



# Inteligencia artificial y sesgos algorítmicos: una reflexión sobre las prácticas de la verdad y la dominación

## Artificial intelligence and algorithmic biases: A reflection on the practices of truth and domination

Fecha de recepción: 11 de mayo de 2025 - Fecha de aprobación: 15 de agosto de 2025

Juan Manuel Hernández Aguilar<sup>1</sup>

Universidad nacional abierta y a Distancia. UNAD - Colombia

### Resumen

Este artículo expone la estrecha relación entre los saberes producidos por estadísticas ligadas a la política actual y el biopoder en la gestión de poblaciones desde la inteligencia artificial (IA). Siguiendo a Foucault en sus escritos *Vigilar y castigar*; *Seguridad, territorio, población*; *El nacimiento de la biopolítica* y *Defender la sociedad*, se argumenta que la IA funciona como una herramienta de control social, que facilita la vigilancia y la regulación del comportamiento mediante el entramado de discursos. La recopilación masiva de datos y el monitoreo constante de la población permiten que la IA opere como un mecanismo de biopoder que organiza y normaliza la vida social, e instaura condicionamientos en los discursos. El uso del análisis de datos de la IA lleva a cuestionar si las relaciones de dominio son inherentes al empleo de la tecnología. Para ello, se hace una breve referencia al origen de la técnica en la prehistoria y se contrasta con el uso de herramientas digitales, como en el caso de la IA. En ese sentido, se sostiene que el análisis de datos es utilizado por el sistema capitalista para establecer un régimen de vigilancia de la información, fenómeno que se ha denominado "Capitalismo de la vigilancia y tecnofeudalismo".

### Palabras clave

Sesgos; verdad histórica; inteligencia artificial; dependencia; dominación.

<sup>1</sup> Licenciado en Literatura y Profesional en Filosofía, Magíster en Filosofía y Doctor en Filosofía de la Universidad del Valle – Colombia. Actualmente docente en la Universidad Nacional Abierta y a Distancia y líder del grupo de investigación Cibercultura y Territorio. filosofoliterato@gmail.com ORCID 0009-0005-3103-7302

## Abstract

This article explores the close link between the knowledge produced by statistics related to current politics and biopower in population management through artificial intelligence (AI). Following Foucault's work in *Discipline and Punish, Security, Territory, and Population: The Birth of Biopolitics, and Society Must Be Defended*, it argues that AI acts as a tool of social control, facilitating surveillance and the regulation of behavior through the weaving of discourses.

## Keywords

Biases, historical truth, Artificial Intelligence, dependence, domination.

## Introducción

En el *Manual de ética aplicada en inteligencia artificial*, Olmeda e Ibáñez (2022) abordan los orígenes de un sesgo inherente a la IA, desde la génesis de sus formas de aprendizaje algorítmico. Se evidencia el sesgo respecto al dato histórico: la máquina parte de volúmenes históricos que reflejan prejuicios e interpretaciones propias de otras épocas. Los modelos resultantes de la IA, contruidos a partir de estos datos, pueden perpetuar patrones de contextos pasados, sin ajustarse necesariamente a la temporalidad actual. Por ello, se considera que el sesgo emerge desde el momento mismo de la recolección de información. Las bases de datos constituyen información que se recoge o se omite en función de intereses y fines específicos. Tomando en cuenta los posibles sesgos, partimos del sesgo de representación, que se refiere a un muestreo insuficiente al no incluir un número adecuado de elementos de los distintos segmentos; en el sesgo de aprendizaje y de medida se evidencia la errónea selección de atributos al medir un problema y las opciones de modelado de la métrica sobre cómo se amplifican las opciones; el sesgo de evaluación, definido por los datos de referencia empleados no tanto para entrenamiento, sino para la calidad de la información que representa la población; el sesgo de agregación, donde se utilizan datos de diversos grupos y se asume que el mapeo es similar para todos ellos, y, finalmente, el sesgo de implementación, presente cuando existe una desconexión entre el problema que el modelo resuelve y su uso real.

Estos defectos metodológicos y no éticos constituyen una respuesta a un cierto perjuicio de la manera en que la IA suele pensar, si cabe decir que “piensa”. Aquí la atención debe centrarse en la viabilidad y la efectividad, pero no en la eliminación de efectos metodológicos que muestran la existencia de técnicas de detección y de no erradicación de los sesgos. Estos acaban afectando la manera en que se narra y representa lo histórico, desde los hechos interpretados a partir de datos influenciados por tradiciones culturales y estados de opinión. En *Power and prediction: The disruptive economics of artificial intelligence*, Agrawal et al., (2022) señalan que estos sesgos conllevan a desviaciones sistemáticas a través de cálculos que impactan lo técnico y operativo sin considerar sus defectos. Estos problemas trasladan lo técnico a lo discriminatorio, manifestándose aspectos colectivos y personales en los juicios de valoración cultural.

En *Salud para quién: interseccionalidad y sesgo de la inteligencia artificial para el diagnóstico clínico*, Amaya-Santos et al. (2024) indican que la IA se ha investigado desde un enfoque transdisciplinar, y se abarcan disciplinas como la ética, la economía del comportamiento, la psicología, la sociología, el derecho, la política, el feminismo interseccional y las ciencias informáticas. Este análisis permite examinar diversos sesgos presentes en el sistema, desde los que afectan la evaluación hasta los que surgen en los datos de entrada, así como en los procesos de ajuste y salida. Los autores mencionan que la información puede estar afectada por múltiples tipos de sesgo: de métrica de rendimiento, de interpretación de resultados, cognitivos, algorítmicos, de confirmación y retroalimentación, demográficos, históricos, de medición, de etiquetado de estereotipos, así como de contenido, representación y muestreo.

Concluyen que la IA mantiene una regulación que juega un papel central en la forma de abordar los discursos. Además, esta normativa regula, mediante la formulación, la magnitud y el modo de tratar el bienestar individual, con el objetivo de responsabilizar a los sujetos por medio de refuerzos políticos. La confiabilidad, según Amaya-Santos et al. (2024), constituye “un cuello de botella crítico” (p. 4), debido a que la IA carece de transparencia informativa y en la explicación de sus resultados, estableciéndose mediante estándares que generalizan, sugieren y determinan las políticas públicas a través de sus respuestas,

influyendo en el desarrollo de esquemas armonizados. Es fundamental considerar la idea de responsabilidad sobre el poder, la comprensión y la justificación de las decisiones del sistema como una perspectiva que debe garantizar la ética en la investigación.

## Capítulo 1: Límites y configuraciones del discurso

En *Inteligencia artificial y sesgos*, Bertocchi (2022) advierte sobre la manera en que la IA opera mediante patrones que incorporan información generada y refinada a través de predicciones. La autora se centra en los sesgos cognitivos y algorítmicos, subrayando que este sistema automatizado refleja los valores de quienes desarrollan y entrenan la IA. Estos sesgos desempeñan un papel esencial en el entrenamiento, ya que constituyen la fuente de información que indica hasta qué grado el sistema cumple su tarea y si alcanza conclusiones válidas. El problema radica en que jueces y fiscales, con su bagaje cultural, experiencias de vida y opiniones previamente formadas, pueden incorporar los sesgos de su contexto y vivencias en las sentencias que dictan.

Cabe advertir que la IA, en caso de ser incorporada al derecho, suscita el debate sobre si termina por reproducir el sentido del derecho o simplemente lo presenta renovado, omitiendo la reflexión sobre su propio fundamento, pero ensalzando su epistemología actualizada. Bertocchi concluye que incluso los sesgos constituyen parte de la formación de la IA y difícilmente pueden eliminarse. Como afirma la autora, "El sistema de IA nace, se desarrolla y aprende a partir de datos, y estos datos provienen de humanos, quienes a través de variables los incorporan para que el sistema resuelva. Pero estas resoluciones no son infalibles" (2022, p. 369).

Al interrogar la IA, conviene esforzarse en comprender las técnicas mediante las cuales se enuncia la verdad sobre lo humano, reconocida como parte de un dispositivo que permite mostrar la manera en que hoy día se construye la historia y se resignifican los conceptos que regulan el actuar humano. Los diagnósticos erróneos sustituyen explicaciones de expertos, las reducen o trivializan, para perseverar en la manera en que la IA establece discursos avala-

dos, reproduciendo el orden institucional de los expertos y formulando nuevos lineamientos simbólicos y colectivos. La IA traza el modo científico tutelado por instituciones disciplinarias, como colegios, escuelas, hospitales, y técnicas de control de población, es decir, biopolítica, entendida aquí como la función de invadir la vida calculando dominios sociales. Estos elementos se articulan en un entramado de dispositivos que, al igual que en la lógica biopolítica descrita por Foucault, producen sujetos regulados por un lenguaje que tiende a englobar todo.

En el caso de la IA, esta verbalización abarcadora se centra en la organización y clasificación de la información, codificando y jerarquizando las formas en que se interpretan los sesgos. Así, la verdad generada por la IA no existe fuera de los marcos en que surge el discurso, y este marco opera, de manera biopolítica, fijando de antemano los límites de lo decible: qué discursos serán legitimados y cuáles quedarán excluidos. "En efecto, nunca hay que olvidar el principio de que el mercado es un regulador económico y social general; pero eso no quiere decir, sin embargo, que sea un elemento dado de la naturaleza que pueda encontrarse en la base misma de la sociedad" (Foucault, 2007, p. 172).

## **Capítulo 2: La IA como dispositivo biopolítico: poder, verdad y agencia en la era algorítmica**

La IA, en tanto tecnología biopolítica gubernamental, optimiza, predice y modela la conducta de poblaciones. De este modo, puede individualizar (perfilando usuarios) y totalizar (gestionando poblaciones) a la vez, elementos claves en la biopolítica. En un contexto neoliberal, la IA se convierte en infraestructura invisible para orientar comportamientos, anticipar necesidades y moldear decisiones. Por tanto, no existe una IA neutra: siempre está inscrita en relaciones de poder-saber.

Lo que permanece oculto y lo que puede compartirse a partir de las respuestas de la IA equivale a la técnica del mercado que determina cuáles ideas sustentan el ejercicio de la verdad. Así se desarrolla la sabiduría y las nuevas formas de creación académica del sujeto, que terminan instaurando la nueva verdad de mercado a manera de base para las acciones reguladoras y el pensamiento es-

estructurado por el propio mercado. Es en esta relación donde la problemática de la IA implica servidumbre y sustitución del dominio autónomo. Esto sustituirá el propio pensamiento individual, imponiendo una relación autónoma —a nivel genealógico— con los objetos de comprensión que serán asumidos como respuestas prescriptivas para el sujeto.

En suma, el modo en que el individuo se autogobierna por medio de estos algoritmos constituye la experiencia contemporánea de la IA, resultado de discursos expertos y de prácticas institucionales judiciales, pedagógicas y estatales, amalgamadas para definir lo que bordea los significados y queda anclado a los mecanismos de normalización académica y disciplinaria. Esta materialización demuestra cómo se adoptan modos de hablar e imponer en el sujeto formas de enunciación del problema de la verdad, reconstruyéndola en función del significado actualizado en la nueva discursividad histórica. Estrategias promovidas por Gobiernos como los de la Unión Europea, China y Estados Unidos colocan la experiencia individual frente a un modelo tecnológico desnaturalizado de la experiencia vivida.

A partir de la IA, el individuo puede adaptar su conducta, encontrar respuestas para su pensamiento y ser gobernado por entidades externas, es decir, por el discurso que circunscribe la herramienta elegida para priorizar un tipo de comportamiento, en detrimento de la capacidad de regresar a significados renovados. ¿Estamos dominados absolutamente por la IA? ¿Existe alguna forma de agencia en la relación con la IA? En Foucault, el poder no anula la agencia; el poder no es solo opresión, sino una red de relaciones que produce sujetos y posibilidades de elección. Esto implica que, incluso en un marco de control (como el impuesto por la IA a través de algoritmos, interfaces y datos), existen espacios para maniobrar, resistir, reinterpretar y subvertir.

Foucault denominaría a la IA un dispositivo: un ensamblaje de discursos, instituciones, saberes técnicos y prácticas que organiza qué pensar, cómo actuar y cómo hablar. Desde esta perspectiva, la agencia no es una posición “externa” al sistema, sino el uso estratégico de las reglas y posibilidades que el mismo sistema provee. De ahí que sea necesario resignificar su empleo: emplear la IA para propósitos alternativos a los establecidos por la corporación

o institución que la diseñó. Una estrategia puede ser interrogar sus “verdades”, utilizando sus respuestas como punto de partida para debates; también, la crítica y la investigación, generando contranarrativas mediante el entrenamiento de modelos propios o introduciendo datos y perspectivas ausentes en los datos de entrenamiento. Por último, el autogobierno: desarrollar conciencia sobre cómo la IA modela la conducta y decidir periódicamente cuándo y de qué modo será pertinente dejarse orientar por ella.

En este sentido, recobrar el cuidado de sí propuesto por Foucault respecto a la IA implica cultivar una relación crítica y consciente con estas tecnologías: no limitándose a consumir resultados, sino comprendiendo sus lógicas de funcionamiento, sus límites y sesgos. La agencia surge entonces cuando el sujeto no es únicamente usuario, sino actor activo que se forma a sí mismo en la relación con la IA.

Aquí resulta fundamental la articulación entre poder y libertad; al tomar la IA como modelo que orienta el deber ser del individuo respecto al poder y al cuestionar dicho poder, se explicita la posibilidad de influir en la conducta de los otros, conforme a la definición de la acción basada en la aparente toma de decisión y la conversión de narrativas en hechos que buscan sustituir la iniciativa individual. Por ello, no sorprende que actualmente muchas personas hayan dejado de realizar búsquedas tradicionales para usar la IA como principal fuente de consulta. La tendencia es recurrir a la IA para establecer “verdades” que determinan cuáles discursos o hechos resultan relevantes: actúa como “psicólogo” o “consultor”, y arrastra a un código de sociedad que raras veces es problematizado ni examinado, aceptando o rechazando los resultados conforme a los sesgos implícitos en el diálogo contemporáneo. Esta decisión impone normas de conducta, sin cuestionar dichas expectativas ni revalorar las formas históricas ya establecidas, enfrentando a un ser humano presa de la técnica y del nihilismo de la época.

Cabe recordar que la *techné* es considerada un conocimiento inferior a la episteme por su carácter principalmente práctico. Así, el objetivo final cuestiona cómo se constituye la verdad en los sistemas de IA: como una derivación de discursos que ya no requieren del sujeto, sino de modificaciones en la forma de

preguntar y de pensar, sin atención al contexto individual ni al aspecto particular de la experiencia, que es el que da significado en el diálogo entre sujeto y entorno.

Uno de los aspectos para afinar es el modo en que la dominación normaliza y naturaliza este ejercicio de subjetividad mediada por el lenguaje de la IA. Es en ese contexto en el que elementos de la producción de textos y exposiciones teóricas funcionan como instrumentos de estrategias de manufactura que concluyen en un proceso de facilidad permanente, acorde con las estrategias que ofrece la IA. Los sujetos, en ese aspecto, adquieren competencias aparentemente libres a partir de nuevos servicios que ya no los encierran en la posición de reforzar su tesis personal o común, sino de producir un aspecto de lo inmediato, prefabricado y de un orden social que se autoengendra y perpetúa una misma lectura mediante los propios agentes. En el momento en que se regulariza esta práctica, la posibilidad de que se caiga en sesgos económicos y sociales termina por establecer las condiciones bajo las cuales se piensa la misma verdad. La propia economía y las herramientas tecnológicas convierten a este nuevo individuo en alguien que no busca ir más allá de la multiplicación de las formas de optimización del sistema, donde se redefinen o construyen prácticas y reglas de juego que inciden en la sujeción ambiental.

Para Foucault, la reflexión sobre la tecnología no se limita a una mera manera de operar, sino que implica la regulación del medio mismo en el que los individuos intervienen e interactúan, y configuran las condiciones de posibilidad de tales acciones. La tecnología —entendida como un entramado de técnicas, saberes y dispositivos de poder— no solo orienta la conducta, sino que también establece un campo normativo que define lo aceptable y lo excluido. En este sentido, la IA puede leerse como un dispositivo que, al centralizar y organizar la producción de enunciados, promueve una nueva economía discursiva en la que los discursos se monopolizan y su regularidad se consolida en función de los criterios de verdad y pertinencia inscritos en la propia máquina. Así, la regulación técnica deviene también regulación política, al fijar las formas legítimas de intervención e interacción en el espacio social.



A este respecto, es importante establecer que ya no existen procesos ni formaciones permanentes, sino que se proporciona un nuevo régimen que termina por moldear incluso las relaciones sociales, económicas, de producción y de comprensión de la pareja. Ese control, que se realiza de manera libre, no obedece a relaciones sociales. En su lugar, surgen análisis que no se centran en la necesidad de buscar la oferta o la demanda, sino que más bien señalan un nuevo dispositivo de gobierno que busca que se hable bajo la presuposición de los discursos. Este es el cálculo de la IA: una acción reproducida mediante una técnica sin sujeto que termina por homogeneizar los discursos, integrándolos en una síntesis de intervenciones avaladas por instituciones que dirigen la verdad hacia una función. De esta forma se configuran los modos de validación y las relaciones que sustentan su esquema, sustituyendo el pensamiento como forma de cuestionamiento social.

### **Capítulo 3: Del cuidado de sí al capitalismo de la vigilancia: la inteligencia artificial como huésped inquietante**

¿De qué manera, entonces, este ambiente artificial favorece que la autorregulación del sujeto repare en sus prácticas del cuidado de sí? La IA tiene la capacidad de incrementar el capital humano mediante la creación de textos estandarizados y la unificación de la innovación de una sociedad interesada en satisfacer sus necesidades, capitalizando no su pensamiento, sino el modo en que este discurso económico hegemónico comanda el nacimiento de una producción de la información que se repite y establece lo que se debe pensar, el conocimiento y los simbolismos. El concepto de discurso económico hegemónico alude, en clave foucaultiana, a un régimen de producción, circulación y legitimación de enunciados en el que lo económico no se refiere al valor monetario, sino a la administración estratégica de lo decible: su rarefacción, distribución y control. Esta economía discursiva se torna hegemónica cuando consolida un conjunto dominante de verdades que, al naturalizarse, definen el horizonte de lo pensable e invisibilizan o subordinan otras formas de saber. Su hegemonía no emana de un único centro de poder, sino de la convergencia de múltiples dispositivos —estatales, corporativos, técnicos— que configuran las condiciones de posibilidad del discurso y fijan los criterios de legitimidad que sostienen un orden social determinado.

¿Es, pues, la IA una guía para la acumulación de juegos y factores que encuentran actualmente una manera de operar? ¿Hasta qué punto esos correlatos de prácticas se convertirán en discursos históricos y tendrán que cambiar no la economía, sino la forma y las técnicas de economía y de gobierno que se articularán en los sujetos con el fin de vender el humo sobre el trabajo de sí y procurar incrementar sus activos materiales? Desde este punto, el lenguaje de la IA y su gestión, tanto en la práctica como en las relaciones, potencian de forma genérica no al individuo, sino a la especie; no el trabajo como potencia, sino al individuo como mero agente de reproducción. No se trata de recurrir a las críticas sobre lo que se considera el ser humano al interrogarse en su instante de vida, sino de mostrar cómo queda expuesta esta incapacidad para conceptuar.

Se desarrolla un *modus operandi* tecnológico como nueva concepción de vida mediante la movilización de tecnologías como fin ético de autorregulación de los sujetos. Esto es lo que determinará el comportamiento que se busca dirigir y que ya no se sostiene en conductas fundamentadas en razones y pragmatismo, sino que sigue criterios proporcionales como referencia del trabajo, el desarrollo del trabajo y la formación de tejidos sociales que establecerán las nuevas líneas de fuga que determinan el último pensamiento; es decir, el pensamiento forjado en este nuevo acontecimiento. Hablar, entonces, en términos foucaultianos de técnicas de conducción del comportamiento implica reconocer que la razón de Estado ya no se limita a intervenir sobre objetos o dominios externos, sino que opera a través de prácticas y dispositivos que inciden directamente en la forma en que los sujetos se relacionan consigo mismos y con los otros. Siguiendo esta lógica, los individuos no solo interiorizan esos modos de regulación, sino que incorporan en su propio discurso el lenguaje y los esquemas de inteligibilidad producidos por la IA. En este proceso se generan archivos —entendidos como en la arqueología del saber o como sistemas organizados de enunciados— y monumentos —no como restos inertes del pasado, sino como configuraciones discursivas que conservan y fijan determinados regímenes de verdad— que consolidan y perpetúan la normatividad inscrita en los dispositivos de la IA. Así, la tecnología no solo modela interacciones presentes, sino que erige formas discursivas duraderas que actúan como referencia y marco para las prácticas futuras.

Ahora bien, el modo en cómo se gestiona y se potencia la vida hacia un discurso más productivo, más eficiente y seguro, establece finalmente el modo en que se forman los discursos. Se regula el modo de someter las contingencias del discurso y sus mecanismos de poder y contrapoder a una fórmula mediante el copiar y pegar un despliegue, como posibilidad de un conjunto de instituciones y procedimientos, de análisis y reflexiones calculadas que se ejercen de manera específica para establecer un límite al saber de la población. El sujeto se deshace en su intento de interpretarse a sí mismo, para quedar como un vacío que funge como espacio susceptible de ser llenado por preguntas que no se quieren responder o por comportamientos irreflexivos. *La arqueología del saber* de Foucault establece que las relaciones de fuerza pueden generar cambios en las condiciones de enunciación y en las claves de formación de los discursos y en la red de relaciones de habla. Este nuevo fenómeno pone de manifiesto la diferencia entre preguntar —como apertura de horizontes— y confundirlo con los sesgos que condicionan la forma en que se pregunta, lo que se problematiza y prohíbe las instancias de la significación, asegurando en códigos la condición de mantener lo aceptable como una manera de estabilizar el pensamiento.

Como vemos, las relaciones de dominación se manifiestan en el discurso a través de los algoritmos; sin embargo, esta influencia no se limita al plano estrictamente discursivo. La IA interviene también en el análisis de datos procedentes del flujo de información disperso en el sistema, un proceso que puede incorporar sesgos heurísticos al momento de acceder o consultar la información. A continuación, profundizaremos en cómo estas relaciones de dominio impregnan el uso de la tecnología, introduciendo una ambivalencia en la forma en que dependemos de la sistematización y el procesamiento automatizado. No obstante, cabe preguntarse si, en última instancia, estas mismas relaciones de dominación no generan un cierto grado de dependencia estructural del ser humano respecto de la máquina.

Quizás uno de los rasgos más inquietantes de la IA sea la posibilidad de dominación y esclavitud del ser humano. Si bien es cierto que la IA aún no domina completamente todas las esferas de lo humano, se debe reconocer que el individuo es cada vez más dependiente del uso aplicado de la tecnología en

diversos campos y, con base en ella, ha logrado alcanzar un mayor grado de bienestar en ciertas labores para las cuales se ha entrenado la IA. Incluso se puede afirmar que, gracias a la tecnología, el individuo goza de una mejor calidad de vida. No obstante, a modo de hipótesis, todo cambio en las condiciones materiales del humano conlleva una optimización de la calidad de vida, pero también la posibilidad de dominación. Es decir, todo avance en términos de progreso técnico debe entenderse como un mejoramiento de las condiciones de vida del ser humano, pero también como la posibilidad de ejercer control sobre otros. Este otro puede ser entendido como el dominio de otras personas o, incluso, del entorno natural. En ese sentido, poco importan los medios que se empleen para ejercer dominación, pues lo que interesa resaltar son las condiciones bajo las cuales se producen los efectos de poder: la IA puede resultar tan efectiva para la dominación como lo fue la técnica del fuego o el uso del lenguaje. Los tres son instrumentos para mejorar el bienestar del ser humano, pero también sirven para ejercer un control sobre sus condiciones de vida y el modo de desarrollo cultural. El individuo es, en suma, una especie biosocial, porque gran parte de nuestra estructura física es, de hecho, producto de la relación ininterrumpida entre naturaleza y cultura. En este sentido, la utilización de herramientas coadyuvó al desarrollo de nuestro cerebro.

Algunos acontecimientos en la historia de la técnica sugieren que la evolución cultural está vinculada al uso de técnicas con las cuales los antiguos homínidos pudieron modificar gradualmente la naturaleza. De esta forma, la utilización de piedras como el sílex fue aprovechada para fabricar herramientas y armas: desde cuchillos hasta raspadores, hachas y puntas de flecha; este material fue empleado para elaborar numerosos implementos. El descubrimiento del fuego le permitió al homínido la técnica de cocción, mejorando las condiciones higiénicas del consumo de alimentos. Igualmente, la agricultura permitió a los homínidos seleccionar artificialmente semillas para cultivar y cosechar frutas y verduras, incorporándolas a su dieta diaria. La invención de la rueda facilitó el desplazamiento de objetos. El uso del lenguaje se constituyó en un recurso comunicativo para representar y asimilar el mundo circundante; probablemente lo mismo sucedió con la pintura rupestre, considerada una protoescritura utilizada como extensión de la memoria, permitiendo registrar procesos del entorno natural.

El uso de técnicas y herramientas ha proporcionado históricamente al ser humano las condiciones necesarias para establecer, de forma gradual, un dominio sobre otras personas y la naturaleza. En este sentido, lo que muchos paleoantropólogos han denominado revolución neolítica puede compararse con la revolución de la información que actualmente distintos científicos cognitivos denominan IA. No obstante, la IA es algo más que una herramienta: ha transformado profundamente nuestra relación con la realidad, tal como ocurrió en su momento con la producción del hacha de sílex, el descubrimiento del fuego o la invención del lenguaje.

La opción clásica comprende la IA como problema filosófico y alude a la pregunta formulada por Alan Turing en 1949: *¿pueden pensar las máquinas?* Sin embargo, ¿qué significa pensar? Turing consideraba que cuando las máquinas realizaban cálculos en un lenguaje computacional binario, ese tipo de operación podía compararse con los procesos mentales que forman los pensamientos. Así, Turing sentó las bases de la teoría de la mente computacional. Esta teoría es un modelo explicativo que sostiene que la mente procesa información mediante procedimientos algorítmicos.

El desarrollo histórico de la IA presenta tres etapas. La primera, de 1956 a 1970: se caracteriza por grandes expectativas y limitaciones de recursos y conocimientos; la segunda, de 1971 a 1981, incorpora conocimientos heurísticos y produce algunos prototipos funcionales; y la tercera, de 1981 hasta el presente, implica la integración de sistemas de IA en la industria, incluyendo sistemas expertos, bases de datos inteligentes, robótica, redes neuronales y aprendizaje automático. Lo que sigue se concentra en esta última etapa, donde se extiende el uso comercial, industrial y científico de la IA, abarcando máquinas virtuales que operan mediante interactivos recursos de internet y sistemas de información. Así, la IA puede considerarse una rama de la informática cuyo propósito es fabricar máquinas capaces de realizar tareas que normalmente requieren inteligencia humana, mediante algoritmos y sistemas especializados, y que posibilita funciones como aprender, razonar o autocorregirse. Su objetivo es replicar o simular la inteligencia humana en cuatro dimensiones: el pensamiento humano, el pensamiento racional, la acción humana y la acción racional. Para

todos estos procesos, se apoya en dos técnicas de procesamiento: el *machine learning* y el *deep learning*.

Hoy en día, el *machine learning* (aprendizaje automático) y el *deep learning* (aprendizaje profundo) son técnicas empleadas en empresas de todos los sectores. El *machine learning* implica alimentar un ordenador con datos para que, a través de técnicas de análisis, la máquina aprenda a desarrollar un objetivo sin necesidad de una programación específica con millones de líneas de código, por lo que se denomina *aprendizaje automático*. El *deep learning* es una variante del *machine learning* que imita la arquitectura neuronal del cerebro humano y en el que una red neuronal artificial procesa los datos a través de múltiples capas, lo que permite a la máquina profundizar en su aprendizaje identificando conexiones y alterando los datos introducidos para alcanzar los mejores resultados (Santiago y Berbel, 2024).

Sin embargo, es necesario precisar que la crítica no se dirige a la IA como instrumento que facilita tareas, sino a su configuración como producto que promete una supuesta emancipación al transponer el pensamiento a su algoritmo. Este proceso, mediado por el entrenamiento del sistema, tiende a convertirla en un dispositivo de potenciación de la vida en el sentido del biopoder, pero en la práctica opera como un instrumento que alimenta la ilusión de servir mejor para comprender el poder. En realidad, tal dinámica no libera al pensamiento de la normatividad, sino que lo vincula aún más estrechamente a ella, reforzando sus lazos con las prácticas sociales y los efectos de estas. La IA es un elemento que legitima nuevas ventas; no resulta extraño que se hable de una proliferación de múltiples inteligencias artificiales cada día, las cuales terminan por normalizar, mediante el lenguaje y sus sesgos, un dominio sutil que modifica las prácticas de pensamiento bajo el avance de intereses económicos. Esta nueva tecnología es producto de un acontecimiento de racionalidad que asume el gobierno de las conductas, lo que supone el abandono de la preocupación genealógica de sí mismo, inclinándose ya no por el análisis, sino por el dinero a cambio de respuestas genéricas. En el artículo *La historia del algoritmo. Los "fallos" de la inteligencia artificial*, Julio César Guanche (2023) resalta que, aunque la IA no sea racista, pasa por alto la formación estructural

de la historia que incluye rasgos racistas persistentes, según los programas de historia de cada nación. Basta preguntarle a Leopoldo II de Bélgica en [Habloconlahistoria.es](http://Habloconlahistoria.es) su opinión acerca de las prácticas violentas en el Congo y la IA dará una respuesta superficial, políticamente correcta y desconocerá el trasfondo del racismo y el colonialismo. Fenómenos del pasado que reaparecen e impiden formular un verdadero diagnóstico que permita considerar auténticas las prácticas de sí.

Según Margaret Boden (2017), la IA tiene dos objetivos fundamentales: uno es tecnológico y el otro, científico. El primero emplea ordenadores para realizar tareas concretas; el segundo desarrolla conceptos y modelos de IA que contribuyen a resolver cuestiones sobre seres humanos y demás formas de vida. Así también, la IA ha permitido a psicólogos y neurocientíficos crear teorías sobre la entidad mente-cerebro y modelos sobre el funcionamiento cerebral y psicológico. Dicho de otro modo, permite estudiar "a qué cuestiones computacionales (psicológicas) responde y qué clases de procesamiento de la información le permiten hacerlo". Otra reflexión que aborda Boden (2017), y que permite regresar sobre el argumento inicial de este artículo, es qué tipo de relación mantenemos con un objeto tecnológico como la IA, cuestión que influye en nuestro modo de pensar la realidad y comprender, de alguna manera, la dignidad y la responsabilidad humana frente a los cambios disruptivos introducidos por cada avance técnico, los cuales repercuten en la reflexión filosófica, ya que las discusiones sobre IA han influido en muchos pensadores contemporáneos.

En la actualidad, numerosos filósofos apoyan sus planteamientos sobre la mente en conceptos de IA para analizar problemas de la historia del pensamiento, como el problema mente-cuerpo, el enigma del libre albedrío y los misterios de la conciencia. No obstante, estas reflexiones resultan controvertidas y aún no existe consenso acerca de si algún sistema de IA podría poseer verdadera inteligencia, creatividad o vida. Boden (2017) advierte que la técnica de la IA cuestiona y deja en una zona gris el concepto de humanidad y su porvenir, pues para muchos científicos cognitivos, la IA superará la inteligencia humana en todos los campos. Frente a esta realidad, algunos pensadores adoptan una postura optimista, mientras que otros muestran inquietud: ¿en qué lugar que-

dará la dignidad y la responsabilidad humana ante la acelerada trayectoria de la tecnología? Es aquí donde la reflexión filosófica debe comprender de forma crítica y razonada las implicaciones de este “huésped inquietante” en que se ha transformado la técnica de la IA para el futuro de la humanidad.

La IA es un huésped inquietante porque no solo genera comodidad, sino que nos expone a problemas humanos, económicos y sociales. En este sentido, la tecnología no solo ha mejorado las condiciones laborales, sino que también podría sustituir en el futuro a numerosos trabajadores en múltiples ámbitos industriales y de servicios: ciberseguridad, salud, finanzas, comercio, agricultura, transporte, comunicación, educación y robótica. La tecnología, en esa medida ambivalente, puede convertirse en un gran apoyo para el hombre, ayudando a eliminar cargas pesadas e incluso a encontrar cura para enfermedades. Le ha permitido a la humanidad conocer, dominar y acercar la naturaleza a nuestro saber, así como mejorar las condiciones de vida. Pero también se ha transformado en “instrumento de dominio, de muerte, de explotación de la naturaleza y de destrucción de la misma, de manipulación y dominio de las mentes y de las conciencias, de condicionamiento y disminución de las libertades” (Villalibre y Berciano, 1995, p. 61).

A continuación, se ampliará la idea de ambivalencia inherente a la tecnología, con un enfoque crítico respecto a la técnica de procesamiento de información que funda la IA, y que muestra sus vínculos con el capitalismo de la vigilancia y el tecnofeudalismo. La ambivalencia tecnológica no implica, en realidad, juicio valorativo sobre la técnica de la IA, sino una dependencia tecnológica respecto a la herramienta en diversos campos humanos, como el análisis de datos. Actualmente, el análisis de datos es una valiosa tecnología de apoyo en biología, química, medicina o ecología, como sucede en la bioinformática, y va más allá del campo investigativo. El procesamiento de datos informáticos es utilizado para examinar los datos del consumidor mediante la *big data* y el análisis estadístico para prever tendencias electorales en la sociedad. Como se observa, la ambivalencia y dependencia pueden llegar a ser confusas, pues la tecnología de datos también sirve para la defensa nacional, por ejemplo, para dirigir la trayectoria de drones y monitorear misiles en caso de conflicto bélico.



En *Capitalismo de la vigilancia*, Shoshana Zuboff (2020) presenta una imagen relevante del panorama derivado de la tecnología capitalista actual. El libro refiere cómo las grandes empresas tecnológicas utilizan experiencias humanas y datos personales de los usuarios para predecir el comportamiento social. Este sistema es diferente al industrial, comercial o financiero. Funciona solo a través de internet y la mano de obra son los propios usuarios, quienes gratuitamente ponen su vida privada al servicio de las empresas tecnológicas. El sistema amenaza la privacidad individual porque los términos de uso de las páginas web no siempre explican claramente sus implicaciones. Además, las aplicaciones buscan generar adicción y son formas dominantes de interacción social, por lo que la mayoría de usuarios consiente el uso de los datos en plataformas de análisis. Así, se plantea que el capitalismo de la vigilancia lleva a una economía feudal de la información, es decir, a un tecnofeudalismo, dado que las plataformas se encaminan a convertirse en nuevos feudos (Durand, 2021).

Más allá de la lógica territorial para el acaparamiento de los datos originales, el bucle informacional de los servicios digitales crea, como hemos señalado, una situación de dependencia para la sociedad. No solo porque los algoritmos que observan las prácticas de las personas tienden a convertirse en medios de producción indispensables para la existencia cotidiana, sino también porque el registro de los individuos en las plataformas digitales permanece en el procesamiento del tráfico y el flujo de información compartido dentro del sistema.

## Conclusiones

Es evidente que, cada día, dependemos más de la tecnología, por lo que resulta ingenuo emitir juicios de valor desde una postura excesivamente optimista que celebre los progresos de la IA sin reservas. Igualmente, tampoco pueden desestimarse los avances de esta tecnología, aunque también se emplee para la guerra o la destrucción. La biotecnología, por ejemplo, podría erigirse en la clave para erradicar graves problemas como el hambre, contribuyendo al mejoramiento de la agricultura y la medicina, la detección de enfermedades a largo plazo e, incluso, a la conservación ecológica. En definitiva, el problema de la técnica no reside en ella misma, sino en la idea de que podría llegar a

sustituir el análisis de expertos que, en su labor, resisten en la práctica a un poder político de la verdad, manteniendo la pregunta sobre sí mismos como sujetos afectados. Continuamente estamos alimentando los sistemas de información con nuestros datos y, con ello, cediendo ininterrumpidamente nuestra libertad a un sistema basado en la vigilancia, cada vez que ingresamos en una red social o completamos un formulario en línea. De este modo, estamos alimentando a un nuevo leviatán (metáfora del poder estatal al que se entrega potestad y libertad para eliminar cualquier acto de desorden social que derive en un estado de naturaleza): el capitalismo de la vigilancia, visión que impide la existencia de una relación entre libertad y política. No es que la máquina destruya empleos, sino que hemos creado herramientas para alimentar una ilusión colectiva de estatificación de la historia. Las nuevas formas discursivas de establecer verdad se someten a una particularidad racional transformada en sentido común, que busca llenar de contenidos vacíos de significado la apuesta estética de la existencia humana, es decir, la capacidad de moldearnos a nosotros mismos.

## Referencias

- Agrawal, A., Gans, J., & Goldfarb, A. (2022). *Power and prediction: The disruptive economics of artificial intelligence*. Harvard Business Review Press.
- Amaya-Santos, S., Jiménez-Pernett, J., y Bermúdez-Tamayo, C. (2024, agosto). ¿Salud para quién? Interseccionalidad y sesgos de la inteligencia artificial para el diagnóstico clínico. *Anales del Sistema Sanitario de Navarra*, 47(2). <https://dx.doi.org/10.23938/assn.1077>
- Bertocchi, N. G. (2022). Inteligencia artificial y sesgos. *Revista de la Escuela Judicial*, 3(3), 359–372. <https://doi.org/10.59353/rej.v3i3.64>
- Boden, M. (2017). *Inteligencia artificial*. Turner Publicaciones.
- Durand, C. (2021, 29 de enero). Tecnofeudalismo: La nueva gleba digital. *Viento Sur*, (173). <https://vientosur.info/tecnofeudalismo-la-nueva-gleba-digital/>
- Foucault, M. (2007). *El nacimiento de la biopolítica*. Fondo de Cultura Económica.

- Guanche, J. L. (2023, 15 de mayo). *La historia del algoritmo. Los "fallos" de la Inteligencia Artificial*. Unesco. <https://www.unesco.org/es/articles/la-historia-del-algoritmo-los-fallos-de-la-inteligencia-artificial> <https://nuso.org/articulo/capitalismo-de-vigilancia/>
- Olmeda, M. V., e Ibáñez, J. C. (2022). *Manual de ética aplicada en inteligencia artificial*. Comercial Grupo Anaya.
- Santiago, C. J. M., y Berbel, A. N. (2024). La Geo Inteligencia Artificial (GeoAI) como una herramienta de participación e innovación en el aula: el Alcázar Real, un estudio de caso en el espacio urbano de Ciudad Real. *Ensayos*, 39(2), 97–112.
- Villalibre, M. B., y Berciano, M. (1995). *La técnica moderna: reflexiones ontológicas*. Universidad de Oviedo.
- Zuboff, S. (2020). Capitalismo de la vigilancia. *Política exterior*, 34(194), 7–12.



# ESPACIO SOCIOLOGICO

