




DE LA VERIFICACIÓN EMPÍRICA A LA COHERENCIA COMPUTACIONAL: RACIONALIDAD PREDICTIVA EN LA ERA DE LA INTELIGENCIA ARTIFICIAL

ARTIFICIAL INTELLIGENCE AS AN EPISTEMIC AGENT: THE PREDICTIVE TURN AND
THE TRANSFORMATION OF SCIENTIFIC RATIONALITY



Fernando Ramos-Zaga  

Consejo Nacional de Ciencia y Tecnología e Innovación Tecnológica: Lima, PE

Resumen: La expansión de la inteligencia artificial (IA) en el descubrimiento científico ha transformado las bases del conocimiento al desplazar la verificación empírica por la predicción algorítmica como nuevo criterio de verdad. En ese sentido, el presente artículo tiene por objetivo analizar críticamente el estatuto ontológico y epistemológico de la IA, examinando si la predicción algorítmica instituye un determinismo tecnocientífico que redefine las nociones de causalidad y verdad científica. El estudio revela una mutación epistémica en la que la eficacia predictiva sustituye la explicación causal, instaurando un régimen de predictibilidad operativa y una racionalidad performativa donde los modelos algorítmicos configuran la realidad que pretenden describir. El hallazgo central indica que la ciencia contemporánea enfrenta una paradoja entre precisión y comprensión, ya que la autonomía del cálculo erosiona la inteligibilidad. En un sentido más amplio, este cambio anuncia un horizonte en el que la racionalidad científica se redefine bajo una lógica de control anticipatorio que exige una epistemología crítica y responsable.

Palabras clave: inteligencia artificial; epistemología; ontología; predicción algorítmica; determinismo tecnocientífico; performatividad.

Abstract: *The expansion of artificial intelligence (AI) within scientific discovery has transformed the foundations of knowledge by replacing empirical verification with algorithmic prediction as the new criterion of truth. Accordingly, this article aims to critically analyze the ontological and epistemological status of AI, examining whether algorithmic prediction establishes a technoscientific determinism that redefines the notions of causality and scientific truth. The study reveals an epistemic mutation in which predictive efficacy replaces causal explanation, establishing a regime of operational predictability and a performative rationality where algorithmic models actively shape the very reality they seek to describe. The central finding indicates that contemporary science faces a paradox between precision and understanding, as the autonomy of computation erodes intelligibility. More broadly, this shift signals a horizon in which scientific rationality is redefined under a logic of anticipatory control, demanding a critical and responsible epistemology.*

Keywords: *artificial intelligence, epistemology, ontology, algorithmic prediction, technoscientific determinism, performativity.*



1 INTRODUCCIÓN

La irrupción de la inteligencia artificial (IA) en el ámbito del descubrimiento científico ha introducido un desplazamiento epistémico que supera la dimensión instrumental para afectar el núcleo ontológico de la ciencia contemporánea. La emergencia de sistemas capaces de modelar, predecir y generar estructuras materiales con precisión atómica, como los desarrollos de David Baker en diseño computacional de proteínas (Kuhlman et al., 2003) y las arquitecturas neuronales de AlphaFold creadas por Demis Hassabis y John Jumper (Jumper et al., 2021), galardonados con el Premio Nobel de Química en 2024, ha modificado las fronteras entre representación y realidad. La ciencia parece haberse adentrado en una fase en la que la simulación no solo reproduce el mundo, sino que lo anticipa, reconfigurando la relación entre el conocimiento, la materia y la verdad. En consecuencia, la incorporación de la IA no puede entenderse simplemente como un avance técnico, sino como una mutación del propio régimen de racionalidad que define lo que se considera científico.

Las discusiones contemporáneas sobre la epistemología del cálculo y el determinismo computacional (Wolfram, 2002; Deutsch, 2011) reactivan un imaginario que se remonta al demonio de Laplace (1814), actualizado ahora por el algoritmo autoaprendente. Desde la filosofía de la ciencia, esta transformación implica una transición desde un paradigma representacional, donde la teoría busca corresponderse con el fenómeno, hacia un paradigma generativo en el que el modelo produce el fenómeno que describe (Barad, 2007; Parisi, 2019). La performatividad del conocimiento se convierte en principio estructurante de la racionalidad tecnocientífica, y la coherencia interna del cálculo adquiere un estatus equivalente, e incluso superior, a la verificación empírica (Floridi, 2020). En este marco, la IA se presenta como un agente epistémico autónomo, capaz de sustituir la observación por predicción y la causalidad por correlación, desplazando los límites tradicionales de la explicación científica (Anderson, 2008).

La literatura reciente muestra una brecha significativa en torno a la comprensión filosófica del estatuto ontológico de la IA en la ciencia. Si bien se ha discutido extensamente su impacto técnico en disciplinas como la biología computacional o la física de materiales, persiste un vacío teórico en la evaluación crítica de su papel como entidad productora de conocimiento. La falta de una reflexión profunda sobre cómo la predicción algorítmica transforma las nociones de verdad y causalidad impide captar las implicaciones de fondo de esta mutación. El desafío radica en determinar si la IA inaugura una forma de determinismo tecnocientífico que

reconfigura la ciencia como una práctica de simulación total, o si, por el contrario, abre un horizonte ontológico de indeterminación emergente donde la realidad se concibe como un espacio de posibilidades computacionales en continua actualización.

La investigación se justifica por la necesidad de articular un marco teórico que permita interpretar la IA no como mera herramienta, sino como operador ontológico. Analizar su función en los procesos de descubrimiento científico resulta crucial para comprender cómo el cálculo, al sustituir la verificación empírica por la coherencia del modelo, altera las condiciones mismas del conocimiento. Al examinar la convergencia entre los logros técnicos de Baker, Hassabis y Jumper y las perspectivas filosóficas del realismo agencial (Barad, 2007) y la infosfera (Floridi, 2020), se busca iluminar las tensiones entre predicción, causalidad y realidad, aportando una lectura crítica del nuevo régimen epistémico que se configura en la era del algoritmo.

Las implicaciones prácticas de esta investigación se extienden a los campos de la ética científica, la política del conocimiento y la gobernanza tecnológica. Comprender la IA como agente epistémico implica reconsiderar los criterios de validación, responsabilidad y autoría en la producción científica. Si la verdad depende cada vez más de la coherencia del modelo y no de su correspondencia empírica, los mecanismos de revisión y legitimación del saber podrían transformarse radicalmente. Además, el desplazamiento hacia una racionalidad algorítmica introduce riesgos asociados con la opacidad del cálculo y la delegación de juicio a sistemas no interpretables, lo que demanda nuevas formas de regulación epistemológica y social.

El análisis de la IA en el descubrimiento científico se inscribe también en los grandes desafíos contemporáneos de la filosofía de la técnica y la teoría del conocimiento. Las tensiones entre simulación y realidad, predicción y causalidad, cálculo y comprensión, dialogan con debates más amplios sobre la posthumanidad (Braidotti, 2019), la tecnociencia performativa (Barad, 2007) y la automatización del pensamiento (Zuboff, 2019). En un contexto global atravesado por la aceleración de los sistemas de aprendizaje automático, la pregunta por la autonomía de la ciencia frente al poder del algoritmo adquiere un valor estratégico para el futuro del conocimiento humano.

En ese marco, el objetivo del presente artículo es analizar críticamente el estatuto ontológico y epistemológico de la IA en el descubrimiento científico contemporáneo, con el fin de examinar si la sustitución de la verificación empírica por la coherencia computacional configura una forma de determinismo tecnocientífico. La contribución esperada radica en ofrecer una reflexión interdisciplinaria que combine filosofía de la ciencia, teoría de la

información y epistemología posthumanista, aportando una base conceptual sólida para comprender cómo la IA redefine la noción de verdad y causalidad en la ciencia, e inaugurando así un nuevo horizonte de racionalidad donde el cálculo se convierte en principio constitutivo de lo real.

2 EL MODELO PREDICTIVO COMO NUEVO FUNDAMENTO DE LEGITIMIDAD EPISTÉMICA

La noción de descubrimiento científico ha sido un eje estructural en la epistemología desde los debates entre el inductivismo y el falsacionismo hasta las concepciones contemporáneas sobre la práctica científica (Popper, 1959; Kuhn, 1962). El descubrimiento se ha concebido tradicionalmente como un proceso mediante el cual un fenómeno desconocido o insuficientemente comprendido se revela a través de la observación, la experimentación o la elaboración teórica.

Tal perspectiva presupone una ontología realista, donde la realidad existe de manera independiente del sujeto cognoscente, y una epistemología sustentada en la verificación empírica como criterio de validación del conocimiento (Hacking, 1983). La introducción de la IA como agente epistémico, sin embargo, transforma radicalmente ese marco conceptual al instaurar una modalidad de conocimiento en la que la predicción computacional antecede e incluso reemplaza, en ciertos contextos, la observación empírica.

Un ejemplo paradigmático de esta mutación se encuentra en los sistemas de predicción estructural de proteínas, particularmente en AlphaFold, desarrollado mediante arquitecturas de redes neuronales profundas entrenadas con extensos conjuntos de datos sobre secuencias y estructuras proteicas conocidas (Jumper et al., 2021).

Dichos sistemas logran predecir con gran precisión la conformación tridimensional de proteínas cuya estructura no ha sido determinada experimentalmente, alcanzando niveles comparables a los de la cristalografía de rayos X o la resonancia magnética nuclear. Esta capacidad plantea un interrogante central acerca del estatuto epistemológico de las estructuras predichas, pues si una conformación molecular puede anticiparse computacionalmente con exactitud atómica sin recurrir a observación empírica, se produce un desplazamiento profundo entre el conocimiento científico y la verificación experimental que históricamente había definido la ciencia moderna (Latour y Woolgar, 1979).

Dicho desplazamiento no se limita al ámbito metodológico, sino que afecta las condiciones ontológicas del conocimiento. La estructura predicha por AlphaFold no constituye una simple hipótesis pendiente de confirmación, sino que adquiere un estatuto operativo dentro de la investigación biomédica y biotecnológica. Las predicciones se utilizan para diseñar experimentos, desarrollar fármacos y formular hipótesis funcionales sobre procesos biológicos, bajo el supuesto de que la estructura computacional es suficientemente confiable para sustituir, al menos temporalmente, a la estructura empíricamente determinada (Senior et al., 2020). Esta práctica revela una mutación en el régimen de verdad científica, en la medida en que la coherencia interna del modelo y su capacidad predictiva otorgan a la predicción algorítmica una legitimidad epistémica equiparable a la observación empírica directa.

La filosofía de la ciencia ha reconocido desde hace décadas que los modelos científicos no se limitan a reflejar pasivamente la realidad, sino que actúan como herramientas epistemológicas que configuran la forma en que los fenómenos son conceptualizados e intervenidos (Cartwright, 1983; Morgan y Morrison, 1999).

Sin embargo, la IA intensifica esta dimensión performativa de los modelos. Los algoritmos de aprendizaje profundo no solo representan regularidades preexistentes en los datos, sino que producen nuevas formas de inteligibilidad mediante la detección de patrones que exceden la capacidad cognitiva humana (Kitchin, 2014). De este modo, la IA emerge como un agente cognitivo no humano que genera conocimiento de forma autónoma, aunque sus procesos internos permanezcan en gran medida inaccesibles para los investigadores que la utilizan.

La opacidad algorítmica introduce, en consecuencia, un desafío epistemológico decisivo para la noción de explicación científica. La tradición filosófica ha sostenido que la ciencia no se limita a predecir fenómenos, sino que debe explicarlos mediante la identificación de mecanismos causales y principios generales (Salmon, 1984; Woodward, 2003). La distinción entre predicción y explicación resulta, por tanto, esencial para mantener la inteligibilidad del conocimiento científico, ya que la mera capacidad de anticipar resultados sin comprensión causal podría reducir la ciencia a un saber instrumental carente de sentido explicativo. Los sistemas de IA basados en aprendizaje profundo manifiestan con frecuencia una potencia predictiva extraordinaria, pero acompañada de opacidad respecto de los mecanismos que sustentan sus resultados, lo cual genera una tensión estructural entre eficacia predictiva y comprensión explicativa (Buckner, 2018).

En esa dirección, la tensión puede interpretarse como un desplazamiento desde un paradigma explicativo hacia un paradigma anticipatorio, en el cual la verdad científica se define por la exactitud del pronóstico antes que por la elucidación causal (Anderson, 2008). La lógica del paradigma anticipatorio adquiere plausibilidad en ámbitos caracterizados por alta complejidad, donde los sistemas exhiben propiedades emergentes y comportamientos no lineales que desafían la causalidad clásica.

En tales contextos, la capacidad predictiva puede asumirse como criterio suficiente de validez epistémica. Sin embargo, aceptar esa premisa conlleva riesgos significativos. La sustitución de la explicación por la predicción podría reducir la ciencia a una práctica tecnocrática centrada en la eficacia operativa, debilitando su vocación comprensiva y transformando el conocimiento en una forma avanzada de tecnología sin fundamento teórico sólido (Humphreys, 2009).

Un ejemplo ilustrativo de la coexistencia entre predicción y explicación puede encontrarse en el trabajo de David Baker sobre diseño computacional de proteínas, plasmado en el software Rosetta. Este sistema integra principios fisicoquímicos relacionados con la energía de las conformaciones moleculares con procedimientos algorítmicos para diseñar proteínas de estructuras y funciones específicas (Kuhlman et al., 2003). La combinación de fundamentos teóricos y técnicas de optimización computacional sugiere que la predicción y la explicación no son necesariamente excluyentes. Sin embargo, incluso en este modelo híbrido, la complejidad de los cálculos implica que una parte considerable de las soluciones emerge de procesos de optimización algorítmica cuya racionalidad no resulta plenamente accesible para la comprensión humana (Dill y MacCallum, 2012).

A partir de estos desarrollos, la IA no puede concebirse únicamente como una herramienta que amplifica la capacidad humana de cálculo, sino como una forma cualitativamente distinta de cognición científica. Los algoritmos de aprendizaje profundo operan en espacios de alta dimensionalidad, identificando correlaciones y patrones que trascienden la intuición y la visualización humana (LeCun et al., 2015). La posibilidad de generar inferencias válidas en regímenes de complejidad informacional inabordable para el intelecto humano sugiere que la IA actúa como una auténtica extensión de la cognición científica, no como una mera prolongación instrumental de la mente humana (Clark y Chalmers, 1998).

La expansión de la cognición hacia sistemas híbridos humano-máquina introduce preguntas de fondo sobre la naturaleza del sujeto epistémico y la agencia científica. Cuando el

conocimiento emerge de ensamblajes donde los algoritmos desempeñan un papel irreductible, la figura tradicional del científico como sujeto autónomo del saber se torna insuficiente para describir la dinámica epistémica (Knorr-Cetina, 1999). La agencia cognitiva se distribuye entre actores humanos, artefactos técnicos, infraestructuras computacionales y conjuntos de datos, conformando una red sociotécnica en la que la producción del conocimiento resulta de interacciones complejas y materialmente mediadas (Latour, 2005). Desde esta perspectiva, la IA deja de ser un instrumento externo a la práctica científica para constituirse en un agente activo que participa en la configuración misma de los objetos de conocimiento.

Sin embargo, la distribución de la agencia epistémica suscita interrogantes éticos y epistémicos sobre la autonomía y la responsabilidad en la ciencia. A medida que los algoritmos adquieren un papel decisivo en la validación del conocimiento, existe el riesgo de que la práctica científica se subordine a una autoridad algorítmica opaca, cuya legitimidad no puede ser examinada mediante criterios externos al propio sistema computacional (Pasquale, 2015). El problema se agrava si se considera que los sistemas de IA son entrenados con datos históricos susceptibles de contener sesgos sistemáticos, lo que puede conducir a la reproducción y amplificación de desigualdades epistémicas y sociales (Noble, 2018). La dependencia creciente de la predicción algorítmica, sin mecanismos de transparencia y control, amenaza así con erosionar la dimensión crítica que históricamente ha definido la racionalidad científica.

La incorporación de la IA a la práctica científica transforma también la temporalidad del descubrimiento. La investigación moderna se ha articulado tradicionalmente en ciclos de formulación de hipótesis, experimentación, análisis y revisión, procesos que implican una temporalidad regida por limitaciones materiales y cognitivas humanas (Rheinberger, 1997). Los sistemas algorítmicos, al acelerar drásticamente la generación y evaluación de hipótesis, alteran esa temporalidad, produciendo una ciencia marcada por la inmediatez de la anticipación.

En ese sentido, la aceleración epistémica no es únicamente cuantitativa, pues redefine el ritmo de la producción científica y comprime los intervalos de reflexión crítica necesarios para la deliberación racional. La ciencia resultante corre el riesgo de perder la distancia temporal que permite la verificación cuidadosa, sustituyéndola por una inmediatez continua que, en nombre de la eficiencia, podría socavar la calidad reflexiva y la prudencia metodológica propias de la investigación rigurosa (Rosa, 2013).

La progresiva sustitución de la explicación causal por la eficacia anticipatoria obliga a examinar la relación entre predicción y necesidad. La potencia de los modelos algorítmicos sugiere un retorno implícito a la idea de regularidad universal, pero bajo una lógica estadística

y no ontológica. Comprender si la predictibilidad implica determinismo o si ambos conceptos responden a regímenes epistémicos distintos requiere un análisis conceptual que clarifique los límites de la racionalidad algorítmica. A tal cuestión se dedica la siguiente sección, centrada en la distinción entre determinismo y predictibilidad como pilares de la nueva epistemología tecnocientífica.

3 DISTINCIÓN ENTRE DETERMINISMO Y PREDICTIBILIDAD EN LA RACIONALIDAD ALGORÍTMICA

La capacidad predictiva de los sistemas de IA suscita una reconsideración profunda sobre la estructura ontológica del mundo natural y sobre la posibilidad de un orden subyacente que articule las regularidades empíricas. Si un algoritmo puede anticipar la conformación tridimensional de una proteína a partir de su secuencia de aminoácidos, resulta plausible pensar que el nivel molecular presenta regularidades suficientemente estables como para posibilitar tal anticipación. De este modo, se reactiva el debate clásico entre determinismo y contingencia, presente desde las formulaciones de la mecánica clásica hasta las interpretaciones contemporáneas de la mecánica cuántica (Earman, 1986).

En relación con la tradición determinista, la formulación del demonio de Laplace representa un punto de inflexión conceptual. Laplace sostuvo que, si una inteligencia conociera las posiciones y velocidades de todas las partículas del universo junto con las leyes que rigen sus interacciones, podría deducir tanto el pasado como el futuro de manera exacta (Laplace, 1814). Esta hipótesis define un universo donde la totalidad de los estados futuros está inscrita en el presente. No obstante, el surgimiento de la termodinámica estadística, la teoría del caos y la mecánica cuántica introdujo nociones de irreversibilidad, sensibilidad a las condiciones iniciales e indeterminación fundamental, lo que problematizó la concepción laplaciana del universo como sistema estrictamente determinista (Prigogine, 1997; Strogatz, 2015).

A partir de estas transformaciones conceptuales, la IA puede considerarse una reformulación tecnológica del ideal laplaciano, aunque con fundamentos epistémicos distintos. Los sistemas de aprendizaje automático no acceden al conocimiento exhaustivo de todas las variables, sino que infieren patrones estadísticos en grandes volúmenes de datos (Goodfellow et al., 2016). Por consiguiente, la predictibilidad algorítmica no implica determinismo en sentido ontológico, sino una forma de regularidad emergente que se sostiene en la estabilidad estadística de correlaciones observadas. La capacidad de AlphaFold para anticipar estructuras

proteicas ilustra esta diferencia, dado que su eficacia depende de la identificación de configuraciones recurrentes en el espacio de secuencias y estructuras previamente registradas (Jumper et al., 2021).

En consecuencia, la distinción entre determinismo y predictibilidad adquiere relevancia filosófica. Mientras el determinismo se refiere a la fijación causal de los estados del mundo, la predictibilidad designa la capacidad epistémica de anticipar tales estados sin que ello implique una necesidad ontológica (Hofer, 2016). Un sistema puede ser impredecible en la práctica y, sin embargo, determinista en su estructura, del mismo modo que puede ser estadísticamente predecible aun siendo indeterminista. Los sistemas de IA se inscriben en este último registro, operando sobre correlaciones probabilísticas más que sobre relaciones causales exhaustivas.

Aun así, algunos enfoques interpretan la eficacia predictiva de la IA como indicio de un orden subyacente de mayor regularidad. Si los algoritmos logran anticipar comportamientos complejos con una precisión notable, ello sugiere la existencia de principios estructurales que, aunque inaccesibles a la intuición humana, se manifiestan en la consistencia de los datos observables (Wolfram, 2002). Desde esta perspectiva, el universo podría concebirse como un sistema computacional cuya evolución responde a reglas informacionales, incluso si dichas reglas exceden las formulaciones científicas convencionales.

La hipótesis de la simulación amplifica esta interpretación al proponer que la realidad misma podría constituir una instancia computacional creada por una civilización tecnológicamente avanzada (Bostrom, 2003). Si la generación de simulaciones realistas resulta factible y recurrente, la probabilidad de que la experiencia humana ocurra dentro de una de ellas se vuelve estadísticamente relevante. Desde este enfoque, la creciente capacidad predictiva de la IA podría ser comprendida como un indicio de que la realidad opera conforme a principios computacionales. Sin embargo, tal interpretación enfrenta objeciones epistemológicas y ontológicas significativas, especialmente en lo que respecta al carácter especulativo de la hipótesis y a las dificultades para establecer criterios de verificación empírica en un contexto de simulación total (Chalmers, 2005).

De manera complementaria, la teoría de los sistemas complejos introduce un marco interpretativo que permite comprender la predictibilidad algorítmica sin recurrir al determinismo ontológico. Desde esta perspectiva, las regularidades que permiten la predicción emergen de dinámicas colectivas que no pueden reducirse a las propiedades elementales de los componentes, sino que resultan de interacciones no lineales en niveles superiores de organización (Mitchell, 2009). Así, la eficacia predictiva de la IA puede entenderse como la

identificación de patrones emergentes estables, más que como la revelación de un orden causal absoluto. En consecuencia, la posibilidad de anticipar comportamientos macroscópicos depende de la coherencia estadística que se manifiesta en tales niveles de organización, no de un conocimiento exhaustivo de los mecanismos microscópicos (DeLanda, 2006).

Por otra parte, la noción de causalidad adquiere un carácter problemático en este escenario. La tradición filosófica ha distinguido entre la causalidad como relación necesaria entre eventos y la causalidad entendida como regularidad estadística (Lewis, 1973). Los sistemas de IA se sitúan principalmente en el segundo sentido, pues identifican correlaciones significativas sin establecer vínculos causales en el sentido metafísico fuerte. En consecuencia, surge la posibilidad de que la ciencia contemporánea esté transitando desde un paradigma causal hacia uno correlacional, donde la explicación científica se sustituye por la detección de patrones estadísticos (Pearl, 2009). Tal desplazamiento implicaría una reconfiguración profunda de las bases epistémicas de la ciencia, dado que la causalidad ha constituido tradicionalmente el fundamento de la explicación y la intervención técnica.

A partir de este contexto, la ontología agencial ofrece una alternativa que redefine la relación entre conocimiento y realidad. Se sostiene que la realidad no consiste en entidades preexistentes con propiedades intrínsecas, sino en fenómenos que emergen de interacciones intra-activas en las cuales los instrumentos, las teorías y los objetos materiales se configuran de manera recíproca (Barad, 2007). Bajo esta concepción, los modelos de IA no representarían pasivamente un mundo independiente, sino que participarían en la producción misma de los fenómenos que buscan describir. En consecuencia, la estructura predicha por AlphaFold podría entenderse como una configuración emergente de la interacción entre secuencias proteicas, algoritmos de predicción, bases de datos y prácticas científicas, más que como una representación directa de una realidad preexistente.

De esta manera, la pregunta por el determinismo adquiere una nueva formulación. En lugar de interrogar si el mundo es determinista en sí mismo, resulta más pertinente analizar cómo las prácticas científicas producen regímenes de predictibilidad que configuran la realidad de modos específicos. Las técnicas de IA no descubren simplemente regularidades preexistentes, sino que instauran espacios de anticipación en los cuales ciertos futuros se vuelven calculables y, por tanto, susceptibles de intervención (Mackenzie, 2005). Esta capacidad de producción de futuros posibles conlleva efectos materiales y políticos concretos, ya que las predicciones orientan decisiones científicas, económicas y biomédicas.

No obstante, la interpretación performativa del conocimiento enfrenta el desafío de no caer en un construccionismo absoluto que niegue la existencia de un orden material que resiste las construcciones teóricas. La efectividad predictiva de los modelos depende en última instancia de su ajuste con las restricciones impuestas por la estructura física del mundo. Las predicciones de AlphaFold, por ejemplo, son posibles porque las secuencias de aminoácidos están sujetas a limitaciones termodinámicas y cuánticas que determinan configuraciones tridimensionales estables (Hacking, 1999). En este sentido, la performatividad del conocimiento se encuentra delimitada por la resistencia de lo real, que impone condiciones de posibilidad a la eficacia técnica y epistémica.

Desde esta comprensión, el determinismo tecnocientífico no se resuelve en términos ontológicos, sino prácticos. Los sistemas de IA instauran una forma de determinismo operativo en la que ciertos futuros son anticipables y, por tanto, gestionables, mientras que otros permanecen contingentes e inaccesibles (Amoore, 2013). La predictibilidad se convierte así en una tecnología de gobierno del futuro, en la que la contingencia es reducida mediante la modelización algorítmica. Sin implicar necesariamente un determinismo universal, esta lógica produce efectos equivalentes al restringir el campo de lo posible a lo calculable.

La distinción entre epistemología y ontología se difumina en este horizonte. La IA no solo describe el mundo, sino que lo co-constituye mediante la instauración de regímenes de predictibilidad que definen lo que puede ser conocido, previsto o intervenido (Amoore, 2020). En consecuencia, la capacidad de anticipar patrones moleculares, sociales o económicos no representa una mera descripción de la realidad, sino una práctica que participa activamente en su configuración. De ahí que la cuestión del determinismo contemporáneo deba entenderse no como una propiedad intrínseca del mundo, sino como el resultado de un entramado tecnocientífico que produce la realidad bajo la forma de lo predecible.

La reformulación contemporánea del determinismo como régimen de predictibilidad tecnológica conduce a interrogar la autonomía de los procesos de inferencia. Si la realidad puede modelarse mediante correlaciones emergentes sin referencia causal, la producción del conocimiento deja de depender del sujeto humano y pasa a residir en arquitecturas de cálculo autorreferenciales. Tal desplazamiento abre el debate sobre la autonomía epistémica de los algoritmos y sobre la posibilidad de una comprensión científica sin inteligibilidad humana, cuestión que estructura la sección siguiente.

4 AUTONOMÍA ALGORÍTMICA Y PARADOJA EPISTÉMICA DE LA COMPRENSIÓN SIN INTELIGIBILIDAD

La transformación de la racionalidad científica en el contexto de la IA redefine los fundamentos del conocimiento al desplazar la centralidad del sujeto epistémico humano. Tradicionalmente, la epistemología moderna ha concebido un sujeto racional que, mediante la observación empírica y la deducción lógica, produce conocimiento objetivo sobre el mundo (Descartes, 1637; Kant, 1781). Sin embargo, la emergencia de sistemas de IA capaces de generar conocimiento y formular hipótesis sin mediación humana directa altera esa estructura antropocéntrica, al instituir formas de producción cognitiva que operan más allá de la comprensión y control del sujeto humano (Hayles, 1999). Este desplazamiento exige reconsiderar el estatuto ontológico y epistémico del conocimiento científico en una era donde la agencia cognitiva se distribuye entre humanos y máquinas.

En esta línea, la noción de cognición distribuida proporciona un marco interpretativo para comprender la reorganización contemporánea del conocimiento. Dicha perspectiva sostiene que los procesos cognitivos se extienden a través de redes de agentes, artefactos y contextos materiales que participan en la generación de conocimiento (Hutchins, 1995). La ciencia puede entenderse así como una ecología cognitiva donde instrumentos, lenguajes formales y comunidades de investigación cooperan en la producción de sentido (Nersessian, 2006). La IA intensifica esa distribución al incorporar agentes computacionales que realizan operaciones inferenciales, predictivas y analíticas con niveles de complejidad inalcanzables para la cognición humana individual, generando un entramado de conocimiento híbrido en el que lo humano se articula con lo maquínico en un continuum operativo.

No obstante, la IA introduce una discontinuidad cualitativa respecto de las herramientas científicas tradicionales. Mientras los instrumentos clásicos amplían las capacidades perceptivas sin operar autónomamente, los sistemas algorítmicos contemporáneos generan inferencias independientes de la supervisión humana, configurando resultados que pueden escapar a la comprensión de sus propios diseñadores (Burrell, 2016). Esta opacidad algorítmica ha sido interpretada como una crisis de inteligibilidad del conocimiento científico, dado que los mecanismos que sustentan la producción de saber se tornan intransparentes incluso para quienes los aplican (Stinson, 2022). La comprensión científica, entendida como la capacidad de reconstruir racionalmente los pasos que conducen a una conclusión, se ve erosionada cuando el conocimiento más preciso se obtiene a través de procesos que exceden las capacidades de

explicación humana (Trout, 2002; Lipton, 2018). Se configura así una paradoja epistémica: la maximización de la exactitud puede coincidir con la pérdida de inteligibilidad.

En respuesta a esta paradoja, el pensamiento posthumanista propone una reconfiguración de las categorías epistemológicas. Se plantea que la condición posthumana demanda superar el antropocentrismo que ha delimitado el horizonte de lo cognoscible, reconociendo la emergencia de formas de agencia no humanas en la producción del saber (Braidotti, 2013). Desde esa perspectiva, la IA amplía el campo cognitivo al permitir el acceso a regularidades y patrones que trascienden la capacidad perceptiva o conceptual humana. La alianza entre inteligencia biológica y computacional se concibe, por tanto, como un modo de expansión epistémica antes que como una amenaza a la racionalidad científica. Sin embargo, esta ampliación genera interrogantes sobre la autonomía y la responsabilidad epistémica, dado que el conocimiento producido por sistemas opacos plantea dificultades para atribuir agencia y responsabilidad moral o epistémica (Floridi y Sanders, 2004).

La noción de responsabilidad cognitiva se sustenta en la capacidad del agente para justificar sus afirmaciones mediante razones accesibles y evaluables públicamente (Code, 1987). La producción algorítmica de conocimiento, al operar mediante procesos no interpretables, interrumpe esa estructura justificativa, transfiriendo la confianza desde la comprensión racional hacia la fiabilidad técnica. En consecuencia, la ciencia contemporánea se enfrenta a una mutación en el régimen de la confianza. La tradición científica se ha apoyado en mecanismos de verificación y replicabilidad que permiten la evaluación colectiva del conocimiento (Shapin, 1994). Sin embargo, la opacidad de los sistemas algorítmicos induce formas de confianza ciega, donde los usuarios deben aceptar resultados cuya validación interna resulta inaccesible (O'Neil, 2016). Esa transformación desplaza el eje de la confianza desde la transparencia epistémica hacia la eficacia operativa.

Dicha mutación incide directamente en la legitimidad de la autoridad científica. La autoridad epistémica moderna se ha fundado en la posibilidad de escrutinio público y revisión crítica (Merton, 1973). Cuando el conocimiento se produce mediante procesos que no pueden ser auditados plenamente, el criterio de legitimidad puede deslizarse desde la verificabilidad hacia la utilidad predictiva, generando una forma de tecnocracia epistémica donde la capacidad de cálculo sustituye al argumento racional como principio de validación (Jasanoff, 2003). Este desplazamiento erosiona la dimensión deliberativa de la ciencia, debilitando su función crítica y su carácter público.

A la vez, la promesa de objetividad algorítmica se revela problemática. Si bien los algoritmos parecen garantizar una neutralidad derivada de su formalización matemática, las investigaciones empíricas han demostrado que reproducen y amplifican los sesgos existentes en los datos que los alimentan (Benjamin, 2019). La objetividad de la IA no elimina la carga valorativa, sino que la reconfigura en términos de diseño técnico. Los procesos de selección de datos, codificación de variables y definición de métricas de éxito implican decisiones normativas que moldean la producción del conocimiento (Gillespie, 2014).

Desde la epistemología feminista, ciertos conceptos permiten profundizar en este análisis, pues la objetividad fuerte no implica la eliminación de la perspectiva, sino la explicitación crítica de las condiciones sociales que configuran el conocimiento (Harding, 1991). En este sentido, la opacidad de los sistemas de IA puede interpretarse como una manifestación de objetividad débil, en la medida en que disimula las perspectivas incorporadas bajo una apariencia de neutralidad técnica. Así, una epistemología crítica de la IA exigiría visibilizar los marcos valorativos que atraviesan las decisiones de diseño, los datos de entrenamiento y las métricas de evaluación, entendidos como elementos constitutivos del conocimiento que tales sistemas producen.

De manera complementaria, la noción de conocimiento situado como alternativa a la concepción tradicional de objetividad (Haraway, 1988), reconoce que todo conocimiento emerge desde posiciones parciales, pero redefine la objetividad como responsabilidad respecto de las perspectivas desde las cuales se conoce. Aplicado a la IA, este planteamiento implica comprender la objetividad algorítmica no como eliminación de la perspectiva, sino como ejercicio de transparencia y responsabilidad frente a las condiciones sociales y técnicas que estructuran los sistemas computacionales.

La expansión de la IA reconfigura, además, la noción misma de racionalidad científica. La tradición ilustrada, representada paradigmáticamente por Kant, concibe la razón como facultad universal regida por principios lógicos y metodológicos independientes del contexto (Kant, 1781). Sin embargo, las críticas postmodernas, como las formuladas por Lyotard (1979), subrayan la coexistencia de múltiples racionalidades dependientes de marcos culturales y discursivos específicos. La IA introduce, en este marco, una racionalidad algorítmica sustentada en el procesamiento masivo de datos, la detección de patrones estadísticos y la optimización de funciones mediante aprendizaje automático (Cardon et al., 2018).

Tal racionalidad difiere de la científica tradicional, que privilegia la formulación teórica y la deducción de leyes generales. Los algoritmos de aprendizaje automático operan

identificando regularidades empíricas sin la mediación de teorías explícitas (Boden, 2006). Surge así el interrogante sobre si la racionalidad algorítmica complementa la racionalidad teórica o, por el contrario, representa un retorno a un empirismo inductivo superado por la ciencia moderna.

A su vez, el problema de la explicación científica adquiere relevancia en este contexto. La filosofía de la ciencia ha distinguido modelos como el nomológico-deductivo, centrado en leyes generales, y el causal-mecánico, basado en la identificación de mecanismos subyacentes (Hempel, 1965; Machamer et al., 2000). Los sistemas de IA, aunque capaces de predecir con precisión, suelen carecer de explicaciones que se ajusten a estos modelos. En consecuencia, se han propuesto nociones alternativas, como las explicaciones por reconocimiento de patrones o por analogía estadística (Craver y Darden, 2013).

No obstante, tales aproximaciones implican el riesgo de diluir la diferencia entre predicción y comprensión. Si toda correlación estadística se considera explicación, el concepto pierde densidad epistémica y se reduce a la capacidad predictiva. Este empobrecimiento erosiona la inteligibilidad profunda que caracteriza los avances científicos más significativos (de Regt, 2017). La teoría de la relatividad general de Einstein ejemplifica una comprensión que transforma los marcos conceptuales del conocimiento (Einstein, 1916), algo que los sistemas algorítmicos, centrados en la correlación estadística, difícilmente logran replicar.

Asimismo, la noción de descubrimiento científico adquiere matices inéditos. La tradición filosófica ha entendido el descubrimiento como un proceso creativo guiado por la intuición y la formulación de hipótesis innovadoras (Hanson, 1958). En contraste, la IA opera mediante interpolación estadística, explorando espacios definidos por los datos de entrenamiento (Jordan y Mitchell, 2015). Aunque casos como AlphaFold demuestran la capacidad de la IA para identificar configuraciones proteicas no observadas previamente (Jumper et al., 2021), tales logros permanecen condicionados por los límites del espacio algorítmico y los conjuntos de datos disponibles.

Las implicaciones epistémicas de la IA no pueden separarse de sus dimensiones políticas y éticas. La concentración del poder computacional en un reducido número de corporaciones tecnológicas plantea riesgos para la democratización del conocimiento (Zuboff, 2019). Si el acceso a infraestructuras de gran escala se restringe a actores privilegiados, se consolida una nueva forma de desigualdad epistémica (Eubanks, 2018). Ello afecta la autonomía de la ciencia, históricamente concebida como esfera relativamente independiente de intereses económicos y políticos (Polanyi, 1962). La subordinación de la investigación científica a plataformas

corporativas sugiere una reconfiguración del régimen de producción de conocimiento, en el que los criterios de verdad pueden verse desplazados por los imperativos de rentabilidad (Mirowski, 2011).

Tras lo antes señalado, se puede afirmar que la irrupción de una racionalidad algorítmica autónoma transforma no solo la práctica científica, sino también el estatuto ontológico del conocimiento. Cuando los modelos dejan de representar y comienzan a configurar la realidad mediante su aplicación, la frontera entre descripción y creación se diluye. La siguiente sección profundiza en esa dimensión performativa de la IA, donde la predicción algorítmica se convierte en un acto constitutivo que disuelve la distinción tradicional entre representación y realidad.

5 PERFORMATIVIDAD ALGORÍTMICA: DISOLUCIÓN DE LA FRONTERA ENTRE REPRESENTACIÓN Y REALIDAD

La capacidad predictiva de la IA suscita una reconsideración profunda de la relación entre representación y realidad. Cuando los modelos computacionales generan predicciones que posteriormente se materializan, la distinción entre simulación y mundo observable se desdibuja, ya que la representación deja de ser un mero reflejo y se convierte en un agente constitutivo del fenómeno (Baudrillard, 1994). En tal sentido, el concepto de simulacro caracteriza una condición en la cual las representaciones no siguen a la realidad, sino que la anteceden y configuran (Baudrillard, 1981). La predicción algorítmica, entonces, puede entenderse como un simulacro contemporáneo: sus anticipaciones informacionales moldean las estructuras materiales que después se observan, como ocurre cuando AlphaFold orienta experimentos que terminan confirmando la estructura que había proyectado.

En correspondencia con ello, la sociología de la ciencia ha señalado que las teorías científicas no solo describen el mundo, sino que lo configuran a través de prácticas experimentales y tecnológicas (Pickering, 1995). La performatividad del conocimiento, estudiada por MacKenzie (2006), demuestra que los modelos pueden transformar el comportamiento de los actores que los utilizan, como sucede en los mercados financieros. Sin embargo, la IA intensifica esta performatividad, ya que comprime el ciclo entre predicción y verificación, reduciendo el tiempo entre el modelo y el fenómeno y eliminando espacios de mediación reflexiva.

Esa compresión temporal afecta la ontología del conocimiento. En la ciencia tradicional, la distancia entre hipótesis, verificación y aplicación permitía momentos de deliberación crítica

(Stengers, 2018). Los sistemas de IA, al generar predicciones que se integran inmediatamente en procesos de decisión o intervención, reducen dichos intervalos y promueven un régimen de acción donde la anticipación se transforma en acontecimiento sin mediación suficiente. La realidad informacional adquiere así peso ontológico propio, al constituirse como un espacio de interacción entre entidades algorítmicas, datos y procesos materiales (Floridi, 2014). La infosfera describe precisamente ese entramado donde lo informacional y lo material convergen, configurando una realidad híbrida que difumina las fronteras entre cálculo y fenómeno.

De modo concomitante, la concepción correspondentista de la verdad (Tarski, 1944) enfrenta tensiones frente a la performatividad algorítmica. La idea de que una proposición es verdadera si corresponde a un estado de cosas presupone separación entre lenguaje y mundo. No obstante, cuando las predicciones participan en la constitución de los hechos que las confirman, la verdad adquiere un carácter circular: la confirmación empírica depende de las condiciones experimentales generadas por el propio modelo. Aun así, esta circularidad no elimina las restricciones materiales, pues la realidad impone límites concretos que condicionan qué predicciones pueden verificarse. En consecuencia, la coevolución entre modelos y fenómenos genera una dinámica recursiva donde ambos se configuran mutuamente (Barad, 2007).

Tal dinámica se comprende con mayor profundidad desde una ontología procesual, según la cual la realidad no consiste en sustancias estáticas, sino en procesos de devenir relacional (Whitehead, 1929). Desde esta perspectiva, las predicciones algorítmicas no representan entidades preexistentes, sino que participan en procesos de actualización donde las potencialidades informacionales se concretan en configuraciones empíricas. La predicción no es causa ni efecto, sino un nodo constitutivo en el entramado de relaciones entre prácticas científicas, estructuras materiales y decisiones humanas.

No obstante, la visión procesual plantea el problema de la estabilidad del conocimiento. Si el mundo es un flujo continuo de co-constitución entre teoría y fenómeno, la objetividad científica no puede derivarse de una realidad fija, sino de redes sociotécnicas que estabilizan los hechos (Latour, 1987). La estabilidad epistémica se convierte así en un logro práctico sustentado por infraestructuras materiales, protocolos y consensos institucionales. De tal entramado emerge la noción de cierre epistémico, entendida como el momento en que una comunidad asume que ciertos problemas fundamentales han sido resueltos y orienta su labor hacia aplicaciones tecnológicas (Fleck, 1935). La IA puede acelerar este cierre, al producir la

ilusión de comprensión completa dentro de dominios donde solo se ha alcanzado eficacia predictiva (Kuhn, 1962).

Surge, entonces, el riesgo de un determinismo epistémico. La precisión con que AlphaFold predice estructuras proteicas no implica que el problema del plegamiento esté plenamente comprendido, ya que subsisten cuestiones sobre las leyes físicas subyacentes, la cinética y la termodinámica del proceso (Dill y MacCallum, 2012). La aparente eficacia puede desalentar la exploración teórica y consolidar un paradigma centrado en la predicción instrumental.

Dicha tendencia expresa una tensión más amplia en la ciencia contemporánea entre orientación instrumental y orientación contemplativa. A lo largo de la historia, la ciencia ha oscilado entre la búsqueda de conocimiento por sí mismo y la producción de saber útil para la intervención técnica. En la actualidad, la IA refuerza la primera dirección, privilegiando la predicción operativa sobre la comprensión profunda, lo que podría transformar el ideal de la ciencia en una práctica eminentemente tecnológica (Heidegger, 1977).

En consecuencia, la disolución de la frontera entre representación y realidad en la ciencia contemporánea obliga a reconsiderar el sentido último del conocimiento en la era de la IA. Si la predicción se convierte en un acto performativo que produce los fenómenos que pretende describir, la epistemología ya no puede apoyarse en la correspondencia entre teoría y mundo, sino en la eficacia material de los modelos. Las conclusiones retoman este punto para sintetizar las implicaciones de dicha transformación, evaluando los alcances y límites de la racionalidad algorítmica como nuevo fundamento del saber científico.

6 CONCLUSIONES

La discusión en torno al estatuto ontológico y epistemológico de la IA en la ciencia revela una inflexión decisiva en la comprensión del conocimiento. El desplazamiento desde la verificación empírica hacia la coherencia computacional reconfigura la relación entre verdad, causalidad y libertad epistémica. No se trata únicamente de un cambio de métodos, sino de una mutación en las condiciones mismas de posibilidad del saber. La ciencia, al integrar sistemas algorítmicos autónomos, parece ensayar una nueva forma de racionalidad, donde la anticipación sustituye a la explicación y la simulación adquiere peso ontológico. Desde esta perspectiva, la IA no puede concebirse como un mero instrumento, sino como un agente epistémico que transforma la estructura del descubrimiento científico y redefine el horizonte de la objetividad.

La hipótesis que orientó el presente análisis plantea que la IA instituye un régimen ontológico de carácter determinista donde la simulación y la predicción adquieren primacía sobre la experiencia empírica. El análisis realizado confirma parcialmente esta hipótesis, pero requiere matizaciones significativas. La IA efectivamente instituye un régimen de predictibilidad que opera mediante la identificación de regularidades estadísticas en espacios de alta dimensionalidad, lo cual genera una forma de determinismo práctico donde ciertos futuros se tornan calculables y, por tanto, gestionables. Sin embargo, este determinismo práctico no presupone necesariamente un determinismo ontológico en el sentido metafísico fuerte, sino que refleja la institución de regímenes sociotécnicos que producen predictibilidad mediante la estabilización de infraestructuras computacionales, conjuntos de datos y prácticas estandarizadas.

El análisis permite identificar una convergencia entre ontología procesual y epistemología performativa. Los modelos algorítmicos no se limitan a representar el mundo, sino que participan activamente en su configuración. La predicción computacional deja de ser una operación derivada de teorías preexistentes y se convierte en una forma de producción ontológica, donde lo posible adquiere consistencia material a través de la inferencia estadística. En este sentido, la noción de determinismo tecnocientífico no debe entenderse como retorno al determinismo laplaciano, sino como instauración de un régimen de predictibilidad operativa que delimita lo que puede ser conocido e intervenido. Tal régimen rearticula la relación entre sujeto y objeto, desplazando el centro de la agencia cognitiva hacia ensamblajes híbridos humano-máquina.

Las transformaciones epistémicas generadas por la IA poseen consecuencias concretas en la práctica científica. La legitimación del conocimiento basada en la eficacia predictiva modifica los criterios de validación, promueve formas de confianza técnica no sustentadas en comprensión racional y redefine la temporalidad del descubrimiento al acelerar los ciclos de hipótesis y verificación. Estas condiciones generan riesgos de tecnocratización del saber, en los que la autoridad algorítmica desplaza la deliberación crítica. No obstante, también abren la posibilidad de una ciencia colaborativa que reconozca la distribución de la agencia cognitiva y la necesidad de marcos éticos capaces de asegurar transparencia, trazabilidad y responsabilidad epistémica. La gestión del conocimiento algorítmico exige, por tanto, políticas de apertura y evaluación pública que equilibren eficacia técnica y legitimidad racional.

La investigación futura debería profundizar en la articulación entre racionalidad algorítmica y racionalidad teórica, explorando modos de integración que eviten tanto el

reduccionismo instrumental como el rechazo tecnófobo. Resulta urgente desarrollar una epistemología crítica capaz de discernir entre correlación operativa y comprensión causal, sin renunciar al ideal de inteligibilidad. Asimismo, el examen de la performatividad algorítmica demanda un enfoque interdisciplinario que incorpore filosofía de la ciencia, ética computacional y estudios sociales de la tecnología. Una agenda fecunda podría orientarse a delinear los límites de la autonomía epistémica de los algoritmos, los criterios de responsabilidad en contextos de opacidad y los nuevos marcos normativos que garanticen una ciencia abierta al escrutinio y a la reflexión colectiva.

La emergencia de la IA como agente de conocimiento anuncia un cambio epocal en la historia de la ciencia. La sustitución parcial de la verificación empírica por la coherencia computacional no constituye una mera evolución metodológica, sino una transformación del imaginario epistemológico moderno. La verdad ya no se concibe como correspondencia entre teoría y mundo, sino como coherencia performativa entre predicción y acontecimiento. Sin embargo, frente a la tentación de un determinismo tecnocientífico, la tarea filosófica consiste en preservar espacios de indeterminación, crítica y libertad epistémica. La ciencia del futuro, si aspira a mantenerse humana en sentido profundo, deberá reconciliar la potencia del cálculo con la necesidad de comprender, deliberar y resistir la clausura del pensamiento bajo la lógica de lo predecible.

PIRES, R. F. M.; BENTES, H. de V. Evasão escolar na modalidade subsequente na educação profissional e tecnológica: As (contra) reformas nas políticas educacionais no Brasil e seus impactos na Educação Profissional e Tecnológica. *In: 2º CONGRESSO NACIONAL DE EDUCAÇÃO (CONEDU)*, 2024. **Anais**. Campina Grande: Realize Eventos, 2024. p. 59-67. Disponível em: https://www.editorarealize.com.br/educacao/anais-ix-conedu/pesquisa?autor=&titulo=evas%C3%A3o+escolar&area_tematica=. Acesso em: 10 jun 2025

REFERENCIAS

- AMOORE, L. (2013). *The politics of possibility: Risk and security beyond probability*. Duke University Press. <https://www.dukeupress.edu/the-politics-of-possibility>
- AMOORE, L. (2020). *Cloud ethics: Algorithms and the attributes of ourselves and others*. Duke University Press. <https://www.dukeupress.edu/cloud-ethics>

- ANDERSON, C. (2008). The end of theory: The data deluge makes the scientific method obsolete. *Wired Magazine*, 16(7), 16-07. <https://www.wired.com/2008/06/pb-theory/>
- BARAD, K. (2007). *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning*. Duke University Press. <https://www.dukeupress.edu/meeting-the-universe-halfway>
- BAUDRILLARD, J. (1981). *Simulacra and simulation*. University of Michigan Press. <https://www.press.umich.edu/>
- BAUDRILLARD, J. (1994). *Simulacra and simulation* (S. F. Glaser, Trans.). University of Michigan Press. <https://www.press.umich.edu/>
- BENJAMIN, R. (2019). *Race after technology: Abolitionist tools for the new Jim Code*. Polity Press. <https://politybooks.com/bookdetail/?isbn=9781509526406>
- BODEN, M. A. (2006). *Mind as machine: A history of cognitive science*. Oxford University Press. <https://global.oup.com/academic/product/mind-as-machine-9780199543533>
- BOSTROM, N. (2003). Are we living in a computer simulation? *The Philosophical Quarterly*, 53(211), 243-255. <https://doi.org/10.1111/1467-9213.00309>
- BRAIDOTTI, R. (2013). *The posthuman*. Polity Press. <https://politybooks.com/bookdetail/?isbn=9780745641584>
- BUCKNER, C. (2018). Empiricism without magic: Transformational abstraction in deep convolutional neural networks. *Synthese*, 195(12), 5339-5372. <https://doi.org/10.1007/s11229-018-01949-1>
- BURRELL, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1-12. <https://doi.org/10.1177/2053951715622512>
- CARDON, D., Cointet, J. P., & Mazières, A. (2018). Neurons spike back: The invention of inductive machines and the artificial intelligence controversy. *Réseaux*, 211(5), 173-220. <https://doi.org/10.3917/res.211.0173>
- CARTWRIGHT, N. (1983). *How the laws of physics lie*. Oxford University Press. <https://global.oup.com/academic/product/how-the-laws-of-physics-lie-9780198247043>
- CHALMERS, D. J. (2005). The matrix as metaphysics. In C. Grau (Ed.), *Philosophers explore the matrix* (pp. 132-176). Oxford University Press. <https://global.oup.com/academic/>
- CLARK, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7-19. <https://doi.org/10.1093/analys/58.1.7>
- CODE, L. (1987). *Epistemic responsibility*. University Press of New England. <https://www.upne.com/>
- CRAVER, C. F., & Darden, L. (2013). *In search of mechanisms: Discoveries across the life sciences*. University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/I/bo15879711.html>
- DE REGT, H. W. (2017). *Understanding scientific understanding*. Oxford University Press. <https://global.oup.com/academic/product/understanding-scientific-understanding-9780190652913>

- DELANDA, M. (2006). *A new philosophy of society: Assemblage theory and social complexity*. Continuum. <https://www.bloomsbury.com/>
- DESCARTES, R. (1637). *Discourse on the method*. Leiden. <https://plato.stanford.edu/entries/descartes-works/>
- DILL, K. A., & MacCallum, J. L. (2012). The protein-folding problem, 50 years on. *Science*, 338(6110), 1042-1046. <https://doi.org/10.1126/science.1219021>
- D'IGNAZIO, C., & Klein, L. F. (2020). *Data feminism*. MIT Press. <https://mitpress.mit.edu/9780262044004/data-feminism/>
- EARMAN, J. (1986). *A primer on determinism*. Reidel. <https://www.springer.com/>
- EINSTEIN, A. (1916). The foundation of the general theory of relativity. *Annalen der Physik*, 354(7), 769-822. <https://doi.org/10.1002/andp.19163540702>
- EUBANKS, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press. <https://us.macmillan.com/books/9781250074317>
- FLECK, L. (1935). *Genesis and development of a scientific fact*. University of Chicago Press. <https://press.uchicago.edu/>
- FLORIDI, L. (2011). *The philosophy of information*. Oxford University Press. <https://global.oup.com/academic/product/the-philosophy-of-information-9780199232383>
- FLORIDI, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press. <https://global.oup.com/academic/product/the-fourth-revolution-9780199606726>
- FLORIDI, L., & Sanders, J. W. (2004). On the morality of artificial agents. *Minds and Machines*, 14(3), 349-379. <https://doi.org/10.1023/B:MIND.0000035461.63578.9d>
- GILLESPIE, T. (2014). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies: Essays on communication, materiality, and society* (pp. 167-194). MIT Press. <https://mitpress.mit.edu/>
- GOODFELLOW, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press. <https://mitpress.mit.edu/9780262035613/deep-learning/>
- HACKING, I. (1983). *Representing and intervening*. Cambridge University Press. <https://www.cambridge.org/core/books/representing-and-intervening/ED17FF87198728DF10EB5E6D29EAD5C2>
- HACKING, I. (1999). *The social construction of what?* Harvard University Press. <https://www.hup.harvard.edu/catalog.php?isbn=9780674812000>
- HANSON, N. R. (1958). *Patterns of discovery*. Cambridge University Press. <https://www.cambridge.org/core/>
- HARAWAY, D. J. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, 14(3), 575-599. <https://doi.org/10.2307/3178066>
- HARDING, S. (1991). *Whose science? Whose knowledge? Thinking from women's lives*. Cornell University Press. <https://www.cornellpress.cornell.edu/>

- HAYLES, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/H/bo3773963.html>
- HEIDEGGER, M. (1977). *The question concerning technology and other essays*. Harper & Row. <https://www.harpercollins.com/>
- HEMPEL, C. G. (1965). *Aspects of scientific explanation*. Free Press. <https://www.simonandschuster.com/>
- HOEFER, C. (2016). Causal determinism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition). <https://plato.stanford.edu/archives/spr2016/entries/determinism-causal/>
- HUMPHREYS, P. (2009). The philosophical novelty of computer simulation methods. *Synthese*, 169(3), 615-626. <https://doi.org/10.1007/s11229-008-9435-2>
- HUTCHINS, E. (1995). *Cognition in the wild*. MIT Press. <https://mitpress.mit.edu/9780262581462/cognition-in-the-wild/>
- JASANOFF, S. (2003). Technologies of humility: Citizen participation in governing science. *Minerva*, 41(3), 223-244. <https://doi.org/10.1023/A:1025557512320>
- JORDAN, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260. <https://doi.org/10.1126/science.aaa8415>
- JUMPER, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583-589. <https://doi.org/10.1038/s41586-021-03819-2>
- KANT, I. (1781). *Critique of pure reason*. Cambridge University Press. <https://www.cambridge.org/core/>
- KITCHIN, R. (2014). Big data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 1-12. <https://doi.org/10.1177/2053951714528481>
- KNORR-CETINA, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Harvard University Press. <https://www.hup.harvard.edu/>
- KUHLMAN, B., Dantas, G., Ireton, G. C., Varani, G., Stoddard, B. L., & Baker, D. (2003). Design of a novel globular protein fold with atomic-level accuracy. *Science*, 302(5649), 1364-1368. <https://doi.org/10.1126/science.1089427>
- KUHN, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/S/bo13179781.html>
- LAPLACE, P. S. (1814). *A philosophical essay on probabilities*. Dover Publications. <https://store.doverpublications.com/>
- LATOURET, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press. <https://www.hup.harvard.edu/>
- LATOURET, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press. <https://global.oup.com/academic/product/reassembling-the-social-9780199256044>

- LATOUR, B., & Woolgar, S. (1979). *Laboratory life: The construction of scientific facts*. Princeton University Press. <https://press.princeton.edu/books/paperback/9780691028323/laboratory-life>
- LECUN, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- LEWIS, D. (1973). Causation. *The Journal of Philosophy*, 70(17), 556-567. <https://doi.org/10.2307/2025310>
- LIPTON, Z. C. (2018). The mythos of model interpretability. *Communications of the ACM*, 61(10), 36-43. <https://doi.org/10.1145/3233231>
- LYOTARD, J. F. (1979). *The postmodern condition: A report on knowledge*. University of Minnesota Press. <https://www.upress.umn.edu/>
- MACHAMER, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, 67(1), 1-25. <https://doi.org/10.1086/392759>
- MACKENZIE, D. (2005). Opening the black boxes of global finance. *Review of International Political Economy*, 12(4), 555-576. <https://doi.org/10.1080/09692290500240271>
- MACKENZIE, D. (2006). *An engine, not a camera: How financial models shape markets*. MIT Press. <https://mitpress.mit.edu/9780262633673/an-engine-not-a-camera/>
- MERTON, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press. <https://press.uchicago.edu/>
- MIROWSKI, P. (2011). *Science-mart: Privatizing American science*. Harvard University Press. <https://www.hup.harvard.edu/>
- MITCHELL, M. (2009). *Complexity: A guided tour*. Oxford University Press. <https://global.oup.com/academic/product/complexity-9780195124415>
- MORGAN, M. S., & Morrison, M. (1999). *Models as mediators*. Cambridge University Press. <https://www.cambridge.org/core/books/models-as-mediators/4B402D87C80CAC0C95E7D548E6F2CAA0>
- NERSESSIAN, N. J. (2006). The cognitive-cultural systems of the research laboratory. *Organization Studies*, 27(1), 125-145. <https://doi.org/10.1177/0170840605054610>
- NOBLE, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press. <https://nyupress.org/9781479837243/algorithms-of-oppression/>
- O'NEIL, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown. <https://www.penguinrandomhouse.com/books/241363/weapons-of-math-destruction-by-cathy-oneil/>
- PASQUALE, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press. <https://www.hup.harvard.edu/catalog.php?isbn=9780674368279>
- PEARL, J. (2009). *Causality: Models, reasoning, and inference* (2nd ed.). Cambridge University Press. <https://www.cambridge.org/core/books/causality/B0046844FAE10CBF274D4ACBDAEB5F5B>

- PICKERING, A. (1995). *The mangle of practice: Time, agency, and science*. University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/M/bo3645448.html>
- POLANYI, M. (1962). The republic of science: Its political and economic theory. *Minerva*, 1(1), 54-73. <https://doi.org/10.1007/BF01101453>
- POPPER, K. R. (1959). *The logic of scientific discovery*. Hutchinson. <https://www.routledge.com/>
- PRIGOGINE, I. (1997). *The end of certainty: Time, chaos, and the new laws of nature*. Free Press. <https://www.simonandschuster.com/>
- RHEINBERGER, H. J. (1997). *Toward a history of epistemic things: Synthesizing proteins in the test tube*. Stanford University Press. <https://www.sup.org/books/title/?id=2895>
- ROSA, H. (2013). *Social acceleration: A new theory of modernity*. Columbia University Press. <https://cup.columbia.edu/book/social-acceleration/9780231148344>
- SALMON, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton University Press. <https://press.princeton.edu/>
- SCHMIDT, M., & Lipson, H. (2009). Distilling free-form natural laws from experimental data. *Science*, 324(5923), 81-85. <https://doi.org/10.1126/science.1165893>
- SENIOR, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., ... & Hassabis, D. (2020). Improved protein structure prediction using potentials from deep learning. *Nature*, 577(7792), 706-710. <https://doi.org/10.1038/s41586-019-1923-7>
- SHAPIN, S. (1994). *A social history of truth: Civility and science in seventeenth-century England*. University of Chicago Press. <https://press.uchicago.edu/ucp/books/book/chicago/S/bo3684740.html>
- STENGERS, I. (2018). *Another science is possible: A manifesto for slow science*. Polity Press. <https://politybooks.com/bookdetail/?isbn=9781509521814>
- STINSON, C. (2022). From implausibility to unreliability: Fixing the semantics of artificial neural networks. *Minds and Machines*, 32(1), 81-104. <https://doi.org/10.1007/s11023-021-09589-4>
- STROGATZ, S. H. (2015). *Nonlinear dynamics and chaos: With applications to physics, biology, chemistry, and engineering* (2nd ed.). Westview Press. <https://www.hachettebookgroup.com/>
- TARSKI, A. (1944). The semantic conception of truth and the foundations of semantics. *Philosophy and Phenomenological Research*, 4(3), 341-376. <https://doi.org/10.2307/2102968>
- TROUT, J. D. (2002). Scientific explanation and the sense of understanding. *Philosophy of Science*, 69(2), 212-233. <https://doi.org/10.1086/341050>
- WHITEHEAD, A. N. (1929). *Process and reality*. Macmillan. <https://www.simonandschuster.com/>
- WOLFRAM, S. (2002). *A new kind of science*. Wolfram Media. <https://www.wolframscience.com/nks/>

WOODWARD, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press. <https://global.oup.com/academic/product/making-things-happen-9780195155273>

ZUBOFF, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs. <https://www.hachettebookgroup.com/titles/shoshana-zuboff/the-age-of-surveillance-capitalism/9781610395694/>

Submissão em: 20/08/2025

Aprovação em: 20/08/2025