Su capacidad para procesar y generar texto de manera natural y contextualizada, gracias

a su entrenamiento en vastos conjuntos de datos, lo posiciona como un hito en el campo de la IAG conversacional, con un potencial transformador en la educación.

La gran irrupción de las IAG en el escenario tecnológico contemporáneo está desafiando los fundamentos epistemológicos de nuestra sociedad (Gutiérrez, 2023). La capacidad de estas tecnologías para generar y procesar información a gran escala está reconfigurando los procesos de producción y consumo de conocimiento. En el marco de la Cuarta Revolución Industrial, concepto acuñado por Klaus Schwab, asistimos a una transformación radical de los paradigmas de aprendizaje, de investigación y de innovación, lo que plantea nuevas interrogantes sobre la naturaleza de la inteligencia y el papel del ser humano en una era cada vez más digitalizada (Gutiérrez, 2023). La proliferación de la IAG suscita interrogantes éticos fundamentales, particularmente en lo que respecta a la generación automatizada de contenidos. La versatilidad de estas tecnologías, que abarcan desde tareas simples como la generación de textos breves hasta la elaboración de contenidos complejos y especializados, plantea un desafío multifacético para las sociedades contemporáneas. Los escenarios de aplicación son vastos y variados, lo que exige una reflexión crítica sobre las implicaciones de esta revolución tecnológica. Lo anterior, hace que se ponga en tela de juicio la autoría, originalidad y fiabilidad de dichos contenidos. Gutiérrez (2023) menciona que la abogada y experta en IA, Micaela Mantegna, sostiene que, si bien la originalidad es el resultado de un proceso creativo en el cual el autor, a partir de su subjetividad, realiza una observación minuciosa y comparativa con obras anteriores, dando lugar a una producción única. Si la IA puede simular estos estilos y técnicas ¿qué distingue realmente una obra creada por un humano? La IA podría socavar la idea de que la originalidad es inherentemente humana, al demostrar que la creatividad puede surgir de algoritmos y datos. Esto difumina la noción de autoría, ya que es difícil determinar en qué medida una obra generada por IA es producto de la originalidad. Otro problema ocurre cuando los chatbots generativos, como ChatGPT, funcionan sobre la base de modelos probabilísticos que predicen la siguiente palabra en una secuencia textual a partir de análisis estadísticos aprendidos de grandes cantidades de datos (libros, páginas web, artículos, etc.). Este enfoque, aunque permite generar texto coherente y fluido, conlleva la posibilidad de que el modelo produzca respuestas incorrectas o incoherentes, conocidas como *alucinaciones*. Estas ocurren cuando el modelo extrapola patrones más allá de su dominio de conocimiento o cuando la

ambigüedad de la consulta dificulta la identificación del contexto adecuado. Es fundamental reconocer que las respuestas generadas por estos sistemas son el resultado de un proceso probabilístico y, por tanto, no deben tomarse como verdades absolutas (Van Vaerenbergh, 2024). Además, existe el riesgo de que las decisiones automatizadas perpetúen sesgos y desigualdades si los sistemas de IA no son desarrollados y supervisados de manera adecuada. Un eventual riesgo en el uso de esta tecnología, que también está presente en el uso generalizado de internet, es fomentar la práctica de *copiar y pegar* para elaborar trabajos o ensayos. Algunos estudiantes pueden aprovechar la facilidad de acceso a la información en línea para evitar el esfuerzo de leer y comprender el contenido, lo cual implica dejar de aprender y de hacer uso del pensamiento crítico. Una deriva de lo anterior lo constituye el plagio, ya que los estudiantes pueden encontrar ensayos completos en línea y utilizarlos como propios, sin realizar una investigación o redacción original. Esto puede limitar su desarrollo de habilidades de investigación, comprensión y expresión escrita, dificultando el sentido que tiene la evaluación. Se suma a que se vuelve importante para las personas que utilizan estas herramientas, dado que deben mantener un pensamiento crítico al interactuar con chatbots generativos y verificar la información que reciben, especialmente si se trata de temas que requieren una alta precisión, como datos científicos o cuestiones filosóficas.

El empresario e informático estadounidense, Kai-Fu Lee, considerado uno de los mayores expertos y exponentes en IA, en su libro *Superpotencias de la inteligencia artificial* (2020), aborda varios temas con respecto al uso de esta tecnología, en donde toca temas relevantes y positivos en su implementación en la educación. Lee (2020) destaca que estas nuevas tecnologías permiten a los docentes optimizar el proceso de corrección, delegando en la IA las tareas más rutinarias y liberando tiempo para centrarse en aspectos más cualitativos de la evaluación, como la coherencia argumentativa y la originalidad de las ideas. Lee subraya el potencial de la IA para personalizar la educación. Una de esas ideas, es crear perfiles de estudiantes generados por IA, esto no solo identifican las dificultades de aprendizaje, sino que también permiten una intervención temprana a través de la notificación a los padres. Al analizar datos sobre la participación y el estado emocional de los estudiantes, estas herramientas facilitan la selección de tutores adecuados y optimizan el proceso de enseñanza-aprendizaje. Este enfoque, que combina la tecnología con la tutoría personalizada, ha demostrado ser eficaz en plataformas que conectan a estudiantes de diferentes

países (Lee 2020). Las propuestas de Lee se inscriben en un contexto de innovación educativa en el que China se destaca como pionero. Esta tendencia anticipa un futuro donde la competencia global impulsará la constante evolución de estas tecnologías, permitiendo su adopción y adaptación a las particularidades de cada sistema educativo.

El historiador y escritor israelí Yuval Noah Harari, en su libro *21 lecciones para el siglo XXI* (2018) nos advierte lo mismo que Kai-Fu Lee que, en el pasado muchas escuelas se centraban en que los estudiantes aprendieran de memoria la información. Esto se debía a que en aquel entonces la información era escasa y difícil de obtener. Además, el flujo de información existente a menudo estaba bloqueado por la censura. En contraste, en el siglo XXI estamos rodeados de una cantidad masiva de información, y los censores o canales de difusión ya no intentan bloquearla. En cambio, su enfoque se ha desviado hacia la difusión de desinformación o la distracción con temas insignificantes. Sumado a ello, es preocupante lo fácil que resulta inundar a las personas con informes contradictorios y pistas falsas. Todo lo anterior accede un estudiante a un solo *click*, por lo mismo Harari destaca la importancia del docente en el aula. Harari (2018, Pág. 290) sostiene que la sobreabundancia de información hace que el pensamiento crítico sea una habilidad esencial. Los profesores deben fomentar en sus estudiantes la capacidad de cuestionar, evaluar y conectar ideas, en lugar de limitarse a proporcionarles datos aislados.

Dianne Adlawan (2023) ejecutiva de marketing en ClassPoint, menciona dos riesgos aparentes en la implementación de la IA en la educación. Una de las principales limitaciones de la implementación de algoritmos de IA en la educación es la potencial deshumanización del proceso de enseñanza-aprendizaje. La interacción personalizada y la retroalimentación cualitativa que un docente proporciona son elementos fundamentales para el desarrollo integral del estudiante, los cuales pueden verse comprometidos al reemplazarlos por sistemas algorítmicos, debido a que carecen de la capacidad de adaptar el contenido y el ritmo de las lecciones de manera verdaderamente individualizada. Otro aspecto, es que la proliferación de herramientas educativas basadas en IA en los entornos escolares plantea el riesgo de una excesiva dependencia tecnológica tanto por parte de docentes como de estudiantes. Esta situación podría conllevar la desvalorización de metodologías pedagógicas tradicionales, fundamentales para el desarrollo de habilidades cognitivas superiores como el pensamiento crítico y la resolución de problemas.

La personalización del aprendizaje a través de la IA puede llevar a una experiencia educativa más individualizada, pero también puede aislar a los estudiantes y debilitar los lazos sociales que se forman en el aula. La interacción social es fundamental para el desarrollo cognitivo y emocional de los estudiantes, ya que les permite aprender a trabajar en equipo, a negociar y a resolver conflictos. Un ejemplo de lo anterior, es que las interacciones sociales son fundamentales para el desarrollo de la inteligencia emocional, que incluye la capacidad de reconocer y gestionar las propias emociones, así como las de los demás. La falta de esta puede limitar el desarrollo de estas habilidades, lo que a su vez puede afectar las relaciones interpersonales y el bienestar emocional a largo plazo. Es importante encontrar un equilibrio entre el uso de herramientas educativas basadas en IA y los métodos de enseñanza tradicionales. La IA puede ser una herramienta valiosa para complementar la enseñanza, pero no debe reemplazarla por completo.

Es esencial que estas decisiones consideren una amplia gama de factores, incluyendo los beneficios potenciales y los riesgos asociados. Por esto mismo, la UNESCO (2021) establece tres pilares para integrar la IA en la educación: aprender con IA, aprender sobre IA y aprender para la IA. La primera se enfoca en la utilización de herramientas de IA para personalizar y optimizar los procesos de enseñanza y aprendizaje. Esto implica el uso de sistemas de tutoría inteligente, software de evaluación adaptativa y otras herramientas tecnológicas para adaptar el aprendizaje a las necesidades individuales de cada estudiante; en la segunda, se busca equipar a los estudiantes con conocimientos básicos sobre los principios y aplicaciones de la IA. Se trata de comprender cómo funciona la IA, sus diversas aplicaciones en diferentes campos y su impacto en la sociedad; la tercera área se centra en el desarrollo de habilidades transversales que permitirán a los estudiantes adaptarse a un mundo en constante cambio donde la IA juega un papel cada vez más importante. Estas habilidades incluyen el pensamiento crítico, la resolución de problemas, la creatividad, la colaboración y la capacidad de aprender de manera autónoma.

La periodista de BioBio Chile Sara Jerez (2024), explica en su artículo web que OpenAI ha presentado su último modelo de lenguaje grande (LLM), el OpenAI o1, el cual ha demostrado un rendimiento excepcional en pruebas científicas. Según un artículo publicado en Nature, el o1, anteriormente conocido como *Proyecto Strawberry*, ha superado a científicos con doctorado en la resolución de problemas complejos. Este modelo de última generación ha sido entrenado para

realizar tareas que requieren un alto nivel de razonamiento y comprensión, lo que lo convierte en una herramienta prometedora para diversas aplicaciones en el ámbito científico y tecnológico. La evaluación Graduate-Level Google-Proof Q&A Benchmark (GPQA), compuesta por 448 preguntas de opción múltiple diseñadas por expertos en biología, física y química, representa un desafío significativo para cualquier modelo de lenguaje. Si bien ChatGPT logró un puntaje del 39% en 2023, el modelo o1 ha superado esta marca de manera considerable, alcanzando un 78% de precisión. Este resultado no solo supera el desempeño de los académicos en la prueba general, sino que también revela una capacidad excepcional del o1 en el área de física, donde ha obtenido una puntuación cercana al 93% (Jerez, 2024). Al superar a expertos en un examen tan exigente, el o1 demuestra un nivel de comprensión y razonamiento que anteriormente se consideraba exclusivo de la inteligencia humana. Este avance significativo sugiere que los modelos de lenguaje grande están cada vez más cerca de alcanzar un nivel de desempeño comparable al de los científicos en diversas disciplinas.

Lo anterior, se contrasta con una reciente investigación llevada a cabo por la Universidad de Oxford y el Instituto Reuters (Vega, 2024), basada en una encuesta a 12.000 individuos de seis países. Esta revela una notable brecha en el conocimiento y la adopción de la IA por parte del público en general. Los resultados indican que un porcentaje significativo de los encuestados, incluyendo hasta un 30% de los residentes en el Reino Unido, no están familiarizados con plataformas de IAG como ChatGPT y Microsoft Copilot. Además, el estudio destaca un bajo nivel de uso cotidiano de estas tecnologías, con solo un 2% de los participantes reportando su utilización diaria. En una reciente entrevista con la BBC, Richard Fletcher, director de investigación del Instituto Reuters y autor principal del estudio, señaló una notable discrepancia entre la percepción de las empresas tecnológicas y la realidad del interés público en la inteligencia artificial generativa. Según Fletcher, a pesar de la gran inversión y promoción de estas tecnologías, el público en general no muestra un interés particular en ellas (Vega, 2024).

Un reciente estudio del IBM Institute for Business Value, presentado en el IBM Forum (Lesta, 2024), subraya la necesidad imperiosa de un cambio cultural para aprovechar al máximo el potencial de la IAG. La encuesta, realizada a más de 600 líderes empresariales, revela que el 57% de los CEOs latinoamericanos considera que superar la resistencia al cambio es más crucial que

abordar los desafíos técnicos asociados a la adopción de la IA. Sin embargo, el 41% reconoce que la falta de comprensión de los empleados acerca de las decisiones estratégicas relacionadas con la IA representa un obstáculo significativo para este proceso de transformación. El estudio puso de manifiesto la creciente brecha entre la velocidad del desarrollo tecnológico y la capacidad de las organizaciones para adaptar a sus empleados. Si bien la mayoría de los CEOs latinoamericanos reconocen la importancia de la capacitación, la presión por adoptar rápidamente las nuevas tecnologías, especialmente la IAG, ha llevado a muchas empresas a priorizar la implementación tecnológica sobre el desarrollo del capital humano (Lesta. 2024). Lo anterior, ejemplifica la importancia de educarse sobre IA e IAG, dado que su uso se está masificando y sus avances son cada vez más rápidos.

Dado la importancia de trabajar en la implementación y uso correcto de las herramientas digitales, es que el Consejo Empresarial de Ética de la IA de la UNESCO, coliderado por Microsoft y Telefónica, organizó un panel en el marco de la Segunda Cumbre Ministerial (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, 2024). Este encuentro regional, coorganizado por CAF y AGESIC, reunió a líderes empresariales para discutir el papel fundamental del sector privado en el desarrollo de la IA en América Latina y el Caribe. Los participantes coincidieron en la necesidad de invertir en capital humano para garantizar una adopción exitosa y responsable de esta tecnología. Un panel de expertos de reconocido prestigio compartió sus visiones sobre el futuro de la IA en la región y las acciones necesarias para aprovechar al máximo su potencial, destacando la importancia de fomentar la colaboración entre los diferentes actores del ecosistema de innovación para construir un futuro digital más inclusivo y equitativo. Como resultado de este evento, se lanzó un llamado a la acción para promover la capacitación y la formación en habilidades digitales, con el objetivo de preparar a la fuerza laboral para los desafíos y oportunidades que plantea la Cuarta Revolución Industrial (UNESCO, 2024).

En un esfuerzo por comprender el impacto de la IAG en el sector educativo, la Editorial Santillana (2024) realizó una encuesta a docentes, directivos, padres de familia y estudiantes durante el mes de marzo del año 2024. Los datos recopilados permitieron identificar tendencias y patrones significativos en las opiniones y expectativas de los participantes. La encuesta contó con una amplia participación de la comunidad educativa, obteniendo un total de 1320 respuestas. Los

participantes se distribuyeron principalmente en los grupos etarios de 40 a 49 años (35%) y de 50 a 59 años (31%) entre docentes y directivos, mientras que entre los estudiantes predominó el grupo de 13 años (41%). Los hallazgos indican un alto potencial de la IAG para revolucionar los procesos educativos. La personalización del aprendizaje y el fomento de la creatividad emergen como las principales ventajas percibidas. Además, el estudio evidencia un uso significativo de herramientas de IAG, con ChatGPT a la cabeza. El estudio revela una creciente adopción de la IAG en los niveles educativos medios y superiores. Si bien se valoran los beneficios de esta tecnología en términos de optimización de recursos y personalización del aprendizaje, surgen preocupaciones sobre su posible impacto negativo en las habilidades cognitivas y el incremento de la dependencia tecnológica. Los hallazgos sugieren que los docentes están explorando diversas estrategias para integrar la IAG en sus prácticas pedagógicas, entre las que destacan la evaluación formativa, la creación de marcos normativos y la tutoría individualizada. Si bien se reconoce el potencial de la IAG para mejorar los resultados académicos, persisten interrogantes sobre su impacto en la proactividad estudiantil. La colaboración con las familias, aunque limitada, se centra principalmente en el intercambio de experiencias e ideas. Desde el punto de vista de la evaluación, el impacto de la IAG en los procesos de aprendizaje se encuentra en una fase inicial, predominando métodos cualitativos. Los docentes, a pesar de su creciente adopción de herramientas IAG, enfrentan desafíos como la desigualdad digital, la falta de soporte institucional y resistencias culturales. Si bien reconocen beneficios como la optimización de recursos y la personalización del aprendizaje, también expresan preocupaciones sobre la dependencia tecnológica y la integridad académica (Editorial Santillana, 2024).

Con el objetivo de comprender la percepción y el aprovechamiento de la IA en el ámbito académico, la Vicerrectoría Académica de la UC (Pontificia Universidad Católica de Chile, 2024), ha llevado a cabo dos encuestas consecutivas entre su cuerpo docente. La segunda, realizada en abril de 2024, ha arrojado datos preliminares sobre la evolución de las actitudes y prácticas relacionadas con la IA. Estos resultados, junto con los obtenidos en la primera encuesta (agosto 2023), sirven como base para el diseño de estrategias institucionales que promuevan el uso responsable y efectivo de estas tecnologías en los procesos de enseñanza y aprendizaje. Los resultados de la segunda encuesta, en el que se contó con la participación de cerca de 800 docentes

de todas las áreas de la Universidad, revelan un incremento significativo en la percepción positiva de su potencial pedagógico entre el profesorado, con un 71% reportando un aumento en su valoración respecto al año anterior. Sin embargo, persiste una proporción considerable (26%) que expresa incertidumbre sobre su contribución. En cuanto a las aplicaciones de la IA, se destaca su uso en la generación de recursos didácticos y en la realización de actividades en clase. No obstante, se identifican preocupaciones significativas entre el profesorado respecto a los potenciales riesgos éticos asociados al uso de la inteligencia artificial en la educación. Un alto porcentaje de encuestados identifica la integridad académica como la principal amenaza, destacando el riesgo de plagio y el uso indebido de estas herramientas, siendo este del 91%. Asimismo, se expresan inquietudes sobre la privacidad de los datos, la veracidad de la información generada por la IA y la posible dependencia excesiva de estas tecnologías, que podría perjudicar la creatividad y la profundidad del aprendizaje (UC, 2024).

Si bien las investigaciones de la editorial Santillana y la UC resaltan el valor de la IA como herramienta pedagógica, es fundamental abordar las preocupaciones relacionadas con su uso indebido. La proliferación de herramientas de generación de texto plantea interrogantes sobre la autenticidad y originalidad de los trabajos académicos, generando un manto de dudas entre el profesorado. En este contexto, resulta esencial establecer un marco regulatorio que promueva el uso responsable de la IA y fomente el desarrollo de habilidades de pensamiento crítico en los estudiantes, en donde la Filosofía tiene un papel fundamental. Adicionalmente, la promoción de iniciativas educativas orientadas a la IA y la IAG permite desarrollar una alfabetización digital más amplia, acercando estas tecnologías al público en general. Al difundir conocimientos sobre los principios y usos de la IA, se contribuye a democratizar el acceso a la información y a fomentar una ciudadanía digital más informada. Ejemplo de ello es un taller realizado por la Universidad Andrés Bello llamado “Inteligencia Artificial Generativa como Herramienta Pedagógica: Estrategias y Aplicaciones para Docentes de Educación Superior” (2024), este explora y enseña sobre el uso de la IAG para quienes no están familiarizados con ella. Lo anterior, fomenta una ciudadanía digital más informada y capaz de navegar en un mundo cada vez más complejo y lleno de información, y de esa forma, enfrentar los desafíos del futuro en un entorno digital en constante evolución.

## Conclusión

La acción de vivir supone necesariamente una interacción con el medio, la cual, con mayor o menor conciencia, conlleva una transformación de éste y, por tanto, de los sujetos que habitan en él. En su momento, la invención del cuchillo, de la rueda, del motor a vapor, del teléfono, del automóvil, del avión, entre muchos otros, cambiaron significativamente el modo de vivir, lo que era o no posible hacer en un momento dado. Las creaciones humanas configuran nuevas realidades, inauguran nuevos mundos, que son impensables antes que las nuevas tecnologías aparezcan. El desarrollo de la electrónica y la computación han permitido un rápido avance de la llamada inteligencia artificial, se han masificado las máquinas y dispositivos capaces de asistirnos, de potenciar nuestro pensamiento, de ayudarnos a manejar la información y tomar decisiones. Esta tecnología ha pasado a formar parte de nuestra vida cotidiana de manera tan rápida que, casi no lo hemos advertido, normalizando ampliamente su uso en diversos ámbitos, incluyendo, ciertamente, el educacional.

Este trabajo de tesis surge a partir de la inquietud de tomar conciencia del efecto que ha tenido, tiene y, eventualmente, tendrá, el creciente desarrollo de la IA en el ámbito educacional, cuál puede ser su impacto, sus potenciales beneficios y los riesgos que conlleva, así como también cuáles son las políticas que se han adoptado o se podrían adoptar para su mejor desarrollo. En este ejercicio, se recorrió de manera muy somera algo de la historia del concepto mismo de inteligencia y de representación, conceptos que son clásicos en la historia de la filosofía, de la psicología y de la ciencia cognitiva, más recientemente. Se repasó también algo de la historia de la IA misma, las conceptualizaciones principales que se han tenido de ella y algunas de las críticas que ha generado, para finalmente, mostrar algo de los avances y aplicaciones que ha tenido en los últimos años.

Es realmente fascinante y atractivo considerar cómo los tiempos antiguos de la Grecia clásica, donde pensadores como Sócrates, Platón y Aristóteles dejaron una huella indeleble en la historia del pensamiento humano, cuyo devenir haría posible la emergencia de la IA. Estas figuras no solo exploraron las profundidades de la filosofía, sino que también sentaron las bases para el razonamiento lógico y el desarrollo del conocimiento en el que se apoya nuestra cultura occidental, incluyendo a la IA moderna. La mayeutica de Sócrates, por ejemplo, en donde la idea de cuestionar

y reflexionar sobre la información es un aspecto esencial, sigue vigente como metodología didáctica en educación y también, metafóricamente, en el funcionamiento de la IA, que logra sus objetivos mediante procesos de retroalimentación. Además, la diversidad de perspectivas filosóficas y científicas, representada por pensadores como Descartes, Leibniz, Newell, Simon, Turing, Searle, Varela, entre otros, han dejado un legado intelectual que ha trascendido las barreras temporales, ejerciendo una influencia profunda y duradera en la evolución de la IA. A medida que avanzamos hacia un futuro donde la IA juega un papel cada vez más significativo en nuestras vidas, es crucial recordar y valorar sus raíces que han moldeado nuestra comprensión del pensamiento y la inteligencia.

Así como hay que apreciar el valor histórico y el crecimiento que ha tenido, también hay que reconocer las limitaciones de la IA, la cual opera esencialmente con algoritmos de reconocimiento de patrones, basados en extensos conjuntos de datos, que identifican regularidades estadísticas y las emplean para generar nuevos contenidos. Sin embargo, esta capacidad se encuentra limitada por la naturaleza de los datos recopilados, restringiendo la IA a combinar y recombinar elementos preexistentes. La ausencia de una conciencia subjetiva y de experiencias vividas impide a la IA establecer conexiones conceptuales innovadoras y generar ideas verdaderamente originales, características intrínsecas a la creatividad humana. Por lo mismo, si los datos con los que trabaja no incluyen ejemplos de pensamiento divergente o innovador, la IA más clásica tampoco será capaz de producirlo.

Otro problema es la incompletitud y la inexactitud de los datos otorgados, si los datos que se utilizan para entrenar a un modelo están incompletos, es decir, si falta información importante o si no abarcan todas las posibles variaciones de un fenómeno, el modelo resultante puede ser incapaz de tomar decisiones adecuadas. Asimismo, la falta de información actualizada en la red puede llevar a generar respuestas incorrectas o confusas, pues el sistema, como ChatGPT, está programado para dar respuestas a partir de los datos con los que cuenta, por lo que en ocasiones ira construyendo o inventando sobre la marcha, para poder atender a lo solicitado. Si nos vamos más específicamente al uso de una de IAG en el ámbito educativo, si a un estudiante le piden realizar un ensayo sobre Aristóteles, y este utiliza un chatbot para buscar la información sin verificar las fuentes o libros del mismo autor, se puede encontrar con la sorpresa que no solo el profesor identificó el uso de IA,

sino también que el ensayo estará mal realizado, debido a que el estudiante nunca verificó si lo indicado por el chatbot era correcto o no.

De lo anterior deriva otro problema, la proliferación de herramientas de IA, capaces de generar respuestas concisas y detalladas a una amplia gama de preguntas, suscita preocupaciones en torno al buen desarrollo del pensamiento crítico en los estudiantes. La facilidad y rapidez con que estas herramientas satisfacen las demandas cognitivas de las personas puede fomentar una cultura de dependencia, en la cual se privilegia la obtención de respuestas inmediatas sobre la construcción autónoma del conocimiento. Esta dinámica genera una zona de confort cognitivo, socavando la motivación intrínseca por la investigación y el aprendizaje profundo. Al externalizar el proceso cognitivo a una máquina, los estudiantes pueden perder la capacidad de formular preguntas pertinentes, evaluar críticamente la información y construir argumentos sólidos, lo que a su vez compromete el desarrollo de habilidades esenciales para la resolución de problemas complejos y contingentes que requieran de innovación.

La creciente digitalización de la sociedad exige una adaptación constante de las instituciones educativas. En este contexto, la IA emerge como un agente catalizador de transformaciones profundas en los modelos pedagógicos, ofreciendo nuevas posibilidades para optimizar los procesos de enseñanza-aprendizaje y replantear los paradigmas educativos tradicionales (Uzcátegui Pacheco & Ríos Colmenárez, 2024). La IA (Holmes, Bialik & Fadel, 2019) ha comenzado a ser utilizada en las aulas de manera silenciosa, sin importar si los estudiantes, profesores, padres y responsables políticos lo aceptan o no, los sistemas de aprendizaje inteligente, adaptativo o personalizado se están implementando cada vez más en escuelas y universidades en todo el mundo. El mismo Hawking (2018) destaca que los avances científicos actuales auguran un futuro lleno de posibilidades y desafíos. Para aprovechar al máximo este potencial, es esencial incentivar la participación de los estudiantes en el ámbito científico desde etapas tempranas, ya sea involucrando a los estudiantes en proyectos de investigación adaptados a su edad y nivel de conocimiento, organizando salidas educativas para despertar la curiosidad y el interés por la ciencia y la IA, así como también aprovechando las distintas plataformas educativas en línea para acceder a recursos y comunidades de aprendizaje.

En este contexto, internet se revela como una plataforma inigualable para el aprendizaje, la colaboración y la creación de conocimiento. Al conectar a personas de todo el mundo, internet fomenta la construcción de redes de aprendizaje colaborativo, comparables a las sinapsis de un cerebro global. En la era digital, el docente asume un rol protagónico en la promoción de un aprendizaje activo y crítico. Al fomentar la indagación, el análisis y la evaluación de la información, se impulsa el desarrollo de habilidades cognitivas superiores en los estudiantes. Este enfoque pedagógico trasciende la mera memorización de contenidos, invitando a los estudiantes a construir conocimientos significativos a partir de la reflexión y la argumentación. La enseñanza filosófica se erige como una herramienta fundamental para cultivar el pensamiento crítico y la capacidad de discernimiento, competencias indispensables para la toma de decisiones informadas en un mundo cada vez más complejo.

La IA emerge como una tecnología disruptiva con el potencial de transformar radicalmente nuestras sociedades. Es fundamental comprender que la IA debe concebirse como una herramienta al servicio del ser humano, más que como una amenaza. Al igual que el teléfono móvil se ha integrado de manera imperceptible en nuestra cotidianidad, la IA puede convertirse en una extensión de nuestras capacidades cognitivas, enriqueciendo nuestras experiencias y optimizando nuestras tareas. Para lograr una integración exitosa, es imperativo fomentar la alfabetización digital en la población, promoviendo la comprensión de los principios, aplicaciones y limitaciones de la IA. Sea a través de cursos en internet, sea por los profesores en el aula o incluso iniciativas hechas por juntas vecinales. Los gobiernos de turno saben de su importancia y de cómo afectara en un futuro próximo, por lo mismo, insisten en que su enseñanza se promueva en la población. Así como en su momento se dijo que las maquinas están para facilitar la vida de las personas, actualmente se sostiene la idea de coevolucionar, seres humanos e IA en una corderiva conjunta, influyéndose y transformándose mutuamente de manera profunda. Es fundamental que la sociedad en su conjunto participe en la definición del futuro de la IA, asegurando que esta tecnología se desarrolle de manera alineada con nuestros valores y principios. La educación, la investigación y la colaboración entre diferentes actores sociales son claves para construir un futuro en el que la IA sea una fuerza positiva para el cambio.

La presente investigación da cuenta del potencial de la IA para personalizar la educación a gran escala. Los modelos de aprendizaje adaptativo, basados en algoritmos sofisticados, permiten ajustar los contenidos y las actividades a las características y necesidades individuales de cada estudiante, optimizando así la eficacia del proceso educativo. Esta retroalimentación continua y personalizada fomenta la motivación y el compromiso de los estudiantes. Además, la automatización de tareas rutinarias, como la evaluación y la gestión de datos, libera tiempo a los docentes para poder así centrarse en aspectos más cualitativos de la enseñanza, como la promoción del pensamiento crítico y la resolución de problemas. La IA permite crear materiales educativos en una amplia variedad de formatos, como textos (ChatGPT), imágenes (Leonardo AI), presentaciones (Gamma) videos (Capcut), mapas mentales (ChatMind) y audio (Notta Ai), adaptándose a las preferencias de aprendizaje de cada estudiante. Por último, la IA facilita el acceso a una amplia gama de recursos educativos, lo que permite a los estudiantes construir conocimientos de manera autónoma y desarrollar habilidades digitales de formas más prácticas.

A pesar de los significativos avances en el campo de la IA y de la IAG, persiste una notable disparidad entre el vertiginoso desarrollo tecnológico y la comprensión pública de estas herramientas. Si bien se han llevado a cabo investigaciones para evaluar su alcance e impacto, los hallazgos revelan un uso deficiente y, en ocasiones, malintencionado por parte de la población. La utilización de la IAG para realizar trabajos sin una investigación previa o la práctica del plagio son claros ejemplos de esta problemática. Ante esta situación, resulta imperativo reducir la brecha existente entre el desarrollo tecnológico y la comprensión pública, fomentando la educación y la alfabetización digital desde temprana edad. Es así como los gobiernos de turno, políticos y especialistas han promovido diversas iniciativas, tales como la realización de encuestas, seminarios, la organización de eventos científicos y la implementación de talleres tanto presenciales como virtuales, con el objetivo de difundir conocimientos y fomentar una cultura digital. De esta forma, se promueve un diálogo abierto y transparente sobre los beneficios y riesgos de la IA en la era actual y en su devenir. Se hace necesario involucrar a diversos actores, en donde la formación continua del docente se erige como un pilar fundamental para la efectiva integración de la IA en los procesos educativos. Las instituciones educativas, en sintonía con las políticas públicas, deben impulsar la capacitación docente en esta área, aprovechando la disposición del

profesorado a adquirir nuevas competencias. Los resultados de las encuestas revelan un interés genuino por parte de los docentes en profundizar sus conocimientos sobre la IA, lo que augura un futuro prometedor para la implementación de esta tecnología en el aula.

Un aspecto que no debe ser subestimado en la discusión sobre la IA es el papel que puede jugar la filosofía en este contexto. Es crucial profundizar en cómo la filosofía puede influir en la forma en que entendemos y desarrollamos tecnologías avanzadas como la IA. Aunque pueda parecer evidente, su rol fundamental es cuestionar, la filosofía debe y tiene que preguntar qué se busca con la IA, cuál es el fin de esta. Al cuestionar los presupuestos ontológicos y epistemológicos que subyacen al desarrollo de esta tecnología, los filósofos aportan una perspectiva crítica que complementa los enfoques más técnicos. Por ejemplo, ¿estamos tratando de mejorar la calidad de vida de las personas, o simplemente buscamos maximizar la eficiencia en los procesos productivos? Estas preguntas no solo son relevantes desde un punto de vista teórico, sino que también tienen implicaciones prácticas en la manera en que se desarrollan y aplican las tecnologías de IA. La filosofía puede enlazar el camino hacia una convivencia humano-máquina, en donde la tecnología se diseñe y utilice de manera que complemente y potencie las capacidades humanas, sin deshumanizarnos. La IA, por sí misma, no se vuelve una amenaza, el que no sepamos utilizarla o no queramos aprender de ella, sí.

## Bibliografía

Adlawan, D. (2023). Los pros y los contras de la IA en la educación y cómo afectará a los profesores en 2023. Fuente: <https://www.classpoint.io/blog/es/los-pros-y-los-contras-de-la-ia-en-la-educacion-y-como-afectara-a-los-profesores-en-2023>

Agre, P. (2005). The Soul Gained and Lost: Artificial Intelligence as a Philosophical Project. En S. Franchi & G. Güzeldere (eds.) *Mechanical Bodies, Computational Minds. Artificial Intelligence from Automata to Cyborgs*. Cambridge, MA/London: The MIT Press

Aristóteles (1995). *Tratados de lógica (Órganon) II*. Editorial Gredos, S, A. Madrid.

Benokraitis, N. V., & Feagin, J. R. (2019). *Affirmative Action and Equal Opportunity: Action, Inaction, Reaction*. Routledge.

Bechtel, W. (1991). *Filosofía de la mente: Una panorámica para la ciencia cognitiva*. Editorial tecnos, S.A.

Bilen, A. M. (2016). *Aprendizaje autónomo de redes neuronales artificiales*. Universidad Nacional de Cuyo: Facultad de Ciencias Exactas y Naturales.

Boden, M (2017). *Inteligencia Artificial*, Inmaculada Pérez Parra, España, p. 11.

Bostrom, N. (2016). *Superinteligencia, caminos, peligros, estrategias*. Primera edición en español: Teell Editorial.

Cherubini, E. (2011). *Empatía y mente extendida*. Universidad Nacional de Córdoba.

Clark, A., & Chalmers, D. (1998). The extended mind. *ANALYSIS*, 58(1), 7–19.

Davidoff, L. (1989). *Introducción a la psicología*. Tercera edición en español por Mcgraw-Hillinteramericana de México, S.A. de C.V.

Diez, J. J. (2023). *Aprendizaje automático en ciencia de datos*. España: Universidad de burgos secretaria general.

Dormido Bencomo, S., & Cruz García, J. M. (1989). *Inteligencia artificial: pasado, presente y futuro*. Pág 12-13

Dreyfus e os Limites da I.A. (2015). Discursus: A Filosofia E Seus Meios. Fuente: [https://forumdediscursus.wordpress.com/contemporanea/dreyfus-e-os-limites-da-i-a/#N\\_1](https://forumdediscursus.wordpress.com/contemporanea/dreyfus-e-os-limites-da-i-a/#N_1)

Emilio, J., Hernández, B., Ortiz, C., & Morales, C. (2021). Cognición Extendida. Fuente: <https://www.cienciascognitivascorporizadas.com/4ea/cognici%C3%B3n-extendida>

Editorial Santillana (marzo del 2024). ¿Qué piensan los docentes de la IA? Fuente: <https://view.genially.com/663a8ed3a5106f00141f09c1/interactive-image-resultados-encuesta-ia-generativa>

Gates, B. (2023). La era de la IA ha comenzado. GatesNotes, the blog of Bill Gates.

Gómez, D. (2012). El impacto de la inteligencia artificial sobre el ser humano y sobre su seguridad. Universidad Militar Nueva Granada. Instituto de estudios Geoestratégicos y asuntos políticos IEGAP.

Gutiérrez, K. (2023). Inteligencia artificial generativa: irrupción y desafíos. Revista Enfoques. 4(2), 57-.

Harari, Y. N (2018). 21 lecciones sobre el siglo XXI. Penguin Random House Grupo Editorial, S. A. U.

Hawking, S. (2018). Breves respuestas a las grandes preguntas. Editorial Crítica.

Historia y evolución de la ia: desde sus inicios hasta las aplicaciones modernas (s.f). ICCSI. Fuente: <https://iccsi.com.ar/historia-y-evolucion-de-la-inteligencia-artificial/>

Holdsworth, J. & Scapicchio, M. (2024). ¿Qué es el deep learning? IBM.

Fuente: <https://www.ibm.com/es-es/topics/deep-learning>

Holmes, W., Bialik, M., Fadel, C. (2019). Artificial Intelligence In Education: Promises and Implications for Teaching and Learning. The Center for Curriculum Redesign, Boston.

Jerez, S. (2024). El nuevo modelo de ChatGPT superó a científicos con doctorados en un examen de ciencias duras. Artículo Web, Biobio Chile. Fuente:

<https://www.biobiochile.cl/noticias/ciencia-y-tecnologia/ciencia/2024/10/03/el-nuevo-modelo-de-chatgpt-supero-a-cientificos-con-doctorados-en-un-examen-de-ciencias-duras.shtml>

Legris, J. (2005). Conocimiento simbólico. Un capítulo de la historia de la metodología científica. Revista perspectivas metodológicas, Vol. 5, N.º 5.

Lee, K. (2020). Superpotencias De La Inteligencia Artificial: China, Silicon Valley y El Nuevo Orden Mundial. Editorial Planeta, S.A.

Lesta, J. (2024). Inteligencia Artificial: una encuesta reveló cuáles son los principales desafíos que afrontan los CEOs de Latinoamérica. Diario en línea, *Ámbito*. Fuente: <https://www.ambito.com/tecnologia/inteligencia-artificial-una-encuesta-revelo-cuales-son-los-principales-desafios-que-afrontan-los-ceos-latinoamerica-n6012003>

Masís, J. (2023). Apostillas fenomenológicas a la filosofía de la inteligencia artificial de Philip E. Agre. Revista de Filosofía. Universidad de Costa Rica, San José.

Newell, A. y Simon, H., “La ciencia de la computación como investigación empírica: símbolos y búsqueda” en Boden, M. (1994), *Filosofía de la Inteligencia Artificial*, FCE, México, p. 124.

OpenAI. (2024). ChatGPT (13 de noviembre del 2024) [Modelo GPT-4].

Fuente: <https://chatgpt.com/share/67352aa0-6168-8011-a837-28b07bc00c02>

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (25 de octubre del 2024). Habilidades, una prioridad para construir una IA ética según las empresas del Consejo Empresarial de UNESCO. Fuente: <https://www.unesco.org/es/articles/habilidades-una-prioridad-para-construir-una-ia-etica-segun-las-empresas-del-consejo-empresarial-de>

Polvorosa, E. (s.f). Explorando la “creatividad infinita” gracias a la inteligencia artificial generativa. Recuperado el 11 de noviembre de 2024. Fuente: <https://www.gmv.com/es-es/node/7794/printable/print>

Pontificia Universidad Católica de Chile (14 de mayo 2024). Encuesta revela percepción del cuerpo académico UC sobre el uso de inteligencia artificial en docencia. Fuente:

<https://desarrollodocente.uc.cl/2024/05/14/encuesta-revela-percepcion-del-cuerpo-academico-uc-sobre-el-uso-de-inteligencia-artificial-en-docencia/>

Raja, Vicente (2019): “Conexionismo”, Enciclopedia de la Sociedad Española de Filosofía Analítica. Fuente: <http://www.sefaweb.es/conexionismo/>

Russell, S. (2019). Human Compatible: Artificial Intelligence and the Problem of Control. Computer Science Division, University of California, Berkeley.

Salcedo Albarán, E. (2004). El experimento mental de la habitación china: máquinas entre la semántica y la sintaxis. Editor Fundación Método.

Sanabria, J. C. (2023). La evolución de la IA a lo largo de los años. Linked ID. Fuente:

<https://www.linkedin.com/pulse/la-evoluci%C3%B3n-de-ia-lo-largo-los-a%C3%B1os-juan-carlos-sanabria-rojas/?trackingId=PYzGtOmLQSWOkmZuD0u%2BbQ%3D%3D>

Sánchez-Martín, FM., Jiménez Schlegl, P., Millán Rodríguez, F., Salvador-Bayarri, J., Monllau Font, V., Palou Redorta, J., y Villavicencio Mavrigh, H (2007). Historia de la robótica: de Arquitas de Tarento al Robot da Vinci (Parte II). Servicio de Urología. Fundación Puigvert. Barcelona.  
\*Instituto de Robótica Industrial (IRI). Univ. Politécnica de Cataluña. CSIC.

Serrano, J. L., y Moreno-García, J (2024). Inteligencia artificial y personalización del aprendizaje: ¿innovación educativa o promesas recicladas? Universidad de Murcia, España.

Seggiaro, C (2014). La noción aristotélica de noûs: conocimiento de los primeros principios y vida contemplativa en el Protréptico de Aristóteles. Facultad de Filosofía y Letras, Universidad de Buenos Aires.

Slegarraga. (2023, 19 octubre). Aprendizaje profundo: el motor detrás de la inteligencia artificial. Pyvotal Solutions. Fuente: <https://www.pyvotal.com/deep-learning>

Terrones, A. (2018). Inteligencia artificial y ética de la responsabilidad. Cuestiones de Filosofía

Unesco (2021). Inteligencia artificial y educación: guía para las personas a cargo de formular políticas. Consultado en junio.

Fuente: <https://unesdoc.unesco.org/ark:/48223/pf0000379376>

Universidad Andres Bello (2024). Inteligencia Artificial Generativa como Herramienta Pedagógica: Estrategias y Aplicaciones para Docentes de Educación Superior. Fuente: <https://view.genially.com/66c6569662f52caf3b09228e>

Uzcátegui Pacheco, R. A. & Ríos Colmenárez, M. J. (2024). Inteligencia Artificial para la Educación: formar en tiempos de incertidumbre para adelantar el futuro. Areté, Revista Digital del Doctorado en Educación, 10(ee), 1–21.

Fuente: [http://saber.ucv.ve/ojs/index.php/rev\\_arete/article/view/29445](http://saber.ucv.ve/ojs/index.php/rev_arete/article/view/29445)

Van Vaerenbergh, S. (2024). Inteligencia artificial para potenciar la creatividad y la innovación educativa. Revista INFAD De Psicología. International Journal of Developmental and Educational Psychology., 1(1), 507–513.

Varela, F. et al. (1997). De Cuerpo Presente: Las ciencias cognitivas y la experiencia humana, Ed. Gedisa, Barcelona.

Vega, M. (2024). Encuesta concluye que a la mayoría de las personas no les podría importar menos la IA. Artículo Web, Biobio Chile. Fuente: <https://www.biobiochile.cl/noticias/ciencia-y-tecnologia/pc-e-internet/2024/05/30/encuesta-concluye-que-a-la-mayoria-de-las-personas-no-les-podria-importar-menos-la-ia.shtml>