

<https://doi.org/10.21555/top.v760.3580>

Simulación y justificación en la era del *deep learning*: AlphaFold como ejemplar paradigmático de modelos opacos

Simulation and Justification in the Age of Deep Learning: AlphaFold as a Paradigmatic Case of Opaque Models

Erick Manuel Rubio
Fundación Universitaria San Martín
Colombia
Universidad del Norte
Colombia
erick.rubio@sanmartin.edu.co
<https://orcid.org/0009-0005-1900-2369>

Leandro Giri
Universidad Nacional de Tres de Febrero
Argentina
Consejo Nacional de Investigaciones
Científicas y Técnicas
Argentina
leandrogiri@gmail.com
<https://orcid.org/0000-0002-7068-9750>

Miguel Fuentes
Consejo Nacional de Investigaciones
Científicas y Técnicas
Argentina
fuentesmig@gmail.com
<https://orcid.org/0000-0003-0833-1245>

Elkin Navarro
Fundación Universitaria San Martín
Colombia
elkin.navarro@sanmartin.edu.co
<https://orcid.org/0000-0001-7567-6409>

Recibido: 25 - 09 - 2025.

Aceptado: 27 - 02 - 2026.

Publicado en línea: 10 - 07 - 2026.

Cómo citar este artículo: Rubio, E. M., Giri, L., Fuentes, M. y Navarro, E. (2026). Simulación y justificación en la era del *deep learning*: AlphaFold como ejemplar paradigmático de modelos opacos. *Tópicos, Revista de Filosofía*, 76, 283-313. <https://doi.org/10.21555/top.v760.3580>



This work is licensed under a Creative Commons Attribution
-NonCommercial-ShareAlike 4.0 International License.

Resumen

Este artículo examina el estatuto epistémico de los modelos computacionales contemporáneos basados en *deep learning* a partir del caso de AlphaFold, con el objetivo de evaluar en qué medida estos sistemas pueden comprenderse dentro del marco tradicional de las simulaciones computacionales o si requieren categorías epistémicas diferenciadas. A partir del análisis del marco propuesto por Paul Humphreys —en particular, las nociones de “plantilla” y “modelo computacional”— se explora la posibilidad de una continuidad formal y epistémica entre las simulaciones clásicas y los sistemas de aprendizaje profundo, atendiendo críticamente a los problemas de opacidad, verificación y validación pragmática, y al papel constitutivo de la representación visual. Se muestra que, si bien las concepciones representacionistas clásicas resultan insuficientes para dar cuenta del funcionamiento interno de modelos como AlphaFold, estos sistemas tampoco pueden reducirse sin más a meros dispositivos algorítmicos. En este sentido, el artículo argumenta que AlphaFold sitúa a la filosofía de la ciencia ante la necesidad de repensar los criterios de justificación epistémica y el alcance mismo de la noción de “simulación” en el contexto del *deep learning*.

Palabras clave: *deep learning*; simulaciones de *deep learning* epistémicamente opacas; aprendizaje profundo; opacidad epistémica; representación científica; validación; verificación; justificación; simulación computacional; modelo computacional; plantilla computacional.

Abstract

This paper examines the epistemic status of contemporary computational models based on deep learning through the case of AlphaFold, with the aim of assessing to what extent such systems can be understood within the traditional framework of computational simulations or whether they require distinct epistemic categories. Drawing on Paul Humphreys' analytical framework—particularly the notions of computational templates and computational models—the paper explores the possibility of a formal and epistemic continuity between classical simulations and deep learning systems, while critically addressing issues of opacity, pragmatic validation, and the constitutive role of visual representation. The analysis shows that, although classical representationalist accounts are insufficient to capture the internal functioning of models such as AlphaFold, these systems cannot be straightforwardly reduced to mere algorithmic devices. In this sense, it is argued that AlphaFold confronts philosophy of science with the need to rethink both the criteria of epistemic justification and the scope of the notion of simulation in the era of deep learning.

Keywords: deep learning; epistemically opaque deep-learning simulations; epistemic opacity; scientific representation; validation; verification; justification; computer simulation; computational model; computational template.

Introducción

Las simulaciones computacionales se han consolidado como una herramienta científica de fuste, surgiendo y afianzándose notablemente a partir de la Segunda Guerra Mundial (Galison, 1996; Winsberg y Alvarado, 2022). En ese período, su impacto fue decisivo en campos como la meteorología, con figuras pioneras como Edward Lorenz y su Royal McBee LGP-30,¹ y en la física nuclear, donde matemáticos como John von Neumann realizaron aportes determinantes en el marco del Proyecto Manhattan, especialmente en el desarrollo de armas nucleares. De hecho, fue en este contexto donde el método de simulación de Monte Carlo, propuesto en gran medida por Stanislaw Ulam, tuvo su génesis en la investigación científica, marcando una nueva forma de trabajar con la aleatoriedad como pilar fundamental. Para Peter Galison (1996, p. 119), las simulaciones computacionales, particularmente el método de Monte Carlo, transformaron las tradiciones de investigación existentes —teórica, experimental e instrumental (Lenhard, 2016, pp. 729 y 736)— y sentaron las bases para un nuevo modo de producción de conocimiento científico.

Desde entonces, se ha vuelto indispensable en un número creciente de disciplinas. La lista de ciencias que hacen un uso extensivo de la simulación por computadora ha crecido hasta incluir la astrofísica, la física de partículas, la ciencia de materiales, la ingeniería, la mecánica de fluidos, la ciencia del clima, la biología evolutiva, la ecología, la economía, la teoría de la decisión, la medicina, la sociología, la epidemiología y muchas otras. Incluso hay algunas disciplinas, como la teoría del caos y la teoría de la complejidad, cuya propia existencia ha surgido junto con el desarrollo de

¹ El Royal McBee LGP-30, adquirido por Lorenz en 1960, constituyó una de las primeras computadoras de propósito general aplicadas al modelado atmosférico; permitió formalizar el estudio de la sensibilidad a condiciones iniciales en sistemas dinámicos no lineales.

los modelos computacionales que estudian (Winsberg y Alvarado, 2022).²

Se consideran no solo instrumentos, sino métodos que han abierto nuevas vías para comprender el mundo y sus fenómenos (Durán, 2017, p. 125). En muchos contextos, el modelado computacional se ha consolidado como una herramienta insustituible para explorar escenarios complejos que desafían el análisis experimental o lógico-formal (Vicsek, 2002). Este tipo de modelado a menudo se enmarca en lo que la literatura reciente denomina “modelado modal”, que busca proveer información sobre lo posible, lo contrafactual o las potencialidades de sistemas; sirve para la exploración de hipótesis y la comprensión de fenómenos que aún no son actuales (Sjölin Wirling y Grüne-Yanoff, 2021). Así, la simulación computacional, desde sus inicios, no solo amplía el alcance del conocimiento científico hacia dominios tradicionalmente inaccesibles, sino que también desafía a la filosofía de la ciencia a repensar las formas de justificación y explicación. Este trasfondo permite comprender por qué la aparición de sistemas como AlphaFold representa no solo una continuidad en la sofisticación y centralidad del modelado computacional, sino también una transformación cualitativa en los modos de producción y validación del conocimiento científico.

Un ejemplo reciente y de gran relevancia de la capacidad predictiva y transformadora de los sistemas de simulación es AlphaFold, un sistema de inteligencia artificial desarrollado por DeepMind que, mediante arquitecturas de tipo *deep learning* basadas en *transformers*,³ predice

² A menos que la bibliografía final indique lo contrario, las traducciones son nuestras.

³ El *transformer* es una arquitectura de red neuronal para modelos de transducción de secuencias, como la traducción automática, que se basa únicamente en mecanismos de atención para establecer dependencias globales entre la entrada (*input*) y la salida (*output*). A diferencia de modelos anteriores, el *transformer* prescinde por completo de la recurrencia y las convoluciones. Su estructura principal consiste en un codificador (*encoder*) y un decodificador (*decoder*) apilados, ambos contruidos con capas de “atención de múltiples cabezales” (*multi-head self-attention*) y redes neuronales de avance (*feed-forward*). Esta arquitectura permite un mayor grado de paralelización en el entrenamiento y reduce significativamente el tiempo necesario para aprender dependencias a larga distancia dentro de las secuencias (Vaswani *et al.*, 2017).

la estructura tridimensional de proteínas a partir de su secuencia de aminoácidos.

En 2024, el Premio Nobel de Química fue otorgado a Demis Hassabis y John Jumper por el desarrollo precisamente de AlphaFold, y a David Baker por sus contribuciones al diseño computacional de proteínas. La importancia científica de AlphaFold radica en que ha alcanzado, en numerosos casos, niveles de precisión comparables a los obtenidos mediante técnicas experimentales consolidadas como la cristalografía de rayos X o la criomicroscopía electrónica, resolviendo así, en gran medida, un problema abierto durante más de cinco décadas en biología estructural (Jumper *et al.*, 2021).

La posibilidad de abordar sistemas complejos mediante simulaciones computacionales permite identificar patrones y comportamientos replicables, incluso cuando se trata de conjuntos masivos de datos (Chirimuuta, 2021). Sin embargo, el caso de AlphaFold marca un punto de inflexión. Ya no se trata únicamente de representar dinámicamente procesos físicos conocidos, sino de generar conocimiento a partir de la detección de correlaciones empíricas que no necesariamente están mediadas por teorías explícitas o modelos representacionales. En este sentido, AlphaFold pone en tensión los marcos clásicos con los que la filosofía de la ciencia ha comprendido los modelos, la representación y la justificación epistémica.

A primera vista, podría parecer más apropiado considerarlo una nueva clase de artefactos científicos: opacos, no representacionales, pero epistémicamente fértiles. En este punto, se puede establecer un paralelo con el modelado modal, donde los modelos a menudo no satisfacen los criterios estándar de precisión representacional, pero aun así ofrecen un conocimiento epistémico valioso sobre las posibilidades (Sjölin Wirling y Grüne-Yanoff, 2021). AlphaFold, al predecir estructuras, puede entenderse como un medio para explorar lo que es objetivamente posible en el plegamiento de proteínas, incluso en ausencia de una teoría causal explícita.

No obstante, también podría sostenerse —como alternativa que será examinada críticamente más adelante— que casos como AlphaFold invitan a repensar la noción de “simulación” en un sentido formal-computacional. Esta lectura, sin embargo, no se apoya en la mera presencia de estructuras matemáticas o en la algoritmicidad del sistema, ya que estos rasgos, por sí solos, resultarían insuficientes para distinguir una simulación de otros dispositivos computacionales. Más bien, lo que

estaría en juego es si la resolución algorítmica de ciertas estructuras, integrada en prácticas de validación y uso experimental, puede cumplir funciones epistémicas análogas a las tradicionalmente atribuidas a las simulaciones. En este artículo nos ocuparemos de AlphaFold no como un objeto aislado, sino como un caso paradigmático de una nueva clase de artefactos computacionales opacos sustentados en redes neuronales profundas, que denominaremos EODS (*Epistemically Opaque Deep-learning Simulations*). La cuestión de si estos sistemas pueden entenderse mediante la noción clásica de “simulación” o no será objeto de análisis a lo largo del trabajo.

Casos análogos —como los grandes modelos de lenguaje (por ejemplo, ChatGPT) o los generadores de imágenes (p. ej., DALL-E)— sugieren que lo que está en juego no se restringe a la biología estructural; porque comparten con AlphaFold un mismo perfil epistémico, en tanto que producen resultados empíricamente fiables mediante arquitecturas algorítmicas opacas, carecen de una teoría mediadora explícita del fenómeno que tratan y se integran a la práctica científica (o técnico-científica) a través de criterios de validación pragmática y de interfaces representacionales diseñadas para el uso humano. En este sentido, estos sistemas señalan una transformación más general en la arquitectura epistémica de la ciencia contemporánea, en la que la producción de conocimiento se apoya crecientemente en sistemas algorítmicos opacos cuya fiabilidad se establece por vías distintas de la comprensión causal tradicional.

Otro ejemplo potente se encuentra en la química cuántica y el diseño de fármacos. Tradicionalmente, predecir las propiedades de una molécula (como su energía de enlace o su reactividad) requiere costosas simulaciones basadas en primeros principios, como la teoría del funcional de la densidad (DFT) (Jensen, 2017). Hoy en día, redes neuronales de grafos (GNNs) son entrenadas con enormes bases de datos de estructuras moleculares y sus propiedades cuánticas calculadas. Estos modelos aprenden a predecir dichas propiedades directamente a partir de la topología 2D de la molécula (Gilmer *et al.*, 2017; Yang *et al.*, 2019), obviando la necesidad de resolver las ecuaciones de Schrödinger. Al igual que AlphaFold, su funcionamiento interno es opaco, pero su precisión predictiva es tan alta que se han convertido en herramientas indispensables para el cribado virtual (*virtual screening*) de millones de moléculas candidatas a fármacos, acelerando drásticamente el proceso de descubrimiento (Goh *et al.*, 2017; Stokes *et al.*, 2020).

En suma, este trabajo defiende una doble tesis: (1) existe una continuidad estructural y epistémica significativa entre las simulaciones computacionales clásicas y los sistemas actuales de *deep learning* —ambos constituyen tecnologías de producción de conocimiento sustentadas cada vez más en la validación empírico-pragmática, la opacidad inferencial y, en muchos casos, la centralidad de la visualización de resultados—; y (2) el reconocimiento filosófico de esta continuidad exige matizar y expandir los marcos conceptuales clásicos de modelo, representación y justificación epistémica, dada la irrupción de retos inéditos (como la radicalización de la opacidad, la falta de teoría mediadora y la centralidad de la interfaz humano-algoritmo) en sistemas como AlphaFold.

Destacar la relación entre ambas dimensiones —más que yuxtaponerlas— permite identificar con precisión qué aspectos se preservan, transforman o desafían en este nuevo paisaje epistémico. El análisis histórico y funcional de la continuidad solo cobra pleno sentido filosófico si revela los límites de los marcos clásicos; la propuesta de un cambio conceptual, a su vez, se fundamenta en el reconocimiento situado de las verdaderas novedades que trae el *deep learning* en el contexto de producción científica. Esta mirada doble evita tanto el rupturismo —que vería a AlphaFold como incomunicable con la tradición filosófica— como el simple continuismo —que reduciría la novedad a mera sofisticación técnica—, y justifica la urgente tarea de expandir las categorías filosóficas para comprender el estatuto epistémico del conocimiento científico, justamente en este nuevo paisaje epistémico.

Entre las principales objeciones cabe mencionar: (a) que la continuidad entre simulación y *deep learning* sería superficial dada la opacidad radical e inexplicabilidad de este último; (b) que los llamados a nuevos marcos conceptuales solo reeditan debates previos sobre justificación pragmática en la modelización. Nuestra estrategia es mostrar, a través del enfoque doble, tanto los elementos que justifican la comparación (validación empírica, creatividad modal, función exploratoria) como aquellos que exigen ampliaciones filosóficas genuinas (ausencia de teoría mediadora, nuevo *locus* para la representación y la justificación en la interfaz humano-algoritmo).

El trabajo se organiza en cuatro partes. En la primera se contextualiza el lugar de las simulaciones dentro de la filosofía de la ciencia atendiendo a las posiciones que las consideran una metodología epistémicamente novedosa y a aquellas que adoptan una postura continuista. En la segunda, se analiza la propuesta de Paul Humphreys

y su noción de “plantillas” y “modelos computacionales” como unidad de análisis evaluando en qué medida AlphaFold puede comprenderse en este marco. La tercera parte aborda la tensión entre las concepciones representacionistas clásicas y los modelos opacos basados en *deep learning*. Finalmente, se examinan los procedimientos de verificación y validación que consolidan el estatuto epistémico de AlphaFold, para concluir con una reflexión sobre las implicaciones que este tipo de artefactos tiene para la filosofía de la ciencia contemporánea.

1. Debate sobre el estatuto epistémico de las simulaciones computacionales

A pesar de su creciente ubicuidad e impacto transformador en la práctica científica, las simulaciones computacionales han suscitado una respuesta ambivalente dentro de la filosofía de la ciencia (Durán, 2017, p. 125; Lenhard, 2016; Morrison, 2012). Mientras en la práctica son cada vez más centrales en el descubrimiento, validación y justificación científica, su estatuto epistémico ha sido objeto de debate. ¿Son simplemente herramientas auxiliares que extienden nuestras capacidades de cálculo y visualización, o constituyen una forma cualitativamente nueva de conocimiento científico que exige repensar nuestras categorías filosóficas tradicionales? Este debate constituye el trasfondo necesario para evaluar hasta qué punto artefactos recientes como AlphaFold pueden ser pensados como parte de esa tradición o, más bien, como un caso que desborda sus categorías.

Autores como Eric Winsberg (2001 y 2010), Peter Galison (1996), Fritz Rohrlich (1990) y Paul Humphreys (2004 y 2009) han defendido la idea de que las simulaciones computacionales no pueden ser reducidas ni comprendidas adecuadamente desde los marcos clásicos de la filosofía de la ciencia. Para ellos, las simulaciones introducen una metodología con características propias, distintas de la teoría, el experimento y la instrumentación. Rohrlich (1990), por ejemplo, sostenía que el poder de cálculo asociado a las simulaciones supera las capacidades cognitivas humanas, y por ello constituyen una “metodología cualitativamente nueva” para las ciencias naturales (citado en Durán, 2017, p. 125). Winsberg (2001, p. 447) argumenta que las simulaciones poseen una “epistemología distintiva”, no reductible a las categorías tradicionales, como los modelos o las leyes (citado en Durán, 2017, p. 126).

Juan M. Durán (2017), en esta misma línea, sostiene que las simulaciones computacionales abren un campo filosófico específico que desafía el modo clásico de concebir la explicación y la justificación en ciencia. A su juicio, la epistemología que requieren las simulaciones no es solo una variación técnica o metodológica dentro del paradigma clásico, sino una ruptura conceptual que obliga a pensar en nuevas unidades de análisis.

Estas posiciones comparten una convicción fundamental, a saber, que la novedad de las simulaciones no es meramente empírica ni tecnológica, sino también filosófica. Tal novedad se expresa en distintos planos: ontológico, al introducir artefactos que no se limitan a representar entidades físicas, sino que operan mediante correlaciones estadísticas; epistémico, al producir conocimiento sin comprensión teórica explícita; metodológico (Winsberg, 2001 y 2010; Durán, 2017), al desplazar los procedimientos clásicos de validación y verificación. Por último, representacional, al cuestionar la centralidad de la transparencia conceptual como criterio de valor científico.

En contraste con estas posturas afirmativas, otros filósofos de la ciencia han mostrado escepticismo frente a la necesidad de revisar las categorías tradicionales. Roman Frigg y Julian Reiss (2009) reconocen que las simulaciones computacionales han introducido elementos novedosos en la práctica científica, pero se resisten a considerar que esta novedad implique una disrupción filosófica significativa. Según ellos, las cuestiones que plantean las simulaciones pueden ser tratadas, sin grandes ajustes conceptuales, desde los marcos ya existentes en la filosofía de los modelos científicos y de la experimentación.

A su juicio, muchas de las características distintivas de las simulaciones, como su componente visual, su carácter iterativo o su dependencia de algoritmos complejos, no constituyen problemas filosóficos genuinos, sino más bien cuestiones de tipo técnico, matemático o incluso psicológico. Así, por ejemplo, afirman que la preferencia por la visualización de resultados no es un tema epistemológico en sí mismo, sino una cuestión de percepción. “Parece ser más bien una cuestión de psicología antes que de epistemología el hecho de que preferimos lo visual a otras formas de representación” (Frigg y Reiss, 2009, p. 607).

Este tipo de argumentación refleja una postura que podríamos denominar continuista, en la medida en que no considera que las simulaciones computacionales exijan una reelaboración sustantiva del aparato conceptual de la filosofía de la ciencia. Aunque Frigg y Reiss

(2009) admiten que las simulaciones “introducen algo nuevo y excitante para la *ciencia*” (p. 595), concluyen que no hay razones suficientes para “reescribir la *filosofía de la ciencia*” (p. 595). Es importante señalar, no obstante, que esta postura fue formulada en un contexto previo a la irrupción masiva de tecnologías basadas en aprendizaje profundo —*deep learning*—. En ese momento, la mayoría de las simulaciones computacionales seguían estando estructuradas por modelos representacionales relativamente transparentes, apoyados en principios físicos o ecuaciones diferenciales explícitas. Así, el escepticismo de Frigg y Reiss puede entenderse como una defensa de la capacidad explicativa de los marcos tradicionales dentro de un paisaje científico todavía dominado por modelos interpretablemente anclados en teorías.

La novedad epistémica de las simulaciones que proponen Frigg y Reiss —al considerarlas meramente una forma de modelos matemáticos computacionalmente ejecutables— resulta problemática incluso en el contexto en el que escribieron. Ya entonces, prácticas como la climatología computacional, con sus mapas de temperatura, precipitación o circulación atmosférica, o la cosmología con simulaciones como la *Millennium Simulation* (Particle Physics & Astronomy Research Council, 2005), mostraban que las visualizaciones no eran simples ayudas perceptivas, sino instrumentos epistémicos centrales para la validación y la exploración teórica (Lenhard, 2016). Estos casos históricos bastaban para cuestionar la tesis de que lo visual es un asunto meramente psicológico. Ahora bien, la fuerte irrupción de EODS, como AlphaFold, intensifica aún más esta dificultad, pues en ellos la visualización constituye prácticamente la única vía de acceso epistémico a modelos internos opacos, lo cual refuerza el carácter indispensable de la representación visual para la práctica científica.

Diversos autores han destacado que la visualización cumple funciones que no pueden ser reemplazadas por descripciones formales o numéricas. Morrison (2015b), por ejemplo, ha argumentado que la forma en que los científicos interactúan con modelos complejos es inseparable de las maneras en que estos modelos se presentan visualmente. En este sentido, la visualización se convierte en una condición epistémica de posibilidad para el trabajo científico en entornos computacionalmente intensivos donde la complejidad se hace manifiesta (Lenhard, 2016). Aunque, podría objetarse, claro está, que la visualización, por más central que sea en la práctica científica, no altera el estatuto epistemológico de

fondo, sino que simplemente facilita el acceso humano a resultados ya generados algorítmicamente.

En efecto, los argumentos de Frigg y Reiss pueden ser refutados en sus propios términos, pues incluso antes del *deep learning* la visualización desempeñaba funciones epistémicas centrales que excedían lo meramente perceptivo. Sin embargo, el punto decisivo no está en esa disputa histórica, sino en el impacto de los desarrollos contemporáneos. AlphaFold representa un quiebre porque introduce una transformación significativa, en tanto que su validez epistémica no puede explicarse apelando a modelos representacionales clásicos ni a simulaciones dinámicas tradicionales ni a herramientas visuales subordinadas a teorías explícitas. Se trata de un tipo de artefacto científico —EODS— que obliga a reconsiderar la manera en que entendemos la justificación y el valor de las simulaciones en la ciencia actual.

Su potencia reside en la capacidad de generar predicciones altamente precisas sin apoyarse en una teoría explicativa explícita ni en una representación interpretable del proceso físico estudiado. En lugar de describir cómo ocurre el plegamiento proteico, infiere directamente estructuras tridimensionales a partir de correlaciones estadísticas inscritas en un espacio de alta dimensionalidad sin semántica biológica explícita. Desde esta perspectiva, pone en tensión las concepciones tradicionales de “simulación”, “modelo” y “representación” para abrir una zona epistémica distinta —cercana a la noción de “régimenes de objetividad” de Daston y Galison (2007)— en la que el conocimiento científico no descansa en la inteligibilidad causal, sino en la eficiencia inferencial y en la validación empírica de sistemas opacos.

No obstante, también puede sostenerse que AlphaFold prolonga, en un sentido formal-computacional, la tradición de la simulación. En efecto, su núcleo operativo se apoya en una estructura matemática precisa: la arquitectura *transformer* (Vaswani *et al.*, 2017; Jumper *et al.*, 2021), cuya lógica de atención y actualización de parámetros constituye un aparato formal-matemático que puede interpretarse como el “modelo” subyacente. En este sentido, AlphaFold no abandona del todo la lógica de las simulaciones; más bien, la radicaliza y la desplaza hacia un plano en el que lo decisivo ya no es la representación transparente de un proceso físico, sino la ejecución algorítmica de una estructura matemática que produce resultados empíricamente validados.

En lo que sigue, analizaremos cómo este doble registro —AlphaFold como ruptura epistémica y AlphaFold como simulación en sentido

formal— fue anticipado, al menos parcialmente, por propuestas como la de Paul Humphreys, quien buscó dotar a las simulaciones computacionales de una unidad de análisis propia, distinta de categorías como “teoría”, “modelo” o “ley”. Su aproximación servirá como punto de partida para mostrar por qué AlphaFold no debería entenderse simplemente como un ejemplo sofisticado de simulación, pero tampoco puede desligarse por completo de esta tradición; se sitúa, más bien, en la frontera misma donde la noción de simulación se redefine.

2. AlphaFold y las *Epistemically Opaque Deep-learning Simulations* (EODS)

Una de las propuestas filosóficas más influyentes para comprender el estatuto epistémico de estas prácticas es la de Paul Humphreys. A diferencia de los enfoques que buscan subsumirlas dentro del esquema tradicional de teorías científicas, ya sea en su versión sintáctica o semántica, Humphreys sostiene que requieren una nueva unidad de análisis filosófico capaz de integrar sus elementos distintivos: opacidad, manipulación algorítmica, temporalidad computacional y representación de resultados. Esta propuesta resulta especialmente relevante en la medida en que ofrece un marco intermedio: permite reconocer la especificidad epistémica de las simulaciones sin tratarlas como una ruptura completamente ajena a la tradición filosófica previa.

Humphreys argumenta que los conceptos clásicos como “teoría”, “modelo”, “ley” o “paradigma” no son adecuados para capturar la especificidad de las simulaciones computacionales modernas. Desde esta perspectiva, las unidades básicas de análisis que la filosofía de la ciencia ha utilizado históricamente resultan insuficientes para caracterizar estas prácticas. Por ello, Humphreys (2004, p. 60) introduce la noción de “plantilla” (*template*) como una herramienta conceptual destinada a reformular, más que simplemente reemplazar, esas unidades tradicionales. Una plantilla se define como un conjunto de restricciones muy generales y abstractas que luego pueden ser especificadas para generar una formulación concreta aplicable a un fenómeno. Como explica Durán siguiendo a Humphreys (2004, p. 60):

La segunda ley de Newton, tal y como la conocemos, se describe para cualquier fuerza, masa y aceleración. Sin embargo para hacer uso de ella aplicándola en un caso particular es necesario primero especificar una función

de fuerza particular, ya sea la fuerza gravitacional, electroestática, magnética o cualquier otra; algo similar ocurre con los restantes elementos de los que habla la ley: masa, aceleración, etc. [...]. A partir de esto es posible interpretar la segunda ley de Newton como una “plantilla teórica”, esto es, como restricciones muy generales que no son aplicables a ningún fenómeno particular sin la previa especificación de algunos de sus componentes. Variando de especificación, entonces, será posible dar cuenta de los distintos fenómenos que puede “contener” la plantilla (Durán, 2017, p. 131).

Del mismo modo, la teoría de la probabilidad funciona como una plantilla cuyos elementos específicos, como funciones de densidad o de distribución, deben completarse en contextos particulares. Estas plantillas, en tanto construcciones sintácticas abstractas, poseen dos funciones principales, según Humphreys, “una de ellas consiste en representar el mundo; la otra, en facilitar el cálculo” (Durán, 2017, p. 132).

El paso decisivo ocurre cuando una plantilla teórica se especifica de tal forma que produce un resultado computacionalmente tratable; en ese momento, se convierte en una plantilla computacional, esto es, en una estructura formal que puede ser ejecutada algorítmicamente. Esta transición —de lo abstracto a lo computacionalmente ejecutable— resulta central para comprender el estatuto de sistemas como AlphaFold. En efecto, su arquitectura *transformer* puede entenderse como una estructura matemática general que, al ser entrenada y especificada mediante grandes volúmenes de datos proteicos, se convierte en un aparato ejecutable capaz de producir resultados empíricamente contrastables. En este sentido, AlphaFold puede interpretarse como una simulación en un sentido formal-computacional, aun cuando sus vínculos con teorías biológicas explícitas sean débiles o incluso inexistentes. Humphreys (2004, p. 78) caracteriza estas plantillas computacionales mediante una quintupla de supuestos de construcción, citada por Durán (2017, p. 134):

©Ontología, idealizaciones, abstracciones, restricciones, aproximaciones©

En esta quintupla, el primer elemento, la ontología, especifica el tipo de objetos, propiedades y relaciones que son representados. Esta especificación sirve de base para los demás componentes. Los

otros cuatro conforman lo que Humphreys denomina el *conjunto de corrección*. Como señala Durán (2017, p. 134), “este conjunto [...] tiene como propósito principal indicar cómo las plantillas deben ser ajustadas cuando los resultados no coinciden con los datos empíricos”.

Dicho conjunto puede implicar desde la relajación de restricciones y el refinamiento de aproximaciones hasta modificaciones más profundas en la ontología de base. Un aspecto especialmente relevante para nuestra discusión es que estas plantillas computacionales no requieren derivarse de teorías científicas tradicionales, lo cual amplía significativamente el espectro de artefactos que pueden ser comprendidos dentro de este marco. Un ejemplo paradigmático de este punto es el proceso de Poisson, que puede funcionar como una plantilla de gran generalidad aplicable a fenómenos muy diversos —desde llamadas telefónicas hasta desintegraciones radioactivas— sin presuponer una teoría causal común que los unifique. Lo que legitima su uso no es la pertenencia a un mismo dominio teórico, sino la posibilidad de ser representados y computados mediante una misma estructura formal (Durán, 2017, p. 135).

Esta idea prepara el terreno para pensar modelos como AlphaFold, que no se derivan de ninguna teoría causal explícita, sino que operan sobre estructuras inferenciales entrenadas estadísticamente. Sin embargo, para comprender plenamente el estatuto epistémico de este tipo de artefactos, es necesario avanzar hacia una unidad de análisis más amplia. Si bien las plantillas computacionales capturan adecuadamente el paso de lo abstracto a lo computacionalmente ejecutable, Humphreys sostiene que no bastan para caracterizar por completo lo que ocurre en una simulación. Por ello, propone una unidad de análisis más robusta, a saber, el modelo computacional, definido como una séxtupla que integra no solo la plantilla y sus supuestos, sino también elementos vinculados con la interpretación, la justificación epistémica y la representación.

Así, el modelo computacional se presenta como una estructura formal compuesta por los siguientes elementos (Humphreys, 2004, pp. 102-103, citado en Durán, 2017, p. 137):

©Plantilla, supuestos de construcción, conjunto de corrección, interpretación, justificación inicial, representación©

La plantilla computacional provee el andamiaje formal básico —ya sea en forma de ecuaciones u otros aparatos matemáticos—, mientras que los supuestos de construcción y el conjunto de corrección se refieren,

respectivamente, a las decisiones necesarias para operacionalizar el modelo y a los ajustes que permiten mantenerlo en sintonía con los datos empíricos. A estos elementos se añaden tres componentes que amplían la dimensión propiamente epistémica del modelo computacional: la interpretación, que especifica cómo deben entenderse sus elementos en relación con el sistema estudiado; la justificación inicial, que explica por qué resulta adecuado emplear ese modelo en un contexto determinado; y la representación, que comprende tanto los resultados del proceso de simulación como las formas de su presentación.

La novedad de esta propuesta radica en que la unidad de análisis no es ninguna de estas partes por sí sola, sino su combinación funcional. En este sentido, los modelos computacionales no pueden ser comprendidos adecuadamente desde una perspectiva reduccionista o parcial. En particular, la representación de los resultados no constituye un aspecto meramente estético o auxiliar, sino una dimensión metodológica central para la interpretación y la validación del modelo. En este punto, la propuesta de Humphreys se aleja tanto de la concepción sintáctica de las teorías científicas —que las concibe como conjuntos de sentencias organizadas deductivamente— como de la concepción semántica —que las identifica con familias de modelos matemáticos abstractos—.

Sin embargo, podría objetarse que este desplazamiento hacia la combinación de elementos corre el riesgo de convertirse en una taxonomía acumulativa más que en una solución conceptual: ¿es suficiente enumerar componentes para dar cuenta de la novedad epistémica, o se requiere un marco más radicalmente distinto? Este interrogante resulta especialmente pertinente en el caso de AlphaFold, donde algunos de los elementos de la séxtupla —en particular la interpretación y la representación— adquieren un estatuto problemático, pues no se apoyan en teorías subyacentes claras, sino en correlaciones opacas aprendidas algorítmicamente. En este contexto, la inclusión de la temporalidad computacional adquiere un papel decisivo: para Humphreys, el modelo computacional no solo representa algo del mundo, sino que ejecuta un proceso computacional que produce resultados en el tiempo.

Humphreys denomina “núcleo de la simulación” a la dimensión temporal del proceso computacional, entendido como un proceso que produce soluciones a un modelo computacional en el tiempo. De este modo, la simulación no se reduce a una estructura formal o representacional, sino que implica una transformación activa de *inputs* en *outputs* mediante algoritmos y procedimientos numéricos. Aunque

toda computación transcurre en el tiempo, en las simulaciones científicas el tiempo de ejecución no constituye un mero subproducto técnico, sino el medio a través del cual se generan dinámicas que sustituyen, representan o anticipan procesos del mundo físico. Es precisamente en este punto donde AlphaFold tensiona el marco de Humphreys; su núcleo temporal no simula el plegamiento proteico paso a paso, sino que realiza inferencias estadísticas masivas que desembocan directamente en configuraciones estructurales finales.

A diferencia de otros enfoques, Humphreys sostiene que las peculiaridades metodológicas de las simulaciones emergen con especial claridad en los procesos de representación de sus resultados. El modo en que los *outputs* de una simulación son presentados —ya sea en forma de tablas numéricas, gráficos estáticos o visualizaciones dinámicas— no es metodológicamente neutro, pues incide directamente en la interpretación del modelo computacional y en las inferencias que pueden extraerse de él. En este sentido, la representación cumple una función epistemológica central en la medida en que media entre los resultados computacionales y la comprensión humana, aun cuando no siempre facilite un entendimiento causal del fenómeno estudiado (Durán, 2017, pp. 137-138).

Esto resulta especialmente relevante en el caso de las EODS como AlphaFold, donde la comprensión humana del proceso inferencial interno es limitada, pero la presentación visual de los resultados permite la producción de inferencias científicas robustas. En estos casos, la representación visual no se limita a comunicar un resultado previamente constituido, sino que participa activamente en su validación epistémica. De este modo, AlphaFold produce una forma de conocimiento que no se apoya en una comprensión teórica o causal del fenómeno ni en una inteligibilidad transparente de su funcionamiento interno. Como señalan Gao y Ganguli (2015), incluso cuando se conocen la arquitectura de la red, sus reglas de actualización y sus parámetros, “aún nos elude una comprensión significativa de cómo funcionan estas redes” (p. 151).

Ahora bien, podría objetarse que la dependencia de AlphaFold respecto de visualizaciones tridimensionales para la validación de sus resultados no lo diferencia de manera categórica de otras prácticas científicas en las que las representaciones gráficas también funcionan como mediaciones indispensables —por ejemplo, en la astrofísica computacional o en la biología de sistemas—. La diferencia, sin embargo, no es meramente instrumental: en AlphaFold la visualización no cumple

un rol auxiliar respecto del modelo, sino que se convierte en la principal vía de acceso epistémico a un sistema cuyo funcionamiento inferencial interno permanece en gran medida opaco.

Esta opacidad epistémica ha sido señalada por diversos autores como una característica distintiva del aprendizaje profundo. Omri Barak (2017, p. 5), por ejemplo, advierte que “el aprendizaje automático nos proporciona niveles de rendimiento cada vez mayores, acompañados de un aumento paralelo de la opacidad”. En estos casos, lo que Humphreys denomina el *predicamento antropocéntrico* se vuelve especialmente relevante. “Las capacidades computacionales superan las humanas, forzando una reevaluación de cómo entendemos y evaluamos estos métodos científicos” (Humphreys, 2009, p. 617). Por supuesto, la opacidad epistémica de las EODS no es absoluta. En principio, es posible auditar los algoritmos, examinar su arquitectura y analizar sus parámetros e hiperparámetros (Rubio *et al.*, 2024). Sin embargo, este acceso técnico no elimina la opacidad relevante desde el punto de vista epistémico. La complejidad combinatoria de estos sistemas y la naturaleza distribuida de sus representaciones internas hacen que cualquier reconstrucción causal del proceso inferencial sea, en el mejor de los casos, parcial e incompleta. En consecuencia, aunque pueda conocerse *cómo* está construido el sistema, sigue siendo difícil —cuando no imposible— explicar de manera inteligible *por qué* llega a un resultado específico en un caso particular.

Esta situación conduce a una redefinición de los criterios de validez epistémica en la ciencia actual. En el contexto de las EODS, ya no ocupan un lugar central la transparencia, la comprensión causal ni la fidelidad representacional, sino la precisión empírica, la eficiencia inferencial y la utilidad predictiva. El valor epistémico de estos sistemas no se fundamenta en su capacidad para hacer inteligibles los mecanismos subyacentes, sino en la estabilidad, robustez y reproducibilidad de sus resultados frente a evidencia independiente. En este sentido, la justificación epistémica se desplaza desde criterios explicativos tradicionales hacia formas de validación orientadas al desempeño.

Esta redefinición de los criterios de validez epistémica encuentra un marco conceptual adecuado en las propuestas recientes sobre modelado modal. En este enfoque, los modelos no se evalúan principalmente por su capacidad de representar fielmente procesos causales reales, sino por su aptitud para generar información confiable acerca de lo que es posible, probable o esperable bajo determinadas condiciones contrafácticas.

Así, la justificación epistémica se apoya en la consistencia interna del modelo, en su desempeño predictivo y en la plausibilidad física de los escenarios que produce, aun cuando estos no se correspondan con una representación literal del sistema estudiado (Sjölin Wirling y Grüne-Yanoff, 2021).

Desde esta perspectiva, AlphaFold puede entenderse como un sistema que genera información modal en forma de predicciones estructurales altamente confiables más que como un modelo que representa dinámicamente el proceso físico del plegamiento proteico. Su aporte epistémico no consiste en describir cómo ocurre dicho proceso, sino en delimitar un espacio de posibilidades estructurales empíricamente plausibles, cuya validez se evalúa mediante su concordancia con datos experimentales independientes. En este sentido, la justificación epistémica de AlphaFold no depende de una explicación causal subyacente, sino de la estabilidad, precisión y utilidad predictiva de los escenarios que produce, en línea con algunas de las estrategias justificatorias propuestas para el modelado modal (Sjölin Wirling y Grüne-Yanoff, 2021).

En este marco, la representación adquiere un estatuto epistémico particular. En AlphaFold, la representación no opera principalmente como una correspondencia interna entre un modelo y un sistema físico, sino como un punto de articulación entre un proceso inferencial opaco y las capacidades cognitivas humanas. Las estructuras tridimensionales generadas por el sistema no hacen visible el mecanismo interno de inferencia, sino que traducen resultados algorítmicos en formatos que permiten su contraste empírico, su evaluación comparativa y su integración en prácticas científicas preexistentes. De este modo, la representación no se sitúa dentro del modelo, sino en la interfaz donde los *outputs* computacionales se vuelven epistémicamente utilizables.

Considerado desde esta perspectiva, AlphaFold no puede ser entendido ni como una mera extensión técnica de las simulaciones computacionales tradicionales ni como una ruptura completamente desvinculada de ellas. Por un lado, comparte con las simulaciones clásicas la ejecución algorítmica de una estructura matemática formal, la dependencia de criterios de validación empírica y el papel metodológicamente central de la representación de resultados. Por otro lado, radicaliza rasgos ya presentes —como la opacidad inferencial y la distancia respecto de teorías explicativas explícitas— hasta el punto de tensionar los marcos conceptuales heredados. En este sentido,

AlphaFold ocupa una posición fronteriza: prolonga la tradición de la simulación computacional al tiempo que revela sus límites conceptuales, preparando el terreno para una reconsideración más amplia del estatuto epistémico de los sistemas basados en aprendizaje profundo.

3. Humphreys y el marco de plantillas y modelos computacionales

En la tradición filosófica sobre modelos científicos, el valor epistémico de un modelo se ha medido, en gran parte, por su capacidad de representar fielmente algún aspecto estructural o causal del sistema estudiado. Bajo este ideal representacionista, un modelo científicamente valioso actúa como mediador que preserva —al menos parcialmente— la estructura relevante del fenómeno, permitiendo así comprenderlo a través de su forma. Como afirma Morrison (2015b, p. 390), esta perspectiva entiende que el mérito de un modelo radica en “salvar” la estructura del fenómeno objetivo, lo cual hace de la correspondencia estructural un criterio central de legitimidad epistémica.

Este ideal de correspondencia está presente también en la caracterización que Barberousse *et al.* (2009) hacen de la simulación como “una correspondencia obvia entre el proceso que opera en el dispositivo simulador y los que operan en el sistema modelado” (p. 559). En la misma línea, Hughes (1999) muestra cómo en ciertos casos —como un modelo hidráulico de una rueda real, un modelo eléctrico de un péndulo amortiguado, una simulación basada en agentes, o un autómata celular concebido como sistema físico— la relación entre procesos del modelo y procesos del sistema representado es clara y trazable.

Hughes contrasta estos ejemplos con una simulación computacional de las interacciones atómicas entre un “pine” y una placa, descrita en Rohrlich (1991). En esta última, aunque la correspondencia directa entre procesos físicos y computacionales no es evidente, se mantiene el ideal representacional. Como advierten Barberousse *et al.* (2009, p. 559), Hughes señala que: “al igual que los sistemas examinados [los tres primeros sistemas simuladores], la computadora es un sistema material, y al utilizarla para responder nuestras preguntas estamos confiando en los procesos físicos que ocurren dentro de ella”. La legitimidad epistémica de la simulación descansa así en la idea de que los procesos computacionales, por su fisicalidad, pueden fungir como sustitutos representacionales de los procesos del sistema estudiado.

En este marco, el uso de la simulación computacional se justifica en la medida en que se postula la existencia de una relación estructural

o formal que vincula tres niveles: el modelo base (teórico), el modelo implementado (o “modelo agrupado”) y la computadora programada. Norton y Suppe (2001) denominan a esta condición “realización” y se exige que los tres niveles “compartan una estructura común” (Barberousse *et al.*, 2009, p. 563). Este requisito opera como un ideal normativo de transparencia estructural, según el cual el modelo no solo debe producir resultados correctos, sino hacerlo mediante procesos que puedan ser entendidos como reflejos, transformaciones o idealizaciones explícitas de los procesos del sistema físico.

Este enfoque está íntimamente ligado a un ideal de inteligibilidad. Un modelo es científicamente valioso en la medida en que permite a los científicos entender el sistema estudiado a través de su estructura representacional. Incluso cuando el cálculo se vuelve demasiado complejo para resolverse de manera analítica, subsiste como ideal regulativo la aspiración a una solución comprensible, capaz de mostrar de manera al menos parcial cómo las propiedades del sistema se reflejan en las propiedades del modelo.

En la práctica, las simulaciones computacionales a menudo requieren aproximaciones no analíticas: el modelo matemático de partida no es directamente implementable en un computador digital y, por tanto, debe sufrir transformaciones, discretizaciones y ajustes que modifican su estructura formal. Esto crea lo que podríamos llamar una brecha de implementación estructural entre el modelo teórico ideal y el artefacto computacional efectivo, una brecha que no constituye una anomalía ocasional, sino una característica constitutiva del modelado computacional.

Esta brecha pone en tensión el ideal representacionalista clásico, pues implica que la validez epistémica de una simulación no depende únicamente de su fidelidad estructural a un modelo base, sino también de la robustez de los procedimientos numéricos, de la calidad de los datos de entrada y de las estrategias de verificación y validación (V&V) empleadas. En consecuencia, la noción misma de “solución científica” se reconfigura (Lenhard, 2016): deja de identificarse exclusivamente con una derivación analítica exacta y pasa a entenderse como una solución operativamente adecuada, cuya legitimidad se mide por su capacidad de producir resultados eficaces para la investigación, incluso cuando los procesos internos permanecen opacos o irreductiblemente complejos.

Esta reconfiguración es una consecuencia directa de la naturaleza de los problemas que abordan las EODS. La opacidad no es un velo

accidental, sino una característica constitutiva del método. Estos modelos operan en espacios de parámetros de altísima dimensionalidad (millones o miles de millones de dimensiones), donde la tarea consiste en aprender una función no lineal que mapea un *input* complejo (e. g., una secuencia de aminoácidos) a un *output* igualmente complejo (e. g., coordenadas atómicas en 3D). La solución que encuentra el algoritmo de optimización —como el descenso de gradiente estocástico (SGD) o el backpropagation (BP)— puede caracterizarse como una región funcionalmente estable en este espacio de alta dimensionalidad. Dicha solución es en gran medida irreductible a una secuencia de pasos lógicos o a una ecuación directamente comprensible para la cognición humana. En este sentido, el modelo no “deduce” la estructura en un sentido clásico, sino que la reconstruye inferencialmente mediante una representación distribuida y correlacional aprendida de los datos. Así, la explicación causal del plegamiento no es eliminada, pero queda desplazada por una inferencia paralela que se implementa en la arquitectura computacional del sistema.

Ahora bien, aquí emerge el dilema central de nuestro trabajo. Para algunos autores, este desplazamiento no exige abandonar la noción de simulación, sino reformularla o ampliarla; lo decisivo ya no sería la correspondencia representacional estricta, sino la resolución algorítmica de estructuras matemáticas (una lectura continuista en sentido amplio). Sin embargo, puede sostenerse también —y AlphaFold lo ilustra con particular fuerza— que estamos ante una transformación más profunda: un nuevo tipo de artefacto científico —las EODS— cuya opacidad y fertilidad epistémica tensionan seriamente las categorías heredadas (una lectura disruptiva, aunque no necesariamente rupturista). Frente a este dilema, la epistemología de los modelos modales cobra un papel relevante, en la medida en que ofrece estrategias conceptuales para justificar los *outputs* de modelos que no son tradicionalmente representacionales sin por ello desligarlos por completo de la práctica científica establecida (Sjölin Wirling y Grüne-Yanoff, 2021).

Este desplazamiento se acentúa en las simulaciones contemporáneas basadas en redes neuronales profundas. Aquí, la carga heurística y el trabajo de ajuste se redistribuyen; ya no recaen exclusivamente en el investigador humano, sino que son parcialmente delegados a la infraestructura computacional, entendida como arquitecturas con millones de parámetros optimizadas mediante BP o SGD. En estos casos, incluso la noción de “correspondencia estructural” entre

modelo y sistema físico se vuelve problemática, pues el proceso interno del modelo ya no refleja, ni siquiera de manera idealizada, el comportamiento causal del fenómeno que estudia, sino regularidades estadísticas funcionalmente eficaces.

Ahora bien, el hecho de que AlphaFold opere sobre arquitecturas formales y ejecute algoritmos matemáticos no es, por sí solo, suficiente para considerarlo una simulación computacional en sentido filosófico. Un criterio basado únicamente en la ejecución algorítmica o en la existencia de una estructura matemática subyacente sería excesivamente laxo, pues llevaría a clasificar como “simulaciones” a artefactos puramente computacionales, como una calculadora o un algoritmo de ordenamiento, trivializando así la noción de “simulación científica”. Lo que distingue a AlphaFold no es simplemente que compute, sino el rol epistémico que desempeña dentro de una práctica científica específica: explora un espacio de configuraciones físicamente posibles, produce información modal relevante sobre sistemas del mundo, y sus *outputs* son sometidos a procedimientos sistemáticos de verificación y validación empírica independientes. En este sentido, AlphaFold puede entenderse como una simulación no por su ontología computacional, sino por su función epistémica, en tanto sustituye, complementa y orienta prácticas experimentales en biología estructural.

Desde esta perspectiva, AlphaFold prolonga la noción de “simulación”, pero solo bajo una concepción ampliada y no trivial de esta; esto es, no como mera ejecución de una estructura formal, sino como un artefacto computacional integrado en un régimen de producción, validación y uso del conocimiento científico, donde la estructura matemática subyacente es condición necesaria, pero no suficiente, para su estatuto epistémico.

4. Opacidad, representación y visualización en las simulaciones contemporáneas

En la práctica científica, las V&V constituyen procedimientos clave para evaluar la fiabilidad de las simulaciones. La verificación busca asegurar que el modelo teórico ha sido implementado correctamente en el código, mientras que la validación consiste en contrastar los resultados de la simulación con datos experimentales u observacionales (Morrison, 2015b). Sin embargo, en sistemas complejos, esta validación se ve limitada por la escasez o inaccesibilidad de datos

empíricos, precisamente la razón que motiva el recurso a simulaciones. Como señalan Morrison (2015a y 2015b) y Winsberg y Alvarado (2022), esta tensión ha impulsado el desarrollo de métricas y protocolos de validación que no se reducen a la comparación directa con datos, sino que se configuran en estrecha dependencia de la propia simulación, poniendo en cuestión la idea de un criterio de validación completamente externo e independiente del artefacto computacional.

Este escenario difumina los límites entre simulación y experimentación convencional. La simulación ya no se presenta solo como una herramienta auxiliar de la experimentación, sino como una parte constitutiva del proceso experimental en sí mismo. Así, el valor epistémico de una simulación no depende exclusivamente de su ajuste a un modelo teórico previo, sino también de su capacidad para integrarse en un ciclo iterativo de predicción, prueba y ajuste que involucra tanto al modelo computacional como a los dispositivos experimentales, desplazando el foco desde la representación fiel de un fenómeno hacia la eficacia operativa del conjunto modelo-simulación-experimento.

Un ejemplo representativo de esta integración lo ofrece el Gran Colisionador de Hadrones (LHC). En este caso, la simulación no es un simple complemento, sino la fundación misma de los experimentos (Morrison, 2015a, p. 288). Las simulaciones son indispensables para interpretar datos, diseñar detectores y seleccionar eventos relevantes, aunque estas simulaciones permanecen ancladas en modelos teóricos explícitos de la física de partículas. De hecho, en palabras de Morrison, “nada se sabe sin simulación” (2015a, p. 288).

Esto implica que los datos denominados “reales” y los generados por simulaciones se procesan de forma conjunta, lo que hace “prácticamente sin sentido cualquier distinción tajante entre simulación y ‘experimento’” en ciertos contextos altamente mediados computacionalmente (Morrison, 2015a, p. 289). La validación, en este marco, se convierte en un proceso continuo e integrado, en el que las simulaciones no solo contrastan hipótesis, sino que participan activamente en la definición de los experimentos y en la interpretación de sus resultados. Esta simbiosis transforma la simulación en una nueva forma de experimentación sin eliminar por completo las diferencias conceptuales entre ambos dominios.

Se consolida así la práctica de la experimentación *in silico*, entendida no como un sustituto de la experimentación empírica tradicional, sino como un dispositivo heurístico que estructura y restringe el espacio

de hipótesis científicas. En el caso de AlphaFold, un biólogo puede obtener una estructura tridimensional plausible para una proteína de función desconocida. Esta estructura no constituye una explicación del proceso de plegamiento, pero sí una hipótesis geométrica altamente informada, capaz de revelar sitios de unión potenciales o similitudes estructurales con otras proteínas. A partir de esta conjetura generada computacionalmente, es posible diseñar experimentos de laboratorio (*wet lab*) más enfocados y eficientes —por ejemplo, mediante mutagénesis dirigida— orientados a evaluar la función inferida.

De este modo, las EODS no solo validan o predicen resultados, sino que reconfiguran ciertos aspectos de la lógica del descubrimiento científico. En lugar de una exploración empírica relativamente abierta, se generan conjeturas altamente restringidas mediante procedimientos computacionales, que luego se someten a corroboración o falsación experimental. El modelo funciona, así, como un dispositivo generativo de conjeturas más que como una representación del fenómeno subyacente. Un caso paradigmático de esta dinámica —aunque no basado en aprendizaje profundo— puede encontrarse en la búsqueda del bosón de Higgs, donde las simulaciones desempeñaron un papel decisivo en la interpretación de los datos y en el diseño experimental. En estos contextos, los llamados “datos experimentales” y los “datos de simulación” se entrelazan de tal modo que resulta difícil sostener una jerarquía epistémica tajante entre ambos.

El caso de AlphaFold encarna con particular claridad este nuevo régimen de validación en las simulaciones actuales. Dada la opacidad interna de su funcionamiento, la posibilidad de una verificación en sentido clásico —esto es, el seguimiento y control inferencial de los procesos computacionales que conducen a los resultados— se ve drásticamente limitada. En consecuencia, su integración en la práctica científica se apoya fundamentalmente en la corroboración empírica sistemática de sus *outputs* más que en la comprensión de los procesos que los generan. Su legitimidad epistémica no proviene de la inteligibilidad de su arquitectura interna ni de una correspondencia representacional con el proceso de plegamiento proteico, sino de la concordancia estable y reproducible entre sus predicciones y las estructuras proteicas determinadas experimentalmente mediante cristalografía de rayos X o criomicroscopía electrónica.

En contextos donde la validación directa resulta difícil o imposible, AlphaFold es evaluado mediante métricas especializadas y a través de

competencias comunitarias como CASP (*Critical Assessment of Protein Structure Prediction*). En estos escenarios, las predicciones se comparan con resultados experimentales aún no publicados, lo que permite evaluar su desempeño predictivo de manera controlada e independiente. Los éxitos reiterados de AlphaFold en estas evaluaciones han mostrado niveles de exactitud comparables a los métodos experimentales (Jumper *et al.*, 2021), lo que consolida su fiabilidad como herramienta predictiva, sin que ello implique una validación del proceso inferencial interno ni una caracterización representacional de su funcionamiento.

La discusión en torno a la validación nos devuelve al punto de partida: la necesidad de contar con categorías filosóficas capaces de dar cuenta de prácticas científicas que ya no se ajustan plenamente a los marcos tradicionales. En este sentido, la propuesta de Humphreys resulta especialmente fecunda como punto de transición. Al concebir las simulaciones como unidades epistémicas específicas, permite desplazar la atención desde la reducción a teorías o modelos clásicos hacia la articulación funcional de plantillas, supuestos, procesos computacionales y formas de representación.

Sin embargo, como el caso de AlphaFold pone de manifiesto, incluso este marco requiere ser revisado en un punto decisivo: la representacionalidad ya no puede seguir operando como una propiedad definitoria de las simulaciones computacionales. En EODS, la validez epistémica se apoya menos en la correspondencia representacional con el fenómeno *target* que en regímenes de validación empírica y en la capacidad de integrar outputs confiables en la práctica científica, lo que obliga a repensar el estatuto mismo de la representación dentro de la simulación.

De este modo, AlphaFold puede entenderse, en un primer registro, como la manifestación paradigmática de una nueva clase de artefactos científicos (EODS): opacos, no representacionales y, sin embargo, epistémicamente fértiles. Alternativamente —y de manera más controvertida— podría sostenerse que estos sistemas radicalizan la noción misma de “simulación”, siempre que esta se redefina en un sentido estrictamente formal-computacional, desvinculado del ideal clásico de correspondencia estructural con un fenómeno *target*. En este segundo registro, lo decisivo ya no sería la representación del proceso físico, sino la ejecución algorítmica de estructuras matemáticas capaces de producir *outputs* empíricamente validados. No obstante, esta ampliación conceptual no es inocua y exige precisiones filosóficas sustantivas, precisamente para

evitar que toda computación exitosa sea trivialmente clasificada como simulación. Ambas lecturas convergen, así, en un mismo diagnóstico: la filosofía de la ciencia contemporánea debe ampliar —pero también afinar— su marco conceptual si pretende comprender cómo se produce y se justifica el conocimiento en la era de los sistemas opacos basados en inteligencia artificial.

Comentarios de cierre

El recorrido realizado en este trabajo ha mostrado que la discusión filosófica sobre las simulaciones computacionales no puede permanecer inalterada ante la irrupción de artefactos como AlphaFold. Las posiciones afirmativas (Winsberg, Galison, Rohrlich, Humphreys, Durán) han defendido, con distintos matices, que las simulaciones constituyen una metodología epistémicamente significativa y, en ciertos sentidos, novedosa (Rohrlich, 1990; Winsberg, 2001). Por su parte, las posturas más escépticas (Frigg y Reiss) han sostenido que estas prácticas pueden comprenderse adecuadamente dentro de los marcos tradicionales de la filosofía de los modelos y la experimentación científica.

El análisis desarrollado sugiere, sin embargo, que la emergencia de sistemas estadísticos de gran escala como las *Epistemically Opaque Deep-learning Simulations* (EODS) introduce tensiones que no pueden resolverse sin ajustes conceptuales. AlphaFold no reconstruye el proceso físico del plegamiento proteico en sentido dinámico o causal; en su lugar, aprende regularidades complejas en espacios de alta dimensionalidad y produce resultados cuya fiabilidad depende de su validación empírica y de su integración efectiva en la práctica científica. Esto plantea un desafío directo a las concepciones representacionistas clásicas, sin que ello implique, automáticamente, la necesidad de abandonar la noción de “simulación”.

En efecto, si bien podría sostenerse que AlphaFold prolonga ciertos rasgos de la simulación en un sentido formal-computacional —en la medida en que ejecuta estructuras matemáticas precisas mediante procedimientos algorítmicos—, este rasgo, por sí solo, no es suficiente para justificar su clasificación como simulación computacional. De lo contrario, cualquier artefacto computacional exitoso podría ser trivialmente subsumido bajo dicha categoría. El problema filosófico relevante no consiste, entonces, en identificar la presencia de cálculo o formalismo, sino en determinar qué condiciones epistémicas y metodológicas permiten distinguir entre herramientas de cómputo y

artefactos capaces de generar conocimiento científicamente significativo sobre sistemas del mundo.

En este contexto, la noción de “representación” que recupera Humphreys adquiere un papel central. La representación ya no se entiende como una correspondencia interna entre modelo y fenómeno, sino como una interfaz funcional que conecta procesos computacionales opacos con las capacidades cognitivas y prácticas de los científicos. Las visualizaciones tridimensionales producidas por AlphaFold no son meros recursos ilustrativos, sino componentes constitutivos de su uso epistémico: es a través de ellas que los resultados se interpretan, se validan y se integran en programas de investigación más amplios.

Los procedimientos de verificación y validación refuerzan este desplazamiento. Al igual que en otros contextos de experimentación altamente mediada por simulaciones —como el Large Hadron Collider, donde “nada se sabe sin simulación” (Morrison, 2015a, p. 288)—, la fiabilidad de AlphaFold se sostiene en ciclos iterativos de contrastación empírica, evaluación comunitaria y métricas especializadas. La justificación epistémica no descansa aquí en la transparencia de los mecanismos internos, sino en la robustez de los resultados y en su capacidad para integrarse productivamente en prácticas experimentales posteriores.

Este escenario invita a repensar la categoría de modelo más allá de los moldes representacionalistas tradicionales. AlphaFold ilustra un régimen epistémico en el que es posible generar conocimiento científicamente valioso sin una comprensión causal inmediata del fenómeno estudiado, pero de manera sistemática, evaluable y socialmente controlada. Comprender este régimen exige reconocer que la inteligibilidad no siempre es condición necesaria de la justificación y que la representación puede concebirse —siguiendo a Humphreys— como una mediación funcional entre cálculo automatizado y uso humano. En este sentido, el análisis converge con los objetivos del modelado modal, orientado a justificar cómo los modelos pueden proveer información fiable sobre lo posible o lo contrafactual incluso en ausencia de teorías causales completas (Sjölin Wirling y Grüne-Yanoff, 2021).

En definitiva, AlphaFold no debe entenderse simplemente como un logro computacional excepcional ni como una ruptura absoluta con las prácticas científicas previas. Más bien constituye un caso límite que obliga a refinar las categorías con las que la filosofía de la ciencia analiza los modelos, las simulaciones y la justificación epistémica.

Enfrentar este desafío conceptual es una tarea central para comprender cómo se produce y se valida el conocimiento científico en un contexto crecientemente mediado por sistemas opacos de inteligencia artificial

Referencias

- Barak, O. (2017). Recurrent Neural Networks as Versatile Tools of Neuroscience Research. *Curr Opin Neurobiol*, 46, 1-6. <https://doi.org/10.1016/j.conb.2017.06.003>
- Barberousse, A., Franceschelli, S. y Imbert, C. (2009). Computer Simulations as Experiments. *Synthese*, 169(3), 557-574. <https://doi.org/10.1007/s11229-008-9430-7>
- Chirimuuta, M. (2021). Prediction Versus Understanding in Computationally Enhanced Neuroscience. *Synthesis*, 199, 767-790. <https://doi.org/10.1007/s11229-020-02713-0>
- Daston, L. y Galison, P. (2007). *Objectivity*. Zone Books.
- Durán, J. M. (2017). Simulaciones computacionales. *Principia*, 21(1), 125-140. <https://doi.org/10.5007/1808-1711.2017v21n1p125>
- Frigg, R. y Reiss, J. (2009). The Philosophy of Simulation: Hot New Issues or Same Old Stew? *Synthese*, 169, 593-613. <https://doi.org/10.1007/s11229-008-9438-z>
- Galison, P. (1996). Computer Simulations and the Trading Zone. En P. Galison y D. J. Stump (eds.), *The Disunity of Science: Boundaries, Context and Power* (pp. 118-157). Stanford University Press.
- Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O. y Dahl, G. E. (2017). Neural Message Passing for Quantum Chemistry. *Proceedings of the 34th International Conference on Machine Learning*, 70, 1263-1272. <https://proceedings.mlr.press/v70/gilmer17a.html>
- Gao, P. y Ganguli, S. (2015). On Simplicity and Complexity in the Brave New World of Large-Scale Neuroscience. *Curr Opin Neurobiol*, 32, 148-155. <https://doi.org/10.1016/j.conb.2015.04.003>
- Goh, G. B., Hodas, N. O., & Vishnu, A. (2017). Deep Learning for Computational Chemistry. *Journal of Computational Chemistry*, 38(16), 1291-1307. <https://doi.org/10.1002/jcc.24764>
- Hughes, R.I.G. (1999). The Ising Model, Computer Simulation, and Universal Physics. En M. Morgan y M. Morrison (eds.), *Models as Mediators: Perspectives on Natural and Social Science* (pp. 97-145). Cambridge University Press. <https://doi.org/10.1017/CBO9780511660108.006>

- Humphreys, P. (2004). *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford University Press. <https://doi.org/10.1093/0195158709.001.0001>
- Humphreys, P. (2009). The Philosophical Novelty of Computer Simulation Methods. *Synthese*, 169, 615-626. <https://doi.org/10.1007/s11229-008-9435-2>
- Jensen, F. (2017). *Introduction to Computational Chemistry*. Wiley. <https://doi.org/10.14293/S2199-1006.1.SOR-CHEM.CLYWQK5.v1>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... y Hassabis, D. (2021). Highly Accurate Protein Structure Prediction with AlphaFold. *Nature*, 596, 583-589. <https://doi.org/10.1038/s41586-021-03819-2>
- Lenhard, J. (2016). Computer Simulation. En P. Humphreys (ed.), *The Oxford Handbook of Philosophy of Science* (pp. 710-736). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199368815.013.42>
- Morrison, M., (2012). Models, Measurement and Computer Simulation: The Changing Face of Experimentation. *Philosophical Studies*, 143, 33-57. <https://doi.org/10.1007/s11098-008-9317-y>
- Morrison, M. (2015a). Legitimizing Simulation: Methodological Issues of Verification and Validation. En M. Morrison, *Reconstructing Reality: Models, Mathematics, and Simulations* (pp. 248- 286). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199380275.003.0008>
- Morrison, M. (2015b). Without It There's Nothing: The Necessity of Simulation in the Higgs Search. En M. Morrison, *Reconstructing Reality: Models, Mathematics, and Simulations* (pp. 287-316). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199380275.003.0009>
- Norton, S. y Suppe, F. (2001). Why Atmospheric Modeling Is Good Science. En P. Edwards y C. Miller (eds.), *Changing the Atmosphere* (pp. 67-106). The MIT Press. <https://doi.org/10.7551/mitpress/1789.003.0006>
- Particle Physics & Astronomy Research Council (2005, 4 de junio). Millennium Simulation—The Largest Ever Model of the Universe. *ScienceDaily*. www.sciencedaily.com/releases/2005/06/050604061156.htm
- Rohrlich, F. (1990). Computer Simulation in the Physical Sciences. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2, 507-518. <https://doi.org/10.1086/psaprocbienmeetp.1990.2.193094>
- Rubio, E., Giri, L. y Ilcic, A. (2023). Desafíos epistemológicos en la era de las redes neuronales artificiales: abordando sistemas complejos

- desde una perspectiva computacional. *Argumentos de Razón Técnica*, 26, 145-178. <https://doi.org/10.12795/Argumentos/2023.i26.06>
- Sjölin Wirling, Y. y Grüne-Yanoff, T. (2021). The Epistemology of Modal Modeling *Philosophy Compass*, 16(10), e12775. <https://doi.org/10.1111/phc3.12775>
- Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., ... y Collins, J. J. (2020). A Deep Learning Approach to Antibiotic Discovery. *Cell*, 180(4), 688-702.e13. <https://doi.org/10.1016/j.cell.2020.01.021>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. y Polosukhin, I. (2017). Attention Is All You Need. En U. von Luxburg, I. Guyon, S. Bengio, H. Wallach, R. Fergus (eds.), *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 6000-6010). Curran Associates Inc. <https://doi.org/10.48550/arXiv.1706.03762>
- Vicsek, T. (2002). The Bigger Picture. *Nature*, 418(6894), 131. <https://doi.org/10.1038/418131a>
- Winsberg, E. (2001). Simulations, Models, and Theories: Complex Physical Systems and their Representations. *Philosophy of Science*, 68, S442-S454. <https://doi.org/10.1086/392927>
- Winsberg, E. (2010). *Science in the Age of Computer Simulation*. The University of Chicago Press. <https://doi.org/10.7208/chicago/9780226902050.001.0001>
- Winsberg, E. y Alvarado, R. (2022). Computer Simulations in Science. *The Stanford Encyclopedia of Philosophy*. E. N. Zalta y U. Nodelman (eds.). <https://plato.stanford.edu/archives/spr2026/entries/simulations-science/>
- Yang, K., Swanson, K., Jin, W., Coley, C., Eiden, P., Gao, H. Guzman-Perez, A., Hopper, T., Kelley, B., Mathea, M., Palmer, A., Settels, V., Jaakkola, T., Jensen, K. y Barzilay, R. (2019). Analyzing Learned Molecular Representations for Property Prediction. *Journal of Chemical Information and Modeling*, 59(8), 3370-3388. <https://doi.org/10.1021/acs.jcim.9b00237>

