Andamios

Revista de Investigación Social
Volumen 23, Número Especial II, 2026

UACM
Universidad Autónoma
de la Ciudad de México
NADA HUMANO ME ES AJENO

Colegio de Humanidades y Ciencias Sociales

Andamios
Revista de Investigación Social

Andamios, Revista de Investigación Social, Volumen 23, Número Especial II, 2026 es una publicación editada por la Universidad Autónoma de la Ciudad de México a través del Colegio de Humanidades y Ciencias Sociales con dirección en Dr. Garcia Diego, núm. 168, col. Doctores, Del. Cuauhtémoc, C.P. 06720, México, Ciudad de México. Tel. 551107 0280, www.uacm.edu.mx, disponible en www.uacm.edu.mx/andamios. Editor responsable: Oscar Rosas Castro. Número de certificado de reserva del título: 04-2004-091014130100-102, ISSN de la versión impresa 1870-0063 e ISSN de la versión electrónica 2594-1917, otorgados por el Instituto Nacional de Derechos de Autor. Número de certificado de licitud de título: 13199 y número de certificado de licitud de contenido: 10772, ambos otorgados por la Comisión Calificadora de Publicaciones y Revistas Ilustradas de la Secretaría de Gobernación.

Andamios. Revista de Investigación Social es una publicación periódica, científica, internacional, arbitrada, indexada y de acceso abierto y gratuito tanto para autores como para lectores ("revista diamante"), cuyo objetivo es contribuir en las tareas de investigación, difusión y enseñanza en materia de ciencias sociales y humanidades. Emplea un sistema estricto de arbitraje en la modalidad "doble ciego" y quienes evalúan son personas de reconocido prestigio en el ámbito de la investigación, tanto nacional como extranjera. Se publica en lengua española y en formato electrónico tres veces al año y está adscrita al Colegio de Humanidades y Ciencias Sociales de la Universidad Autónoma de la Ciudad de México (UACM).

Andamios. Revista de Investigación Social es reconocida por el Sistema Nacional de Publicaciones Científicas y Humanísticas de la Secretaría de Ciencia, Humanidades, Tecnología e Innovación de México. Está altamente posicionada en distintas categorías de revistas académicas como en Scopus y su SCImago Journal & Country Rank en el nivel Q2 del área de Artes y Humanidades, y en el nivel Q3 del área de Ciencias Sociales. Pertenece a múltiples índices y bases de datos, entre los cuales destacan: Scielo-México, Redalyc, Web of Science (WoS) y sus Social Sciences Citation Index (SSCI) y Journal Citation Reports (JCR), En EBSCO y sus Academic Search Ultimate Magazines and Journals, Fuente Académica Plus Database Coverage List, Sociology Source Ultimate Database Coverage List, en International Bibliography of the Social Sciences (IBSS), Political Database of the Americas (PDBA), USC Libraries. University of Southern California, Latindex, Dialnet, Biblat, LatinRev, Matriz de Información para el Análisis de Revistas MIAR, HAPI Search the Hispanic American Periodicals Index (HAPI) for journal article citations about Latin America, the Caribbean, and Hispanics/Latinos in the US, entre otros. En su versión electrónica tiene presencia a través del portal institucional de la Universidad Autónoma de la Ciudad de México (UACM) y el Open Journal Systems (OJS).

Los artículos contenidos en esta publicación son responsabilidad de sus respectivos autores y no comprometen la posición oficial de Andamios, Revista de Investigación Social ni de la Universidad Autónoma de la Ciudad de México.

Se autoriza la reproducción parcial de los contenidos de la presente publicación siempre que se cite la fuente.

En portada: *Un quartier embrovillé*, de Albert Robida

Directorio

Directora

Leticia Romero Chumacero

Editor responsable

Oscar Rosas Castro

Comité Editorial UACM

Álvaro Aragón Rivera
Concepción Delgado Parra
Adrián Espinosa Barrios
Grissel Gómez Estrada
Gezabel Guzmán Ramírez
Jesús Jasso Méndez
Julieta Marcone Vega
Cuauhtémoc Ochoa Tinoco
Norma Angélica Gómez Méndez
Nicolás Olivos Santoyo
Sergio Ortiz Leroux
Cynthia Pech Salvador
Leticia Romero Chumacero
Édgar Sandoval Sandoval
Arturo Santillana Andraca
Ángel Sermeño Quezada
Citlali Villafranco Robles

Equipo de Redacción

Mara Itzel Georgina Montes Margalli
Oscar Rosas Castro

Consejo Editorial

Benjamin Arditi (FCPyS-UNAM, México)
Tatiana Bubnova (IIFL-UNAM, México)
Belem Clark de Lara (IIFL-UNAM, México)
Luiz Augusto Campos (IESP-UERJ, Brasil)
Andrés de Francisco Díaz (Universidad Complutense de Madrid, España)
Rodrigo Díaz Cruz (UAM-I, México)
Ángel José Fernández Arriola (Universidad Veracruzana, México)
Gustavo Fondevila (CIDE, México)
Raúl Fuentes Navarro (Universidad de Guadalajara, México)
Jorge A. González (CEIICH-UNAM, México)
Gustavo Leyva Martínez (UAM-I, México)
María Teresa Muñoz Sánchez (FFyL-UNAM, México)
Denise Najmanovich (UBA, Argentina)
Ana Rosa Pérez Ransanz (IIF-UNAM, México)
Sara Poot-Herrera (Universidad de California en Santa Barbara, Estados Unidos de América)
Octavio Rodríguez Araujo (FCPyS-UNAM, México)
Mario Rufer (UAM-X, México)
Alejandro Sahuí Maldonado (CIJ-Universidad Autónoma de Campeche, México)
José Ma. Sauca Cano (Universidad Carlos III de Madrid, España)
Hugo José Suárez (IIS-UNAM, México)
Isabel Wences (Universidad Carlos III de Madrid, España)
Lauro Zavala (UAM-X, México)

Diseño

Miguel Angel Luna Vilchis / Punto Áureo

Índice

Presentación
Leticia Romero Chumacero 9

Andamios y el acceso abierto: ¿modelo de negocio lucrativo o modelo de ciencia como bien público?
Sergio Ortiz Leroux 13

Autor(IA) y escritura académica
Leticia Romero Chumacero y Adrián Espinosa Barrios 25

La pregunta auténtica heideggeriana que la IA es incapaz de fagocitar
Concepción Delgado Parra 37

¿Qué significa pensar en tiempos de IA? Una perspectiva fenomenológica
Edgar Sandoval 55

Democracia algorítmica e inteligencia artificial: una aproximación
Ángel Sermeño Quezada 69

Entre sintaxis y normatividad: una aproximación lógica y computacional a los problemas éticos de la inteligencia artificial
Jesús Jasso Méndez 85

Primer momento de la Inteligencia Artificial en la edición de revistas académicas de Humanidades y Ciencias Sociales
Oscar Rosas Castro 105

¿Quién responde por la Inteligencia Artificial? Ética, sesgos y responsabilidad en la investigación y la escritura académica
Gezabel Guzmán y Aliber Escobar 119

Dossier
Desafíos de las revistas académicas hoy: Acceso abierto e inteligencia artificial (IA)

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1341

Presentación

Leticia Romero Chumacero*

Andamios. Revista de Investigación Social recibe cada año una cantidad muy importante de artículos, casi el ochenta por ciento de los cuales no se publican. Una de las razones por las que ocurre esto se relaciona con la primera fase de evaluación, la cual consiste en examinarlos con un *software* de detección de similitudes e inteligencia artificial (IA).[1] Ese proceso es supervisado por quienes integran la Redacción y, si detectan plagios o uso no declarado y justificado de IA, la persona proponente recibe un mensaje de correo electrónico donde se le informa que su texto ha quedado fuera del proceso de dictaminación. En más ocasiones de las que nos gustaría admitir, la respuesta es de profunda indignación y agravio.

Con todo, hay quien –tras negarlo– admite el uso de IA, pero solicita no incluir esa información en el artículo, acaso con la certeza de que tal dato colocaría en entredicho la autoría reclamada; alguien más, a falta de argumentos, despotrica contra la revista donde su audacia fue descubierta; no falta quien jura no conocer otra inteligencia que la propia; y también los hay que se retiran en ominoso y sugerente silencio. Pero ciertas personas responden con legítima sorpresa ante el reproche por el uso de IA en su artículo. En estos casos es menester explicarles que, si al usar el procesador de textos de Word aceptan la generosa oferta de Copilot de "reescribir el párrafo", o en un archivo PDF dicen sí a la propuesta de hacer un "resumen automático", o si incluyen en el documento una traducción hecha con el

[1] Cfr.: https://uacm.edu.mx/andamios/Indicaciones-para-los-autores#undefined Las otras dos fases de evaluación incluyen revisiones en la modalidad "doble ciego" y constituyen la etapa donde la mayoría de los artículos son rechazados.

* Profesora-investigadora de la Academia del Posgrado en Estudios Semióticos de la UACM y directora de *Andamios. Revista de Investigación Social*. Es fundadora de la Red Internacional de Mujeres Científicas y miembro del Sistema Nacional de Investigadoras e Investigadores. ORCID: 0000-0003-4960-2339 Correo electrónico: leticia.romero@uacm.edu.mx

traductor de Google... ya ingresaron en los terrenos de la inteligencia artificial. En efecto, hay quien no lo sabe. Esto devela la paradoja de un analfabetismo tecnológico alimentado por la omnipresencia de la tecnología –en este caso, la relativa a la IA–: al figurar en toda clase de dispositivos, ese recurso se antoja invisible y, por ende, no se acredita.

Entre el analfabetismo tecnológico y el fraude intencional, el uso de la IA es hoy uno de los temas caros a la discusión editorial mundial. Cada revista académica ha tomado decisiones relativas a eso. La prohíben, la permiten con salvedades o la admiten sin límites, dependiendo de la forma como se entienden en esos espacios temas como autoría, pensamiento, erudición, democratización del conocimiento, integridad académica y plagio, entre otros.

Sin dejar de celebrar el valor instrumental de la IA, en este número especial *Andamios* se suma al concierto de voces que advierten el riesgo de los sesgos, la univocidad, la concentración de poder, la hueca publicación a destajo y la mercantilización del conocimiento, asociados al uso irracional de esa herramienta. Aquí se reflexiona, pues, en torno a la responsabilidad intelectual como un hecho exclusivamente humano. Las y los autores coinciden en no rechazar el uso de la inteligencia artificial, pero sí impedir que su lógica de cálculo se erija en el modelo exclusivo del conocimiento, del quehacer académico, de la escritura y de las decisiones políticas. También concuerdan al indicar cuán relevante es no dar carta de naturalización a esa tecnología como sustituto de la actividad intelectual pues, si bien poderosa, no deja de ser un instrumento.

Antes de llegar hasta ahí, el artículo que abre el volumen lleva el debate hacia otro asunto editorial por el cual cruzan tanto la ética como el capital. En las revistas académicas, ¿es posible compatibilizar el reconocimiento editorial con el acceso abierto no comercial? *Andamios* es una publicación muy bien situada en los prestigiosos índices y bases de datos internacionales. Ahí, las posiciones más altas son ocupadas por revistas donde se cobran tarifas por concepto del procesamiento de los artículos y se cobra igualmente a quienes desean consultar sus páginas: mucho de negocio hay detrás de esos proyectos editoriales. En contraste, la revista de investigación del Colegio de Humanidades y Ciencias Sociales de la Universidad Autónoma de la Ciudad de México sigue otro esquema. *Andamios* es una revista "diamante", esto es, una publicación científica e internacional de acceso abierto, libre y gratuito para quien escribe y para quien lee.

¿Qué implicaciones éticas, políticas y editoriales representa para una revista financiada por una universidad pública mexicana sentarse ante una mesa internacional para dialogar con quienes pagan por un lugar?, ¿qué supone hacerlo cuando se ha suscrito un modelo no comercial, mas riguroso, donde la ciencia es considerada un bien público?, ¿son compatibles la autonomía académica y las plataformas privadas de evaluación editorial? Y, respecto de la IA: ¿qué implicaciones éticas, políticas y editoriales entraña para una revista académica el uso, declarado o no, de esa herramienta?, ¿cómo inciden sus sesgos en una publicación periódica que aboga por la ciencia como bien público?, ¿qué aspectos de la producción del conocimiento se le pueden delegar y cuáles no?

A continuación, las preguntas y nuestros atisbos de respuesta.

Otoño de 2026, en el 22 aniversario de *Andamios*
y el 25 aniversario de la Universidad Autónoma de la Ciudad de México.

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1341

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1343

Andamios y el acceso abierto: ¿modelo de negocio lucrativo o modelo de ciencia como bien público?

Sergio Ortiz Leroux*

¿Cómo distinguir las diferentes modalidades de revistas académicas? Diversos criterios pueden utilizarse en este ejercicio clasificatorio. Un criterio básico consiste en la distinción entre revistas académicas de *investigación* y de *divulgación*. Las primeras son publicaciones periódicas sometidas a reglas de dictaminación anónimas por pares, llamadas de "doble ciego", que ofrecen salida a investigaciones de mediano y largo alcance de comunidades académicas locales, nacionales o internacionales. Las segundas, en cambio, son revistas que promueven reportes de investigación preliminares, notas informativas, reportajes y entrevistas que no se someten a normas de dictaminación, salvo las consabidas reglas formales de redacción, sintaxis, ortografía y puntuación. Otro criterio muy común que suele utilizarse para distinguir a las revistas académicas es el de las *áreas de conocimiento*. Se trata de diferenciar a las publicaciones en función de si corresponden al campo de las artes y humanidades, las ciencias sociales y jurídicas, las ciencias naturales y exactas, las ciencias de la salud o la ingeniería y arquitectura. A partir de esta clasificación, han adquirido carta de naturalidad en todo el mundo un sinnúmero de revistas especializadas en diferentes disciplinas y subdisciplinas. De ahí que puedan distinguirse revistas especializadas, por ejemplo, en sociología o en filosofía; o publicaciones focalizadas en campos disciplinares, como, por ejemplo, la filosofía política, la filosofía del

* Profesor investigador de la Academia de Ciencia Política y Administración Urbana de la UACM. Miembro del Sistema Nacional de Investigadoras e Investigadores e integrante del Grupo de Investigación de Teoría y Filosofía Política. Fue director de *Andamios. Revista de Investigación Social* del 1° de febrero de 2010 al 28 de febrero de 2013. Actualmente es miembro del Comité Editorial de *Andamios*. ORCID: 0000-0002-6417-4582 Correo electrónico: sergio.ortiz@uacm.edu.mx

derecho, la filosofía social, la filosofía moral o la filosofía de la economía. Un criterio más que puede usarse es la diferencia entre revistas académicas *endógenas* y *exógenas*. Las primeras difunden, prioritariamente, la producción académica de los integrantes de la propia institución de adscripción de la revista. Las segundas, en contraste, suelen dar a conocer, principalmente, investigaciones de profesores e investigadores que no están adscritos a la institución que da acogida a la revista, dado que pertenecen a otras universidades o centros de investigación. La práctica de la endogamia en las revistas académicas es uno de los indicadores que más se vigila desde los índices y bases de datos, pues se trata de evitar que estas publicaciones sean espacios cerrados y autorreferenciales.

A partir de los criterios anteriores se puede afirmar que *Andamios. Revista de Investigación Social* del Colegio de Humanidades y Ciencias Sociales (CHyCS) de la Universidad Autónoma de la Ciudad de México (UACM) es una publicación académica de investigación, en el campo de las Ciencias Sociales y las Humanidades, a partir de las aportaciones de los profesores-investigadores de la UACM, así como de académicos de otras instituciones públicas y privadas nacionales e internacionales. No es, hay que decirlo, una publicación endógena, sino es más bien exógena, ya que menos del 15% del total de artículos de investigación publicados en la sección "Dossier" y en la sección "Artículos" en los primeros 58 números de la revista (otoño-invierno de 2004 / mayo-agosto de 2025) corresponden a autores o autoras de la propia UACM (Rosas, 2025, p. 21).

Sin embargo, existe otro criterio, no menos importante que los anteriores, que suele ser desconocido o menospreciado –salvo por los iniciados o *senséis* en estos asuntos– a la hora de clasificar a las revistas académicas. Se trata de la distinción a partir del criterio que pone acento en el *acceso abierto* de la publicación, especialmente en la diversidad de *formas de acceso*, de *modelos de publicación* y de *modelos de negocio* que posibilitan genéricamente la comunicación científica abierta. Al criterio del acceso abierto y sus modelos prácticos y a los principios que subyacen a los mismos se dedicarán las siguientes líneas, con la mirada puesta en localizar las coordenadas que permiten precisar, y al mismo tiempo evaluar, la especificidad de la revista *Andamios* en el ecosistema de revistas académicas que circulan hoy en día en México, el resto de América Latina, Estados Unidos y Europa.

I. Acceso abierto, modelos de publicación y modelos de negocio. Lo que todos deberíamos saber, pero muchos tememos preguntar

Dice Platón en *La República* que al que no sabe le corresponde preguntarle al que sabe. Si hay una persona que se ha dedicado en nuestro país a investigar desde hace varios años el problema de las transformaciones y desafíos recientes del acceso abierto de las revistas académicas y científicas de México, América Latina, España y Portugal, y a crear, al mismo tiempo, un sistema de indización como *Redalyc*[1] para garantizar dicho acceso más allá de la lógica de lucro del mercado, esa persona es el profesor e investigador mexicano Eduardo Aguado López. En el artículo "La privatización del acceso abierto a veinte años de las tres B", Aguado López (2021) ofrece un recorrido histórico sobre el acceso abierto a partir de diferenciar las formas de acceso, los modelos de publicación y los modelos de negocio que posibilitan la comunicación científica abierta. Por *formas de acceso abierto* se hace referencia a la existencia o no de barreras legales, técnicas o de pago para acceder a las revistas académicas, lo que, a su vez, se traduce en un *modelo de publicación* y supone un *modelo de negocio* que posibilita la comunicación científica abierta. El primero designa a las infraestructuras tecnológicas e institucionales, canales y prácticas que dan forma a un modo particular de comunicación científica; mientras el segundo hace referencia a los recursos que dan sostenibilidad a los modelos de publicación (Aguado, 2021, p. 38).

Entre las formas de acceso abierto destacan principalmente dos modalidades: una a partir de revistas científicas que se distinguen por la revisión

[1] La Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (Redalyc) fue fundada en 2003 como un proyecto académico de la Universidad Autónoma del Estado de México (UAEM), bajo la responsabilidad del cuerpo académico "Difusión y Divulgación de la Ciencia", con el fin de dar visibilidad, mejorar y consolidar la calidad editorial de las revistas de Ciencias Sociales y Humanidades de la región latinoamericana. En 2006 se abrió a todas las áreas del conocimiento e incluyó revistas de la península ibérica. Se trata de un sistema de indización que integra a su índice las revistas de alta calidad científica y editorial de Iberoamérica. Después de 16 años de dar visibilidad y apoyar en la consolidación de las revistas, ahora integra de manera exclusiva a las que comparten el modelo de publicación sin fines de lucro para conservar la naturaleza académica y abierta de la comunicación científica de cualquier región del mundo (Redalyc, 2026).

de pares; y la otra a partir de los llamados repositorios, que son espacios virtuales donde se almacenan, organizan y mantienen datos, archivos o códigos digitales para que puedan ser recuperados o compartidos sin restricciones a través de Internet. El acceso abierto se conceptualizó en una primera etapa en términos de acceso y los tipos que se abordaron fueron la *vía verde*, correspondiente a los repositorios institucionales y temáticos de acceso abierto, y la *vía dorada*, correspondiente a las revistas científicas con artículos revisados por pares. Sin embargo, con el trascurrir del tiempo este esquema dicotómico de acceso abierto comenzó a complejizarse, ya que se dejó de considerar exclusivamente el asunto del acceso y se comenzaron a considerar también elementos de los modelos de publicación y de negocio, lo que derivó en una multiplicidad de prácticas de las revistas académicas. El asunto es complejo, tiene varias aristas y matices y ameritaría, por lo tanto, un mayor desarrollo y precisión técnica. Pero, para los propósitos de este breve ensayo introductorio y de divulgación, bastaría con señalar que un dato o momento clave en esta historia fue el surgimiento del modelo de publicación con *Article Processing Charge* (APC), ya que inició el viraje de un modelo de pagar por leer (vía suscripciones) a otro de pagar por publicar. Revistas científicas y organismos de investigación, asentados principalmente en países del Norte, junto con grupos editoriales privados,[2] hicieron suyo este nuevo modelo de publicación, adaptando las bases del acceso abierto a un modelo de negocio privado en beneficio de ellos, a través, precisamente, del APC.

La transformación abierta por el APC no sería menor, pues provocaría una verdadera metamorfosis en la clasificación y visualización de sí mismas de las revistas académicas y de investigación. La nueva taxonomía, signo de los nuevos tiempos marcados por una economía neoliberal de libre mercado, es la siguiente. En un principio, la categoría de *vía dorada* –como aquí se ha señalado– se asoció a la de las revistas científicas con contenidos abiertos, más allá del modelo de negocio que tuvieran. Después del surgimiento del APC esta vía de acceso abierto conservó su nombre y, con ello, inició

[2] Los veinte principales editores comerciales en el mundo son, según Aguado López (2021, p. 52), los siguientes: Aosis, BMC, Copernicus Publications, De Gruyter, Dove Medical Press, Elsevier, Emerald Publishing, Frontiers Media, Hindawi, MDPI, Nature Publishing Group, PAGEPress Publications, Pensoft Publishers, SAGE Publishing, Sciendo, Springer Open, Taylor and Francis Group, Ubiquity Press, Wiley y Wolters Kluwer.

un cambio relevante en las categorías de clasificación: "con *vía dorada* de AA comenzó a hacerse referencia a la comunicación por medio de revistas científicas con contenidos abiertos, pero con cargos por APC; mientras que las revistas sin APC y sin suscripciones buscaron una especificidad en la categorización del AA, hasta entonces dicotómica (*verde y dorada*), de forma que comenzó a designarse a este modelo de publicación / negocio como *AA diamante*" (Aguado, 2021, p. 40, cursivas mías).

Hasta donde alcanzo a entender (quizá mi entendimiento no es de largo alcance), la oposición entre la vía dorada *revisitada* (APC *dixit*) y la vía diamante, o dicho de otra manera, entre las *tendencias comerciales* y las *tendencias no comerciales* que ha seguido en las dos últimas décadas el acceso abierto tanto en México y el resto de América Latina, como en Estados Unidos y Europa, es el clivaje que ha focalizado y alimentado en la actualidad el debate sobre el sentido y la finalidad de las revistas académicas.[3] En contraste con la vía dorada, la vía diamante se distingue, esencialmente, por los siguientes elementos: *a)* ausencia de cualquier cobro por acceso a las revistas científicas o de cualquier cobro por recepción, dictaminación, publicación o divulgación de artículos de investigación; *b)* participación de instituciones educativas, académicas y sin fines de lucro, como las universidades públicas,

[3] Otro tema relevante en la discusión sobre el sentido y finalidad de la ciencia y de las publicaciones científicas corresponde a los llamados *sistemas de evaluación científica*. En el caso de las revistas académicas, el criterio principal que se utiliza para evaluar el *prestigio editorial* de éstas fue establecido por bases de datos internacionales, que, a su vez, son empresas privadas, como WoS y Scopus, a través de su Scimago Journal Rank (Rosas, 2025, p. 18). Desde ese índice se organiza, con la denominación de *Quartil*, cuáles son las revistas que pueden ingresar a la lista de las más destacadas a nivel mundial. En este marco de referencia, el principal criterio que se usa para definir la lista y el lugar que se ocupa en la misma es el llamado *Factor de Impacto*, número que refiere a la cantidad de citas que se hacen de un texto publicado en un periodo de hasta dos años después de su fecha de aparición. A partir de estas consideraciones, se establecen distinciones entre revistas del *mainstream* (digamos, ahora que estamos en plena fiebre mundialista, de la *Champions League*), y revistas *periféricas* (digamos de la *Liga MX* o de la *Copa del Rey*). Discutir actualmente los sistemas de evaluación científica en general y, en particular, los criterios para evaluar el prestigio editorial de las revistas académicas, es un asunto que no puede pasar desapercibido por autores, dictaminadores, editores y lectores, sobre todo si se considera que estas cuestiones no sólo son de naturaleza técnica o procedimental, sino también están cruzadas por consideraciones éticas, filosóficas y geopolíticas. Quizá en otro lugar profundizaré en esta apasionante, y también preocupante, problemática.

en la gestión y subvención financiera y material de las revistas científicas; y *c)* uso de licencias abiertas para la reutilización de los contenidos y la promoción de una comunidad científica diversa, inclusiva y localmente pertinente que se expresa en el multilingüismo.

II. Principios de los modelos dorado y diamante

Los modelos de publicación y negocio de las revistas académicas no son plenamente neutrales. No son almas bellas hegelianas. Detrás de cada uno de ellos se encuentran presentes implícita o explícitamente un conjunto de principios que responden a consideraciones filosóficas, políticas y éticas diferentes y acaso encontradas. En el caso del modelo dorado, no siempre se hacen explícitos sus principios por parte de editores y patrocinadores de esas publicaciones, pero de las prácticas concretas de las revistas académicas con contenidos abiertos pero con cargos por APC, se puede deducir que defienden una suerte de *mercantilización* de las publicaciones científicas, cuyos costos se cargan no a los lectores, que siguen accediendo gratuitamente a sus contenidos, sino más bien a los autores, que se ven obligados o forzados a pagar tarifas –sea para que su texto sea recibido e ingresado a la fase de dictaminación anónima por pares, o sea, en el mejor de los casos, para ser publicado– cuyos montos dependen comúnmente del prestigio de la revista académica. Éste es medido a partir de bases de datos internacionales, como Wos y Scopus, que, en los hechos, son empresas privadas. Esta mercantilización tiene como consecuencia la *exclusión*, ya que aquellos investigadores que no puedan pagar o, en su caso, se nieguen a pagar esas tarifas porque las consideren moral o políticamente inaceptables, serán automáticamente excluidos de la posibilidad de publicar. Mercantilización y exclusión son, en pocas palabras, las dos caras de una forma explícita de *privatización* del acceso abierto al conocimiento científico generado por las revistas académicas.

En contraste con el modelo dorado, el modelo diamante concibe a la ciencia como un *bien público global* a partir de las premisas y prácticas del acceso abierto no comercial de las revistas académicas y científicas (Cumbre Global sobre Acceso Abierto Diamante, 2023). A la luz de este modelo, la ciencia en general y el conocimiento científico producido por las revistas académicas en particular, son representados como bienes públicos globales

en dos de los sentidos de la palabra *público* en oposición a la palabra *privado*: *a)* en primer lugar lo público alude a lo que es de interés o de utilidad común a todos, lo que atañe al colectivo, lo que concierne a la comunidad, en oposición a lo privado, entendido como aquello que se refiere a la utilidad e interés individual; y *b)* en segundo lugar lo público designa lo que es accesible, abierto a todos, en oposición a lo privado, entendido como aquello que se sustrae a la disposición de los otros (Rabotnikof, 2000, p. 604).

Cinco grandes principios (éticos, filosóficos, políticos, etcétera) rigen a la idea de la ciencia como un bien público global:

1. *Derecho universal*. La ciencia es un bien público global y el acceso a ella es un derecho universal;
2. *Equidad, diversidad y multilingüismo*. La ciencia es inclusiva, multilingüe, accesible, reutilizable y colaborativa;
3. *Propiedad de la academia y patrimonio de la humanidad*. La producción científica es propiedad de la academia y se debe al desarrollo y progreso de la sociedad como patrimonio de la humanidad;
4. *Reconocimiento y valoración*. Las entidades de acreditación, investigación y financiación deben reconocer, evaluar e incentivar los medios no comerciales de producción y circulación del conocimiento científico; y
5. *Colaboración*. La interacción y colaboración entre los agentes no comerciales, publicaciones científicas e infraestructuras abiertas es necesaria para la construcción de ecosistemas de bienes públicos (Cumbre Global sobre Acceso Abierto Diamante, 2023).

III. El lugar y el papel de *Andamios* en el ecosistema de revistas académicas de acceso abierto. A manera de conclusión

¿Qué lugar ocupa la revista *Andamios* en el ecosistema de revistas académicas de acceso abierto de México, América Latina, Estados Unidos y Europa? ¿Qué posiciones públicas ha defendido esta publicación (o algunos de los o las integrantes de su Equipo de Redacción o Comité Editorial) en la discusión académica sobre las formas de acceso, los modelos de publicación y los modelos de negocio?

Desde su nacimiento, en el otoño-invierno de 2004, *Andamios* ha sido una revista de acceso abierto a la totalidad de sus contenidos a través de las versiones electrónicas disponibles en la página de la revista dentro del portal *web* de la UACM y en la plataforma especializada en revistas científicas Open Journal System (OJS). Se encuentra, asimismo, bajo una licencia Creative Commons México 2.5, que permite la reproducción y difusión de sus contenidos para fines educativos o de investigación, sin ánimo de lucro, siempre y cuando éstos no se mutilen y se cite la procedencia (*Andamios. Revista de Investigación Social*) y a la autora o autor.

Jamás la revista ha cobrado (ni cobrará en el futuro) un solo peso a aquellas o aquellos profesores-investigadores internos o externos a la UACM que: *a)* manden sus textos a la redacción de la publicación; *b)* sometan sus textos a dictamen bajo la regla del "doble ciego"; y *c)* publiquen sus textos en las páginas de la revista. Es cierto que la revista *Andamios*, cuando aparecía en versión impresa (número 1 del 2004 al número 35 del 2017), tenía un precio de venta al público en general y promovía suscripciones anuales. Pero, en honor a la verdad, la rendición de cuentas y la transparencia, los ingresos recibidos por esas vías siempre fueron marginales y no cubrían ni el 5% del costo total de su edición, impresión y distribución.

Desde los primeros años de aparición de la revista, el Comité Editorial de *Andamios* se adhirió a las políticas y principios rectores del acceso abierto y se sumó más tarde con entusiasmo a la vertiente del modelo diamante del acceso abierto no comercial. Prueba de ello es la asistencia regular y comprometida del Editor de la revista, Oscar Rosas Castro, a las numerosas reuniones, eventos y simposios académicos, encuentros y congresos de editores de revistas universitarias.

Dos razones, por lo menos, explican la adhesión de *Andamios* al "movimiento", si se me permite la expresión, a favor del acceso abierto y no comercial de las revistas académicas: *a)* la adscripción institucional de la revista. *Andamios* es una publicación del CHyCS de la UACM, una institución de educación superior pública, gratuita e incluyente, que en la exposición de motivos de su Ley defiende, entre otras cosas, la autonomía académica y el carácter público de la educación y del conocimiento científico y humanístico. Ser editada y financiada por una universidad pública y gratuita ha sido la principal garantía de su independencia material, intelectual y también, por

qué no, moral; y *b)* las directoras y los directores e integrantes de su Comité Editorial han sido promotores, desde sus respectivas áreas de conocimiento y adscripciones disciplinares, del carácter público de la educación, así como del carácter abierto de las ciencias sociales y las humanidades.

Ahora bien –y para terminar–, el piso firme de *Andamios* se vuelve más pantanoso si se pasa del asunto del acceso abierto no comercial a la cuestión del prestigio editorial de la revista. Para abordar este aspecto, convendría, como diría Jack "el Destripador", avanzar por partes. *Andamios* realiza su difusión a través de los recursos tecnológicos actuales. De suerte que reconoce el valor de publicar sus contenidos en las bases de datos, índices, catálogos y bibliotecas electrónicas nacionales e internacionales. La lista es muy amplia y variada y puede consultarse en la página *web* de la revista.[4] Lo cierto es que esta proactiva política de difusión de la revista, por llamarla de alguna manera, ya ha rendido sus frutos: hoy en día *Andamios* está *altamente posicionada*, es decir, goza de un *elevado prestigio editorial* en distintas categorías de revistas académicas:

a. Aparece catalogada dentro de Redalyc y Scielo-México;
b. Fue distinguida con la categoría de Excelencia académica por el ex Consejo Nacional de Humanidades, Ciencias y Tecnología (ahora Secretaría de Ciencia, Humanidades, Tecnología e Innovación del Gobierno de México) y forma parte actualmente del Sistema de Clasificación de Revistas Mexicanas de Ciencia y Tecnología de la misma institución; y
c. La empresa Scopus, la más importante base de datos internacional, reconoce a *Andamios*, a través de su famoso Scimago Journal Rank (SJR), en el nivel Q2 del Área de Artes y Humanidades y en el nivel Q3 del Área de Ciencias Sociales.

Todo sería miel sobre hojuelas si no fuera por el dato duro de que Scopus es una empresa privada con fines de lucro que opera a partir del llamado Factor de Impacto (explicado en la cita a pie de página número 3), en el que los mejores lugares en el *ranking* suelen estar reservados a revistas académicas de países del Norte, cuyos modelos de publicación y de negocio

[4] A saber: https://uacm.edu.mx/andamios

suelen ser de acceso abierto, pero con cargos por APC.[5] Las revistas del modelo diamante, en este escenario, difícilmente accederán a los niveles Q1 en las diferentes áreas de conocimiento, porque, entre otras razones, su modelo de publicación y negocio es contrario a los principios que rigen a las revistas académicas internacionales con mayor prestigio editorial. De ahí que las revistas académicas del modelo diamante de México y de otros países latinoamericanos no suelan tener la trascendencia suficiente para estar en los primeros lugares en el llamado Factor de Impacto de la empresa Scopus.

Andamios ha ganado, legítimamente, prestigio editorial en la comunidad académica mexicana e iberoamericana por varias razones: por la regularidad y puntualidad en cada una de sus entregas; por la pluralidad teórica, epistemológica y de áreas de conocimiento de sus dosieres y artículos, así como la originalidad y pertinencia de los mismos; por no ser una revista de carácter endógeno y de autoconsumo; por ser una revista académica que se ha mantenido al margen de las inevitables *grillas* universitarias promovidas por grupos de interés; por el profesionalismo e imparcialidad de sus procesos editoriales internos; por su contribución a la formación y consolidación de grupos y redes temáticas de investigación nacionales e internacionales. Pero también, hay que decirlo, porque está clasificada y bien posicionada en las principales bases de datos e índices nacionales e internacionales. Por eso recibe cada año una cantidad considerable de artículos a dictaminar de autores nacionales e internacionales.

Es momento de que el Comité Editorial de la revista abra una discusión académica seria, argumentada y con la suficiente evidencia empírica sobre los problemas y desafíos planteados en este breve ensayo. ¿Necesariamente son contradictorios los aspectos del acceso abierto no comercial y el prestigio editorial? ¿Será que, con una razón nómada (Pereda, 2021), en algunos momentos se pueda poner acento en el aspecto del acceso abierto no comercial y en otros momentos, en cambio, en la cuestión del prestigio editorial? Quizá un buen punto de partida para iniciar y profundizar en esta discusión de evidente actualidad sea la convocatoria pública de un futuro *dossier* de *Andamios* dedicado, precisamente, a la temática de la "Difusión

[5] No entraré ahora en la discusión sobre la calidad de las revistas académicas del *mainstream*, que suelen ser de los países del Norte Global, y de las revistas académicas *periféricas*, que suelen ser de los países del Sur Global, asunto que podrá abordarse en otro momento.

y divulgación de las Ciencias Sociales y las Humanidades en México y América Latina", donde se podrían abordar puntualmente y con rigor los nuevos retos que supone en nuestra región el acceso abierto comercial y no comercial de las revistas académicas. Al tiempo.

Fuentes consultadas

Aguado, E. (2021). De la esperanza al fracaso. La privatización del acceso abierto a veinte años de las tres B. En A. Becerril-García y S. Córdoba (Eds.). *Conocimiento abierto en América Latina. Trayectorias y desafíos*. pp. 37-78. UAEM/CLACSO.

Cumbre Global sobre Acceso Abierto Diamante (2023). *Manifiesto sobre la ciencia como bien público global: Acceso abierto no comercial.* (2026-07-09). Disponible en: https://globaldiamantoa.org/manifiesto/#/

Pereda, C. (2021). *Pensar a México, entre otros reclamos*. Gedisa / UNAM.

Rabotnikof, N. (2000). Público-privado. En L. Baca, J. Bokser, et al. (Coords.). *Léxico de la política*. pp. 604-607. FCE/FLACSO-México.

Redalyc (2026). *Misión de la Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (Redalyc).* Disponible en: https://www.redalyc.org/

Rosas, O. (2025). Sentido y valores de una revista académica, el caso de Andamios. En *Andamios. Revista de Investigación Social.* Vol. 22. Número Especial. pp. 13-34. DOI: https://doi.org/10.29092/uacm.v22iEspecial.1210

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1343

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1344

Autor(IA) y escritura académica

Leticia Romero Chumacero*
Adrián Espinosa Barrios**

Te das cuenta de que se necesita una buena dosis
de inconsciencia para confiarse a máquinas inseguras,
guiadas imprecisamente; o quizá esto prueba
una imparable tendencia hacia la pasividad,
a la regresión, a la dependencia infantil.
Ítalo Calvino

Aquí se habla de la figura autoral en la escritura académica. Sólo para abrir boca, examinemos algunas noticias frescas: un recién graduado de Berkeley afirmó haber elaborado poco más de 100 artículos académicos en un año; en el mismo lapso, un español publicó 176. Sus detractores señalan la exigua calidad de esos trabajos y los presumen generados con inteligencia artificial (IA) (Ansede, 2023; Down, 2025). Apenas iniciado el año 2026, un detector de uso de IA identificó decenas de "alucinaciones", es decir, datos falsos –por ejemplo, fuentes inexistentes– en artículos publicados en las páginas de influyentes revistas de salud, algunas de las cuales se han visto obligadas a retirar esos textos fraudulentos, pues podrían inducir a tratamientos médicos

* Profesora-investigadora de la Academia del Posgrado en Estudios Semióticos de la UACM y directora de *Andamios. Revista de Investigación Social.* Es fundadora de la Red Internacional de Mujeres Científicas y miembro del Sistema Nacional de Investigadoras e Investigadores. ORCID: 0000-0003-4960-2339 Correo electrónico: leticia.romero@uacm.edu.mx
** Profesor-investigador de la Academia de Filosofía e Historia de las Ideas de la UACM e integrante del Comité editorial de *Andamios. Revista de Investigación Social.* Es integrante del Seminario Universitario sobre Afectividad y Emociones (SUAFEM), de la Universidad Nacional Autónoma de México y miembro del Sistema Nacional de Investigadoras e Investigadores. ORCID: 0000- -0003-3377-4021 Correo electrónico: adrian.espinosa@uacm.edu.mx

peligrosos (Shmatko *et al.*, 2026). En México, el año pasado la Suprema Corte de Justicia de la Nación estableció que "las obras creadas por inteligencia artificial no pueden registrarse como derechos de autor", ¿su argumento?: la ley sólo protege la creatividad humana (INDAUTOR, 2025).

Entre otras cosas, tales noticias dan cuenta de un panorama donde se van tornando sospechosas la construcción y la utilidad del conocimiento académico. Peor aún, dicho conocimiento luce entrampado en un ecosistema internacional que demanda productividad continua bajo el lema *publish or perish*. Abundan las revistas y editoriales dedicadas al predatorio negocio de las *paper mills* –fábricas de artículos–, donde la IA facilita la redacción de miles de manuscritos cuya "autoría" se ofrece en venta. Además, el marco legal del mundo entero es incipiente, tardo en su intento de poner orden al uso de una herramienta digital cuya relevancia cultural quizá es equiparable a la que en su momento protagonizó la imprenta. Muchas preguntas afloran al reflexionar sobre este contexto cuando se forma parte del Comité editorial de una revista de investigación universitaria de alta exigencia. Una de esas cuestiones se aborda en las siguientes líneas: ¿cómo debe entenderse la autoría de un artículo académico en tiempos de la IA?

¿Existe un autor?

Comencemos por el principio: la IA no es inteligencia. Es un producto de la inteligencia humana. El nombre –IA– es eso, una etiqueta que inicialmente se refería a un proyecto con un objetivo específico. Hoy es, más bien, una denominación de *marketing* exitosamente comercializada. La crítica filosófica sobre la posibilidad de que ese proyecto logre su objetivo, es decir, que en efecto sea posible "crear" una inteligencia computacional, es de larga data. En 1972, por ejemplo, Hubert Dreyfus hizo una dura crítica al proyecto de la IA a partir de la fenomenología, de la idea del sentido común y de la comprensión del proceso de adquisición de habilidades necesarias para que un sistema tenga mente (Dreyfus, 1992). Un poco más tarde, el famoso experimento mental del cuarto chino de John Searle estaba destinado a mostrar que una máquina puede tener sintaxis, pero nunca semántica, lo cual es el requisito indispensable para la comprensión y la inteligencia en general (Searle, 1980). Más recientemente, las teorías de la corporización

(Cosmelli y Thompson, 2010; Thompson, 2007) han enfatizado la imposibilidad de una "inteligencia artificial" dada la continuidad vida-mente: para que algo tenga mente, debe ser un organismo vivo.

Sin embargo, como se ha indicado, el objetivo de este escrito no es debatir sobre la presunta posibilidad de que un sistema computacional sea capaz de desarrollar cualidades mentales. El objetivo es llevar a cabo una reflexión filosófica sobre las consecuencias, en la publicación académica, del uso de esa herramienta que lleva por nombre "IA" –en particular, aquella que está basada en modelos de lenguaje generativo o LLM's–.[1] Así, pues, nuestras preguntas iniciales son: ¿los textos producidos por los LLM's tienen autor?; si un texto no tiene autor, ¿es algo más que un agregado acrítico de palabras? En ese caso, ¿qué pasa con el uso académico de estas herramientas?, ¿deberíamos permitir alguna forma de uso de la IA en algún momento del proceso de elaboración de artículos de ciencias sociales y humanidades?

En una famosa conferencia en el *College de France*, dictada en 1969, Michel Foucault (2010) reflexionó en torno a los conceptos autor, obra y escritura. El tema no era nuevo, lo había puesto sobre la mesa un año antes su amigo y colega Roland Barthes en *La muerte del autor* (Barthes, 1987). En parte, la conferencia de Foucault era respuesta a ese texto previo. En el marco del estructuralismo los temas que debatían eran de este tenor: ¿cuál es la función que cumple el autor en la producción académico-literaria?, ¿a quién nos referimos exactamente cuando pensamos en el autor de una obra? Barthes, como el título de su texto expresaba, se apresuró a decretar la muerte de la idea tradicional de autor, mirándolo como un "personaje moderno", producto de una ideología burguesa que restringió toda la importancia del texto a la "genialidad" creativa del individuo.

Hasta antes de la modernidad –digamos, hasta la Edad Media–, afirmó Barthes, la persona detrás del texto no era tan relevante. De ahí que sean escasas las expresiones artísticas –escultóricas, literarias, musicales, pictóri-

[1] Hoy en día –verano de 2026– los LLM's son modelos de propósito múltiple –generan secuencias de texto y sonido, editan imágenes, reconocen patrones, programan código, etcétera–, y de procesamiento multimodal –pueden trabajar sobre texto, audio e imágenes en tiempo real de forma simultánea y fluida–. Aquí nos referiremos específicamente al uso de estos modelos para el análisis y la generación de texto. Los principales modelos que llevan a cabo tareas de este tipo, y que tenemos en mente en este artículo, son ChatGPT 5.4 (Open AI), Claude 4.6 Opus (Anthropic), Gemini 3 (Google), Grok 4.1 (xAI), Llama 4 (Meta) y DeepSeek-V3.

cas– firmadas por quienes las elaboraron y, en cambio, en ciertos periodos predominen los trabajos gremiales o, de plano, anónimos. La comercialización del arte, vinculada con la aparición y el ascenso social de la burguesía, tornó valiosa la adquisición de piezas únicas certificadas a través de la firma de un autor. A eso se sumó la intervención discursiva de quienes se encargaron de sobrevalorar la dimensión autoral como una muestra de autonomía creativa y de expresión individual, particularmente durante el Renacimiento y buena parte del siglo XIX. Sin embargo, al iniciar el siglo XX diversas manifestaciones del arte vanguardista desafiaron la importancia, e incluso la existencia, de un autor identificable y estable. En opinión de Barthes, los lapsos donde no reinó la figura autoral demostraban que la unidad literaria no estaba en esa entidad oscura detrás del origen del texto, sino justo en la parte final del proceso creativo: el lector.

Tiempo después, Barthes matizó su postura. Ya no se trataba de desacreditar la idea de la existencia de un autor, sino de desarticularla. Detrás de ese término reconoció la presencia de, cuando menos, cuatro figuras diferentes:

- La *persona* –el individuo en su dimensión jurídica, civil y biográfica–,
- El *scriptor* –aquel que ha sido construido por la crítica, los manuales escolares y las biografías–,
- El *auctor* –aquel que se identifica y reclama la propiedad legal de lo escrito–
- El *scribens* –el sujeto físico en el acto mismo de la escritura– (Barthes, 2005).

Incluso en un trabajo de 1971 titulado *Sade, Fourier, Loyola* echó mano de la antaño todopoderosa figura autoral para sustentar sus análisis (Díaz, 2016).

Por su parte, Foucault planteó que el escritor sí existe, pero no de la forma en que creemos. Reconoció que el autor no se identifica con el sujeto creado por la crítica, porque es imposible fijar los límites de un corpus textual –¿debemos incluir las notas de lavandería de Nietzsche?–. Sin embargo, en nuestra sociedad en general, y en los contextos académicos y artísticos relevantes en particular, el autor es un "instaurador de discursividad". Es decir, el autor es aquel capaz de generar discusión a partir de sus ideas acadé-

micas o creativas. El autor tiene seguidores y opositores. Es aquel generador de discurso que, a su vez, incentiva la producción de más discurso y, con ello, de más instauradores discursivos. Por supuesto, una definición así resulta especialmente adecuada para autores que son grandes generadores de discurso, como Freud, Marx o Nietzsche. Pero la idea aplica para todos en cierto grado, pues el autor es una función compleja del discurso –no necesariamente la más importante–.

En años recientes, el tema de la autoría en el campo literario ha revelado la existencia de exclusiones simbólicas y reales (Pérez, 2019), por ejemplo, la de las mujeres. Marcadas culturalmente por planteamientos normativos según los cuales no eran capaces de generar discursividad, las escritoras debieron recurrir al anonimato o a seudónimos para no reclamar la propiedad de sus textos y, por tanto, permanecieron ignoradas por la crítica a lo largo de los siglos, incluso en los periodos de mayor auge de la figura autoral. Para seguir el planteamiento barthesiano, diríamos que, durante la mayor parte de la historia, las mujeres de letras fueron *persona* y *scriptor*, pero no *scribens* ni *auctor*. Casos como ése urgen a no hacer un fetiche de la autoría, pero sí a considerarla fundamental en el tablero donde se juegan tanto el poder para decir, como el poder para callar a –e incluso negar la existencia de– quien quiere decir.

Pero volvamos a Barthes y Foucault, quienes parten de la evidencia de la contingencia histórica del autor como lo conocemos hoy. Odas, poemas épicos, tratados, comentarios y leyendas circularon por siglos sin que nadie estuviera especialmente preocupado por establecer su autoría. Uno, dice Foucault, puede imaginar sin mucho problema una sociedad funcional donde los discursos se produzcan y reproduzcan sin echar de menos la función de autor. Es más, hoy en día no todo discurso ostenta autoría o se le reclama tal; un contrato, por ejemplo, tiene firmantes, mas no autores. Por otra parte, tenemos casos donde parece que, en forma deliberada, el autor se diluye o se esconde: los *ghostwriters*, el grupo matemático conocido como Bourbaki, las filósofas italianas del grupo Diotima, la escritura automática y colectiva desarrollada por las vanguardias, los usos mayestáticos propios del discurso papal. En el mismo tenor, un poeta mexicano ha expresado esto como su mantra: “lo importante es que el verso, y no el nombre del autor, pertenezca a la tradición”.

Alguien podría plantear que la necesidad de reconocer la autoría surge sólo donde hay ideas novedosas que expresar: en la ciencia o las artes –y no en todos los casos, como se habrá visto–. Hay autoría de las teorías científicas, en las ideas filosóficas, en buena parte de las novelas y poemas. Sin embargo, en la ciencia parece importar más la verdad que el origen del discurso. Entonces, si aceptamos por ahora que el autor es una idea históricamente reciente y que no en todos los casos de producción del discurso ha sido necesaria –ni siquiera ahí donde ha sido relevante–, ¿no podríamos reacomodar esa idea ante la nueva herramienta computacional que ya nos invade, de modo que pongamos el énfasis en el lector antes que en el productor del texto?, ¿por qué no reducir al autor sólo a una función más o menos compleja, más o menos influyente del discurso sin preocuparnos por la *persona* detrás del *scribens*? En otras palabras, ¿sigue siendo necesario defender la autoría como la conocemos? Respuesta corta: sí y no.

¿Por qué sí es necesario defender la autoría frente a la IA?

Los LLM´s son programas de reconocimiento de patrones y predicción probabilística de secuencias. No resultaría justo escamotear el alto grado de complejidad de estos sistemas, el cual es producto de décadas de arduo trabajo multidisciplinario proveniente de las áreas matemáticas, computacionales, la ingeniería y las neurociencias. La utilidad de la aplicación de esta nueva generación de *software* se evidencia en la carrera por la mejora de los modelos y por su rápida aceptación y uso entre el público en general. Con todo, es necesario repetirlo, los LLM´s no son individuos con mentalidad –bajo ningún estándar jurídico, científico o filosófico–. Por lo tanto, como se mencionó en la introducción, no se les puede adjudicar autoría. Por más que hablemos de reconocimiento, personalidad, inferencia lógica, razonamiento o intencionalidad, los LLM´s no son más que "loros estocásticos": no poseen comprensión genuina puesto que no hay una conexión semántica entre los *tokens* sobre los que se opera probabilísticamente y el mundo al que dichas unidades hacen referencia. En breve: la IA no es capaz de dotar de sentido a un mundo con el que no se relaciona. No hay sentido alguno en atribuir un estilo autoral a Gemini o a ChatGPT, o en afirmar que las versiones de pago, promocionadas con entusiasmo por las empresas desarrolladoras, poseen mejor calidad autoral que sus contrapartes gratuitas.

El problema está en una confusión de base: la confusión entre modelación, simulación e identidad. La IA ha sido útil en la modelación de los procesos mentales. Gracias a esos modelos hoy sabemos mucho más de la mente y hemos desarrollado herramientas computacionales impensables hasta hace unos pocos lustros. En términos de generación de lenguaje, lo que hay es simulación; es decir que sólo está presente un aspecto del lenguaje: aquel que es posible modelar. En definitiva, lo que produce un LLM es texto, pero no discurso.

Ahora bien, ni de la modelación ni de la simulación se sigue identidad. La modelación computacional de un huracán no moja la computadora ni la simulación de una catástrofe causa daños reales y, sin embargo, ambas pueden ser de mucha utilidad. La idea que se vende, con optimismo o pesimismo según las filias y fobias del comentarista en turno, es que a mayor trabajo de ingeniería computacional más cerca estaremos de lograr el punto de no retorno, esto es, la llamada singularidad o la inteligencia de propósito general. Si esa promesa –o amenaza– resulta cierta o no, estará por verse. Sin duda, esas herramientas son asombrosas y útiles, pero siguen siendo simulaciones. Y simulaciones que, más allá de sus virtudes, acarrean múltiples problemas.

En un mundo donde existe una constante presión por la velocidad en la producción, se generan brechas capaces de colocar en desventaja a quienes no utilizan esas herramientas computacionales para la producción académica, ya por convicción, ya por analfabetismo tecnológico, ya por razones financieras. Y la misma intención de producir mucho en poco tiempo ya derivó en malas prácticas. Es el caso del uso no declarado de IA, que las políticas editoriales de las revistas científicas no logran disuadir. Esto coloca en entredicho tanto la originalidad e integridad académicas, como la esencia misma del trabajo de investigación, de sus productos y de las publicaciones universitarias a través de las cuales circula esa farsa.

Por tanto, existen claros retos pedagógicos: hay un alto riesgo al dejar que todo reconocimiento de patrones sea procesado por una IA. Reconocer patrones e interpretar resultados por cuenta propia implica desarrollar habilidades de pensamiento crítico que cada vez serán menos fomentadas. "¿Y qué importa si el resultado, el texto, es producido por una persona con pensamiento crítico o por un LLM?", se podrá objetar. Importa por dos razones: 1) porque este razonamiento se fundamenta en una sobrevalora-

ción del producto por sobre el proceso y la habilidad, y 2) porque todo se termina pareciendo entre sí, sin profundidad reflexiva e invisibilizando a la multiplicidad de voces que conforman lo humano. Así es como las especificidades interpretativas se diluyen en favor de marcos globales uniformadores, con peligrosa capacidad normativa debido a su omnipresencia.

Lo sabemos, todo texto producido por la IA termina exhibiendo el mismo tipo de "discurso" impersonal y neutro, propio de una valoración estadística de la palabra que debe completar cada secuencia de *tokens*.[2] Los partidarios de la IA podrán objetar nuevamente: "Eso también se puede programar". Y en efecto, por default, los LLM´s resultan halagadores y complacientes con el usuario. Como se ha observado muy pronto que esto puede resultar inconveniente, entonces se han desarrollado estrategias para "dotar" a los modelos de "personalidades" diferentes: reflexivas, no complacientes y hasta humorísticas o sarcásticas –como Grok o Monday–. Pero, nuevamente, sigue tratándose de simular personalidad. En todo caso, ¿por qué no mejor dejamos de invertir tanto esfuerzo y recursos en lograr algo que ya es gratuito y que lo puede hacer cualquier persona que se convierte en *autor*? No hay remedio, en caso de dejar toda producción de textualidad a la IA, el estilo propio de cada instaurador de discursividad necesariamente tenderá a desaparecer.

La homogeneización del discurso es problemática. Hasta ahora, las bases de datos con las cuales se entrenan los modelos de IA han sido producidas por seres humanos. Pero con la tasa de producción de texto por parte de los LLM's, eso puede cambiar en pocos años. El texto artificial generado ahí –con todo y sus "alucinaciones"– pronto pasará a nutrir las bases de datos consultadas por la propia IA. Esto puede generar un bucle nada favorable a la diversidad discursiva, a la originalidad y al rigor académico. Y, por supuesto, está el inconveniente de la rezagada legislación de derechos de autor para el uso de las bases de datos humanas. Tal como ha sucedido en otros campos de la tecnología aplicada, como la reproducción asistida o

[2] Esto, además, es epistémicamente peligroso: si los textos son producidos atendiendo a la frecuencia estadística a partir de las fuentes consultadas, entonces se privilegia la reproducción de las ideas ya existentes en demérito de aquellas que puedan resultar novedosas o desafiantes, pues estas últimas no serán relevantes en número para impactar a la predicción del siguiente *token*.

la clonación, este tema tecnológico ya rebasó la velocidad de reacción de las leyes nacionales. Está claro que este desafío sólo podrá ser abordado desde una perspectiva global.

La existencia de un debate sobre el reconocimiento de la IA como autor de los textos producidos nos advierte que tendemos a pensar al *auctor* en términos sólo de *scribens*. Mientras que tiene sentido pensar en un "*scribens* automático", no es así en el caso del *auctor*. Éste, como afirmó Foucault, es un generador de sentido discursivo; lo que escribe genera conversación, diálogo y debate. A su vez, la existencia de un autor detrás de la obra permite que el lector identifique, relacione y clasifique discursos descubriendo nuevos significados sobre el mundo y las personas: el discurso académico pretende decir verdades sobre el mundo. Lo que no reconocen los partidarios del bando de la IA-como-autor, es que las verdades se señalan en ambas direcciones: hacia el mundo, pero también hacia el productor del discurso.[3] Ambos polos epistémicos importan. Como bien lo sabe la tradición hermenéutica, y tal como se vislumbra en la obra tardía de Barthes, autor y lector son parte de un sistema dialéctico y complejo que no es posible separar sin perjuicio del sistema discursivo total. Si rasuramos la autoría quedándonos sólo con el *scribens* corremos el riesgo de matar al discurso favoreciendo sólo al texto. Y si matamos al discurso, otro tanto podrá pasar con la obra y con el lector, cuyas funciones, dicho sea de paso, también están en riesgo por el uso de los LLM´s.

¿Defender al autor?

Entonces, ¿hay que defender al autor tal como lo entendemos en la actualidad? Quizá no. Ya lo señalaron Barthes y Foucault: nuestro concepto de autor es relativamente reciente. Obedece a una lógica de propiedad derivada de un sistema que capitaliza la producción académica y artística. En ese sentido, en la actualidad no sólo importa identificar a la persona detrás del texto por cuestiones de responsabilidad, sino por razones de producción. En el campo académico mexicano la generación de discurso hoy se mide, se regula, se dictamina y se califica. El mérito tiene un número, dictado, entre

[3] En las intervenciones finales a la conferencia citada de Foucault, J. Ullmo afirma con toda razón: "Sólo hay autor cuando hay interioridad" (Foucault, 2010).

otros, por *Google Scholar*. Insumos como ése resultan índices valiosos para clasificar –con nombramientos de Candidato, niveles I, II o III y Emérito en el Sistema Nacional de Investigadoras e Investigadores– premiar, reconocer, "estimular" o desacreditar a la persona detrás del texto. El concepto contemporáneo de autor, en el marco de la investigación universitaria, está atado a los cuartiles, al *copyright* y al ORCID.

Es paradójico que un sistema de producción académica tan enfático en la identificación de los autores –se lucha con denuedo contra el plagio– y en la adjudicación autoral de las obras, priorice tan cuantitativamente la producción, las referencias, las citas y los índices, y termine provocando la reproducción de discursos miméticos, reciclados, muchas veces carentes de estilo propio y de novedad reflexiva. Dicho en breve: se firman como textos originales formatos plagados de ideas genéricas. El autor, como persona, como estilo y como mente generadora de ideas, parece desvanecerse detrás de ese exigente aparato académico-burocrático. La autoría, hay que admitirlo a la luz de esto, ya estaba en crisis antes del surgimiento de los LLM's. Y, sin embargo, seguimos creyendo que la idea de autor es valiosa y, por lo tanto, debe defenderse.

Lo aquí expresado forma parte de un número especial publicado por una revista académica de amplio prestigio internacional dentro del área de las humanidades y las ciencias sociales. En sus casi 22 años de existencia, la revista *Andamios* se ha posicionado como un referente importante dentro de diversas discusiones académicas en múltiples espacios iberoamericanos. Eso se debe, en parte, a dos razones: su accesibilidad y el riguroso esmero puesto en la calificación de los artículos a través de procesos "doble ciego". En efecto, hablamos de una revista alojada en una universidad pública y que patrocina la libre circulación de las ideas, por lo cual no cobra el acceso ni a los autores ni a los lectores. Esto parecerá poco, pero en un mundo editorial centrado en la rentabilidad económica del quehacer académico, la gratuidad y el rigor toman forma de resistencia política. *Andamios* es una de las iniciativas que, como *Scielo* o *Redalyc*, creen posible la reconstrucción del proceso editorial del mundo académico y, con ello, la resignificación del concepto de autor.

En resumen, estamos con Barthes cuando dice que importa el lector como punto sobre el cual se articula la unidad de sentido del texto; pero estamos también con Foucault cuando restituye la importancia del autor al afirmar

que la suya es una función compleja del discurso y que es, además, un instaurador de discursividad –cuando su pensamiento es original y fecundo–. Velar por la conservación del autor es garantizar la existencia del espacio académico como el lugar de encuentro entre racionalidades capaces de discutir la mismidad, pero también la diferencia y la novedad. Autor, obra y lector no son tres elementos absolutamente separados con funciones distintas; son tres elementos de un mismo proceso dialéctico que debemos reconfigurar para fortalecerlo. La dilución del autor o el lector por parte de las prácticas desleales en el uso de los LLM's tiene efectos globales en todo ese proceso.

¿Hay que limitar el uso de la IA en la producción académica? Sin duda. Pero la IA, queramos o no, es hija ilegítima pero reconocible de ese mismo sistema que desde tiempo atrás viene haciendo desaparecer al autor. Por tanto, mientras se limita y regula la producción con IA, se debe reconfigurar la forma como pensamos nuestro modelo de producción académica. De lo contrario, seguiremos fomentando el surgimiento y uso de herramientas que profundizan la crisis de ese modelo. No hay autoría en la IA, pero eso es justamente lo que buscan quienes pretenden defraudar al sistema académico "produciendo" decenas de artículos al año; en definitiva, no es un autor aquel que sólo encomienda la elaboración de texto a través de un *prompt*. Así las cosas, y parafraseando al ya citado Foucault, urge vigilar y legislar.

Fuentes consultadas

Ansede, M. (2023-06-02). Un científico que publica un estudio cada dos días muestra el lado más oscuro de la ciencia. En *El País*. Disponible en: https://elpais.com/ciencia/2023-06-03/un-cientifico-que-publica-un-estudio-cada-dos-dias-muestra-el-lado-mas-oscuro-de-la-ciencia.html

Barthes, R. (1987). La muerte del autor. En *El susurro del lenguaje*. pp. 65-82. Paidós.

Barthes, R. (2005). *La preparación de la novela: Notas de cursos y seminarios en el Collège de France, 1978-1979 y 1979-1980*. Siglo XXI.

Calvino, I. (2013). *Si una noche de invierno un viajero*. Siruela.

Cosmelli, D. y Thompson, E. (2010). Embodiment or Envatment?: Reflections on the Bodily Basis of Consciousness. En M. Siderits,

E. Thompson y D. Zahavi (Eds.). *Self, No Self? Perspectives from Metaphysics, Phenomenology, and Philosophy of Mind*. pp. 361-385. Oxford University Press. DOI: https://doi.org/10.7551/mitpress/9780262014601.003.0014

Down, A. (2025-12-06). Artificial Intelligence Research has a Slop Problem, Academics Say: 'It's a Mess'. En *The Guardian*. Disponible en: https://www.theguardian.com/technology/2025/dec/06/ai-research-papers

Dreyfus, H. (1992). *What Computers Still Can't do: a Critique of Artificial Reason*. MIT Press.

Foucault, M. (2010). *¿Qué es un autor?* Literales.

INDAUTOR. (2025-08-28). *La Suprema Corte de Justicia de la Nación resolvió en amparo directo que las obras creadas por inteligencia artificial no pueden registrarse como derechos de autor en México*. Disponible en: https://www.indautor.gob.mx/comunicados.php

Pérez, A. (2019). Qué es una autora o qué no es un autor. En A. Pérez y M. Torras (Eds.). *¿Qué es una autora? Encrucijadas entre género y autoría*. Icaria.

Searle, J. (1980). Minds, Brains, and Programs. En *Behavioral and Brain Sciences*. Vol. 3. Núm. 3. pp. 417-457. DOI: https://doi.org/10.1017/s0140525x00005756

Shmatko, N., Adam, A. y Esau, P. (2026). GPTZero Finds 100 New Hallucinations in NeurIPS 2025 Accepted Papers. En *AI Detection Resources | GPTZero*. Disponible en: https://gptzero.me/news/neurips/

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Harvard University Press.

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1344

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1346

La *pregunta auténtica* heideggeriana que la IA es incapaz de fagocitar

Concepción Delgado Parra*

Introducción

La inteligencia artificial (IA) está transformando la creación del conocimiento y, en consecuencia, el papel del investigador. No se trata únicamente de un cambio técnico y económico; representa, además, una crisis del modelo civilizatorio del conocimiento. En este nuevo escenario, los desafíos epistemológicos son profundos e indiscutibles. La naturaleza, el origen, los límites y la validez del conocimiento corren el riesgo de replegarse al uso de una infraestructura tecnológica de decisión reducida a acelerar el análisis de datos y redacción de textos en tiempo récord alterando el papel de quien investiga. El investigador deja de formarse como artesano del método y generador de conocimientos disciplinarios para convertirse en gestor de *prompts* y curador de *outputs*. Esto es resultado de que, en su estructura profunda, la IA generativa no crea conocimiento nuevo, aunque produzca estadísticamente combinaciones novedosas, solamente se mueve dentro del "mundo de lo ya interpretado". Procesa, predice, optimiza, pero no *escucha, no se angustia, no habita un mundo, no muere*. No tiene *Dasein*.[1] Por lo

[1] El *Dasein* heideggeriano remite a la existencia humana y significa, literalmente "ser-ahí". Designa el ente singular y concreto que cada uno de nosotros es, caracterizado no por una esencia fija, sino por su capacidad de comprender el Ser y proyectar sus propias posibilidades. Heidegger lo desglosa a través de tres dimensiones fundamentales. La primera, el *ser-en-el-mundo* (facticidad), el *Dasein* nunca existe de manera aislada ni como un sujeto incorpóreo frente a un objeto, siempre se encuentra situado en un entorno concreto. Su

* Profesora-investigadora de la Academia del Posgrado en Ciencias Sociales y Humanidades de la UACM, integrante del Comité editorial de *Andamios. Revista de Investigación Social* y miembro del Sistema Nacional de Investigadoras e Investigadores. ORCID: 0000-0001-7388-4052 Correo electrónico: concepcion.delgado@uacm.edu.mx

tanto, *no piensa* ni elabora *preguntas auténticas* en el sentido heideggeriano. De ninguna manera esto debe interpretarse como un rechazo al desarrollo tecnológico (neoludismo), ni suponer que peligramos ante la idea de que la IA piense mejor que nosotros. Se trata, más bien, de evitar que aceptemos que su modo de operar, particularmente, los procedimientos que utilizan la IA generativa y los grandes modelos del lenguaje, basados en el pensar "calculador", constituyen la única vía para la producción del conocimiento, donde el investigador deviene "operador" de datos.

Bajo la lógica de la IA, que genera una espiral al aplicar algoritmos sobre información que ya fue interpretada previamente, que establece correlaciones sutiles y abstrae nuevos datos para explorar escenarios hipotéticos y extraer múltiples capas de significado sobre la misma base de datos, surge la cuestión acerca de si existen otras formas de pensar que permitan una relación con el mundo existente más allá de un bucle de análisis recursivo. Por supuesto, la respuesta no puede ser lineal, para explorar este dilema proponemos recuperar dos modalidades del pensamiento presentes en la obra de Martin Heidegger. El *pensar calculador* orientado por la lógica de las causas y los efectos, así como por criterios de eficiencia, rendimiento y utilidad. Un pensamiento caracterizado por su capacidad para producir resultados y ofrecer una conclusión rápida y eficaz, que no se detiene a meditar, sino que busca soluciones inmediatas para controlar lo ente (*das Seinde)*. No *das*

relación primaria con las cosas no es meramente empírica, sino práctica, vinculada a la utilidad. La segunda dimensión es la *comprensión* y *el proyectarse*, donde la existencia no estática permanece instalada en un perpetuo proyecto continuo, abierta siempre a una disposición afectiva. Esto significa que el *Dasein* se encuentra en el mundo desde un determinado estado de ánimo –temor, alegría, angustia– que condiciona su modo de comprenderlo. La tercera dimensión corresponde a la *caída* y *la cotidianidad*, donde el *Dasein* tiende a ser absorbido por las tareas prácticas y a "caer" en la cotidianidad. En este estado, la existencia se vuelve inauténtica, pues habitamos el anonimato de lo "Uno" o "el se" (*Das Man*), donde hacemos lo que "se" hace, pensamos lo que "se" piensa y vivimos bajo las normas de la masa, perdiendo nuestra individualidad. Finalmente, el *Dasein* encuentra su sentido último en el cuidado y la temporalidad. Toda preocupación por las cosas, por los otros y por nosotros mismos expresa, en el fondo, el cuidado de la propia existencia. Este cuidado se fundamenta en el tiempo. El *Dasein* es esencialmente temporal, se proyecta hacia el futuro, hereda su pasado y vive en el presente. El sentido más propio y auténtico de la existencia se alcanza cuando el *Dasein* asume su propia muerte –su posibilidad más extrema e insustituible–, lo que permite dejar atrás la vida inauténtica y tomar las riendas de su propia vida (Lozano, 2004, p. 197-212).

seindo es un pensar en sentido propio, porque olvida el ser y solo maneja entes.[2] Y, por el otro, el *pensamiento meditativo*, aquel que se abre a aquello que "da qué pensar", a lo que es digno de ser pensado, que no persigue una utilidad inmediata; es una respuesta a una llamada. Pensar es "dejar ser", corresponder a la interpelación del ser, habitar el misterio. Requiere que nos comprometamos en algo que, a primera vista, pareciera no afectarnos (Heidegger, 2002, p. 19-31). Es un pensar que agradece (*denken-danken*), que permanece en la pregunta, "exige a veces un esfuerzo superior. Exige un largo entrenamiento. Requiere cuidados aún más delicados que cualquier otro oficio auténtico" (Heidegger, 2002, p. 19).

Y aquí es donde situamos la posibilidad de abordar la cuestión de la IA en otro registro, uno que permita re-situar al investigador como protagonista de la producción del conocimiento, no como un ente procesador reforzado, sino como el único que puede habitar la verdad del ser. Recuperar la crítica de Heidegger sobre el pensar frente a la IA implica, ante todo, entender que la esencia del pensar no es un proceso técnico ni un cálculo, más bien *es una escucha y una correspondencia que nos interpela*. Lo más grave de nuestra época –apuntará Heidegger– no es que sepamos cosas, sino que "lo más digno de pensarse se nos retira" (Heidegger, 2005, p. 132), y lo peor es que ni siquiera nos damos cuenta de que esto está sucediendo.

En este marco, restituir al investigador en el centro del conocimiento y reconocerlo como protagonista en la producción del saber constituye la preocupación fundamental de la reflexión que aquí se presenta. La *impo-*

[2] Vicente Lozano elabora una aproximación a la distinción entre el Ser (*Sein*) y el ente (das *Seiend*), desarrollado por Heidegger en *Ser y tiempo*, una interpretación que será útil para quienes no están familiarizados con la obra del filósofo alemán. "Ser es lo que determina al ente en cuanto ente, lo que hace que un ente sea, y ente significa cosa, algo que es. Por ello, todo ente presupone el Ser, que no puede ser pensado como un ente más, sino que es una especie de horizonte significativo desde el cual toda realidad, todo ente, es siempre visto. Ser es pues la idea no capturada conceptualmente que la metafísica supone sin tematizar. Es una idea indefinida de la que no tenemos ningún principio contenido que la precise, pero que está actuando desde el momento en que hay entes, esto es, cosas que no son. Se trata de un significado anterior a todas las distinciones que la tradición ha efectuado sobre los entes y que han dado lugar a las diferentes ciencias. Por ejemplo, la física estudia los entes o cosas que se mueven o alteran, la matemática estudia los entes o cosas numéricas, la historia estudia cómo se han comportado o qué le ha sucedido a un grupo de entes llamados seres humanos" (Lozano, 2005, p. 198).

sición –término traducido del alemán *Gestell*, entendido también como disposición o armazón– de la técnica sobre el ser humano, desplaza todo el tiempo al investigador del plano de la producción epistemológica. Del mismo modo que lo hace con la naturaleza a la que coloca como "fondo de reserva" explotable, medible y subordinada a la exigencia de la disponibilidad. Bajo esta lógica, el desarrollo tecnológico coloca al ser humano en un lugar secundario dentro de la existencia, mientras que la naturaleza queda despojada de su valor intrínseco (Heidegger, 1994, p. 22). Y es, precisamente en este lugar, donde la *pregunta auténtica*, entendida como el preguntar heideggeriano, abre un intersticio para re-pensar el papel del investigador en la producción del conocimiento en un mundo que niega la experiencia del asombro, de la angustia y la perplejidad, aspectos que comprometen la relación con el ser y que la IA es incapaz de fagocitar.

En el horizonte de la obra de Heidegger, la IA representa una de las formas más acabadas del pensar calculador, un modo de operar que *olvidó la pregunta por el ser*. Una especie de máquina que dejó de ocuparse de lo humano. Ciertamente, podemos encontrar fascinación frente a esta tecnología, pero es innegable que está vaciada de preocupación por lo humano. En este espacio, donde el cálculo fractura la voz que pregunta por el sentido de la existencia, se abre un intersticio para re-colocar al ser humano en el centro. Cuando expresamos nuestra preocupación por el desplazamiento del investigador en el proceso de la producción del conocimiento, el señalamiento es claro, se trata de resistir a una lógica instrumentalizada por la nueva tecnología digital que pretende remodelar el sentido del ser humano. El punto de partida de nuestra crítica es el peligro que la IA entraña cuando crea mecanismos dirigidos a organizar (y controlar) una sola forma de ver el mundo; un dispositivo que apunta a someter la autonomía del individuo. La IA no solo es una herramienta, constituye una forma de ver y organizar la vida, una lógica que se aplica a la existencia misma, a nuestras relaciones y a la naturaleza, donde todo queda reducido a la utilidad y la ganancia. Sin embargo, el problema no es que la IA nos domine, pues como mecanismo tecnológico carece de "deseo", es incapaz de preguntarse qué esperamos (o queremos) de este dispositivo; carece de inconsciente, no tiene fallas en el sentido de la muerte, de la finitud.

Lo que nos diferencia como seres humanos en relación con esta máquina, no es nuestra capacidad de conocer, sino nuestra capacidad de *no*

conocer, de preguntar. La IA solo responde, nunca pregunta. Por lo tanto, el verdadero riesgo que corremos frente a esta tecnología digital no radica en el control que pueda ejercer sobre los seres humanos –específicamente, sobre el investigador productor de conocimiento–, sino que, al relacionarnos con la IA, dejemos de hacernos preguntas. Aquí, la respuesta de Heidegger es aparentemente sencilla, pero, a la vez, profundamente radical: usar la técnica sin ser usados por ella. De manera ingeniosa y provocadora, dirá que "la técnica no es nada humano. Y sobre todo la esencia de la técnica no es nada técnico. La esencia de la técnica tiene su lugar en lo que ha de pensarse desde siempre antes de toda otra cosa" (Heidegger, 2005, p. 81). De allí que uno de los aspectos más difíciles de entender en el pensamiento heideggeriano sea la capacidad para dilucidar de qué manera la "pregunta auténtica" ofrece una alternativa a la paradoja moderna de la tecnología. No se trata de una respuesta técnica, más bien apunta a una forma activa de sabiduría dirigida a navegar el mundo moderno sin perder nuestra humanidad.

En este sentido, nuestro interés no está dirigido a encontrar un procedimiento para ajustar los parámetros técnicos de la IA y, de este modo, controlar la información que produce. Lo que guía este trabajo es volver a la *pregunta auténtica* heideggeriana para recuperar la capacidad de *interrogarnos por el sentido del conocer*, expresado en la relación del ser humano con la verdad, con el mundo y con la existencia, principio sobre el que se funda la producción epistemológica. Este ejercicio exige re-situar al investigador en el centro y reconocerlo como artífice de la generación del conocimiento, a partir de su *disposición* para interrogar el mundo. Con este propósito, a continuación, presentamos la ruta metodológica que seguirá el desarrollo de la reflexión. En la primera parte, exploraremos el sentido de la *pregunta auténtica* heideggeriana para distanciarnos de la comprensión ordinaria que la identifica con una simple solicitud de información. A continuación, discutiremos la distancia trazada por Heidegger entre la *pregunta auténtica* y la solicitud de información de modo que podamos discernir sobre lo que significa usar la tecnología sin ser usados por ella. Finalmente, buscaremos pensar si, en la época de la IA, aún es posible resguardar un modo de preguntar que no se agote en la eficacia de la técnica y revele un horizonte epistemológico en el que el investigador ocupe un lugar central por su disposición y apertura ética para poner en relación el pensamiento, el sentido y la verdad.

En síntesis, la búsqueda está orientada a explorar un territorio en el que la *pregunta auténtica* impida que el conocimiento quede atrapado entre un positivismo residual y un giro cualitativo de alcance superficial. Y, así, situarnos en la apertura originaria, donde la investigación y la producción de conocimiento reaparezcan como modos propiamente humanos de habitar la verdad del sentido.

La *pregunta auténtica* en el pensar heideggeriano

Una aproximación rigurosa a la pregunta auténtica heideggeriana requiere tomar distancia de la comprensión ordinaria que la concibe como una mera solicitud de información. En el pensamiento de Heidegger, la pregunta excede su función instrumental y adquiere el carácter de acontecimiento del Ser, una modalidad eminente del pensar en la que se pone en juego la relación del ser humano con la verdad, con el mundo y con su propia existencia.

Pero, cómo podemos identificar los rasgos que caracterizan a la pregunta auténtica. En *Ser y tiempo* (1997), Heidegger señala que la pregunta no se origina en un intelecto meramente curioso; más bien surge de una disposición afectiva ante los fenómenos existenciales que se manifiestan con toda su fuerza en lo cotidiano de nuestro mundo: "El hecho de que los sentimientos puedan trastocarse y enturbiarse solo dice que el 'ser ahí' es en cada caso ya siempre un estado de ánimo" (Heidegger, 1997, p. 151). Esta consideración abre la filosófica de los sentimientos, las emociones y los estados de ánimo a la interpelación analítica existencial de estos fenómenos en el espacio propio de su configuración en el que se producen: la cotidianidad (Cardona, 2003, p. 94). Alguien pregunta auténticamente porque está afectado, porque ha sido tocado por algo que lo desborda. La pregunta no expresa una simple falta de información técnica; acontece cuando las cosas mismas se vuelven, de pronto, enigmáticas. No se trata de una duda instrumental, sino del estremecimiento ante el milagro del ser: una forma de asombro que configura el temple originario de la filosofía. Surge también de la angustia, allí donde el mundo pierde su familiaridad y los entes se tornan insignificantes, obligando al *Dasein* a enfrentarse con su propia finitud, "Lo que angustia no es esto ni aquello [...] sino la posibilidad de algo 'a la mano' en general, es decir, el mundo mismo. [...] El angustiarse

abre original y directamente al mundo como mundo" (Heidegger, 1997, p. 207). De este temple nace la pregunta por el sentido del ser, no como curiosidad académica, sino como necesidad de orientación cuando toda certeza se desvanece. Según Heidegger, la época actual está marcada por la ausencia de la perplejidad. Creemos tener respuesta para todo, o confiamos en que la técnica podrá proporcionarla. El preguntar auténtico comienza, entonces, cuando experimentamos esa falta de preguntas como una privación, una especie de búsqueda a partir de una carencia y de una aspiración de saber (Heidegger, 1956). Sin embargo, la pregunta no surge como iniciativa voluntarista del sujeto; acontece como respuesta a una interpelación, a algo que nos da qué pensar. La pregunta auténtica es, por tanto, un padecer (*Leiden*), un dejar que aquello digno de ser pensado nos alcance.

La pregunta auténtica abre un espacio de manifestación donde comparece algo, no busca llenar un vacío dentro de un sistema de saber, sino que busca el ámbito donde el ser del ente puede mostrarse. A este intersticio Heidegger lo denomina "claro" (*Lichtung*), metáfora utilizada para explicar la verdad como des-ocultamiento, como el espacio abierto y libre de sombras que permite que los entes se muestren y sean inteligibles (Heidegger, 2003, p. 29-41; 2000), el cual debe entenderse como el lugar-abierto que posibilita la manifestación de todo lo que aparece presente (Garrido, 2016, p. 151-152). Poner en relación el "claro" con una pregunta auténtica implica formular una pregunta genuina a la que no se imponen categorías previas, precisamente porque en este ejercicio el pensar se expone a lo interrogado para permitir que lo "oculto" se manifieste en una forma de desocultamiento de lo que permanecía velado. Aquí no se trata de obtener un dato, sino de delimitar un horizonte desde donde lo interrogado pueda ser comprendido.

Por ejemplo, cuando preguntamos *qué es la justicia*, en el sentido auténtico, no se requiere de una definición de diccionario, más bien, la pregunta está dirigida a abrir el horizonte en el que se juega el sentido de la convivencia humana. Por ello, la pregunta ya es una forma de saber, aunque se trata de un saber consciente de su propio no-saber. Es el saber socrático llevado al plano ontológico; el reconocimiento de que no se sabe. Y, precisamente por eso, el pensamiento permanece abierto a que la cosa misma *dé qué pensar*.

Otra dimensión de la pregunta auténtica la constituye el hecho mismo de su imposibilidad de intercambiabilidad; no puede delegarse a una

máquina o a otro *Dasein*, precisamente porque concierne únicamente a quien pregunta. En *Ser y tiempo*, la pregunta por el sentido del ser no aparece como una cuestión externa al sujeto, pues compromete al propio *Dasein* que interroga (Heidegger, 1997, p. 14-24). Por ello, preguntar por el ser equivale a preguntar por el propio modo de existir. Así, la respuesta a una pregunta auténtica nunca deja intacto a quien la formula. Cuando preguntamos por la finitud, la respuesta no se reduce a un concepto, necesariamente nos arrojará a una forma de asumir el propio ser-para-la-muerte. Del mismo modo, quien pregunta por la verdad no encontrará la respuesta en una teoría, la interrogación planteada derivará en una transformación del ser humano con el mundo. Este carácter existencial descrito vuelve imposible que la IA pregunte en sentido propio. La máquina no muere, no se angustia, no habita un mundo y nada le va en aquello que formula. Sus "preguntas" son apenas simulacros estadísticos.

Ante la premura de la respuesta rápida, la pregunta auténtica desdice la prisa de quienes buscan la respuesta definitiva. En lugar de cancelarse en un "sí/no", abre un estar en camino, el marchar por un sendero o un camino de pensamiento: "En el bosque hay caminos [*Wege*], por lo general medio ocultos por la maleza, que cesan bruscamente en lo no hollado. Es a estos caminos que se llama '*Holzwege*' ["caminos de bosque, caminos que se pierden en el bosque"] (Heidegger, 2003, p. 9). Aquí, la respuesta auténtica devuelve la pregunta al preguntante con mayor intensidad. El preguntar entendido como la piedad del pensar nunca se agota ni se convierte en objeto poseído, protege el misterio de lo interrogado. Frente a este preguntar, la IA queda reducida a respuestas tendientes al cierre, en cambio, la pregunta auténtica, permanece abierta al ámbito de lo indisponible.

De este modo, desencadenar la pregunta auténtica implica exponerse a la interpelación de aquello digno de ser pensado para abrir un claro donde el ser del ente puede desocultarse. Lo radicalmente distinto en relación con la IA es que en ese movimiento queda comprometida la existencia misma del preguntante manteniéndolo en un camino inagotable de correspondencia con el misterio. La pregunta heideggeriana constituye apertura al mundo, compromiso existencial y fidelidad al sentido. En este camino, se convierte en el acto fundacional de un saber humano que ninguna máquina es capaz de simular ni falsificar.

La *pregunta auténtica* frente a la solicitud de información

Para Heidegger, la diferencia entre la mera solicitud de información y la pregunta auténtica no corresponde a una diferencia de grado; pertenece al orden de la esencia. Se trata de dos modos radicalmente distintos de habitar el lenguaje y relacionarse con la verdad (Heidegger, 2005; 1997; 1994). En este sentido, la IA solo puede responder a la mera solicitud de información y, al hacerlo, revela por contraste aquello que significa preguntar humanamente.

La pregunta entendida como solicitud de información, a la que la IA puede responder, interpela únicamente a la forma de preguntar propia del pensar calculador y de la existencia inauténtica. Heidegger la caracteriza de manera indirecta al analizar la curiosidad (*Neugier*) y la habladuría (*Gerede*).[3] Su origen se encuentra en una carencia óntica, no ontológica; falta un dato, un eslabón en una cadena causal o una instrucción. Esta modalidad presupone que quien pregunta ya comprende el horizonte de sentido de lo preguntado y requiere únicamente un contenido específico. Por ello, está orientada a cancelarse mediante una respuesta. Bajo la lógica de la determinación, una vez obtenida la información, la pregunta se apaga. Realiza un recorrido algorítmico en el que la pregunta de entrada conduce a una salida, la respuesta, para inmediatamente concluir el proceso. Detrás de la pregunta calculadora no hay nadie, precisamente porque quien cuestiona en ningún momento pone en juego el sentido de su existencia, solamente busca un dato. Por eso, la interrogación puede ser planteada por cualquier "uno" (*das Man*) y respondida igualmente por una máquina en la que la subjetividad no está comprometida en ningún sentido. Se pregunta por hechos, precios, resúmenes, predicciones, asumiendo que el mundo está configurado simplemente por un conjunto de objetos disponibles que pueden ser interrogados

[3] Heidegger aborda la curiosidad como la avidez por ver cosas nuevas constantemente, no se trata de un deseo genuino de conocer, más bien, de la búsqueda de novedades que lleva a cabo el *Dasein* solamente por el placer de saltar de una dispersión a otra, sin detenerse a comprenderlas. Y cuando se refiere a las "habladurías", no lo hace en un sentido despectivo, sino más bien como un fenómeno positivo que constituye la forma de ser del comprender e interpretar al "ser ahí" cotidiano, a la forma habitual en la que el *Dasein* se comunica con la cotidianidad. Aquí, no se trasmite una comprensión profunda de las cosas, sino que se basa en la repetición y en lo "que se dice" por el solo hecho de que los demás lo dicen, proceso que da lugar al entendimiento del mundo (Heidegger, 1997, p. 186-189).

o utilizados. Esta forma de preguntar no modifica a quien la plantea, lo deja esencialmente igual y, aunque puede conseguir más datos, su existencia permanece sin ser tocada. Por ello, la IA es la destinataria perfecta de la pregunta calculadora, porque opera en el mismo plano, procesa entes lingüísticos como objetos disponibles carentes de mundo, sin comprensión del ser y alejados de la posibilidad de que algo le *vaya* en su respuesta.

En el horizonte heideggeriano cambia radicalmente el registro de la pregunta auténtica, toda vez que constituye el acto fundacional del pensar meditativo, el cual debe entenderse a partir de la relación entre un "padecer" y un "corresponder". En las primeras páginas de su obra *¿Qué significa pensar?* (2005), Heidegger lanza un postulado abrupto, en el que declara que "lo que más merece pensarse en nuestro tiempo problemático es el hecho de que no pensamos" (Heidegger, 2005, p. 17). Inmediatamente señala que no se trata de una ofensa ni desvalorización del quehacer científico, lo que explica es que la ciencia no piensa porque se mueve dentro de un carril ya trazado. Para salir de este dilema propone un puente entre el padecer y el corresponder. Si en la primera lección de su texto advierte que entre la ciencia y el pensar no hay un puente, sino un "salto" (Heidegger, 2005, p. 19-23), en la segunda parte aclara que para dar este "salto" fuera del cálculo científico, el ser humano debe cambiar de postura. Abrir un espacio donde el pensar meditativo aparezca como esa relación receptiva y responsiva del padecer y corresponder (Heidegger, 2005, p. 75-109). Aquí, el origen de la pregunta auténtica, referida al padecer y corresponder, no se encuentra en la carencia de dato, pues más bien brota del asombro, de la angustia o de la perplejidad ente el hecho mismo de que las cosas sean. A esto nos remite Heidegger cuando señala que "Lo que más da qué pensar de suyo, lo más merecedor de pensarse [...] es el hecho de que nosotros todavía no pensamos" (Heidegger, 2005, p. 17-18). Quien interroga desde este registro, experimenta que *le falta suelo*, no un objeto determinado. Esta forma de preguntar impide la salida a una respuesta definitiva, antes bien, profundiza en el propio camino del pensar. Precisamente porque impide la clausura de la cuestión, al devolverla con mayor radicalidad a quien la formula. En esta nueva configuración, tanto la pregunta como la respuesta importan al ser que la realiza. Y, en ese tránsito, la experiencia no tiene término, al *Dasein* "le va su ser", preguntar no tiene término, es un camino

y no un problema que se resuelve; de ahí su carácter intransferible. Nadie podrá asumir la pregunta más que el propio *Dasein*, ya que, en ello, le va su propia existencia, la finitud y el estar-en-el-mundo. Asumir la pregunta desde esta postura remite a una índole irreductiblemente personal –esto no debe interpretarse en términos de subjetividad–, se trata de una pregunta auténtica porque abre un claro (*Lichtung*); no solicita información, permite que lo interrogado comparezca en su ser. Desde este marco, cuando se pregunta qué es la verdad, no se demanda una definición, de lo que se trata es de inaugurar el espacio en el que la esencia de la verdad puede manifestarse. Lo que advierte este cambio de registro es que la verdad no se reduce a la exactitud, el preguntar meditativo responde más bien a una llamada, a una escucha que corresponde. No se trata de la respuesta afirmativa que ofrece el sujeto moderno; la contestación surge de la interpelación de aquello que reclama ser pensado. Por ello, Heidegger vincula *pensar* con *agradecer*, esto sucede porque el pensar agradece la llegada de lo que da qué pensar (Heidegger, 2005, p. 139-141). Así, quien sostiene esta apertura siempre es tocado por lo que pregunta, nunca sale indemne, siempre es puesto en cuestión y descentrado. Lo que surge de una pregunta auténtica, del preguntar meditativo, no equivale a un contenido mental, precisamente porque se configura como un modo de ser. De esta manera, el ser humano se convierte en pastor del ser, no en un centro acumulador de información de datos.

Cuando contrastamos la pregunta auténtica con la mera solicitud de información desde el horizonte heideggeriano, advertimos que, a pesar de que la IA es capaz de simular una pregunta profunda, al final, solamente consigue hacer un eco estadístico desprovisto de existencia. La IA no se angustia, no muere, ni tampoco habita un mundo. Esto tiene una connotación fundamental en la reflexión que aquí presentamos, pues el acto humano de preguntar auténticamente constituye un punto de control indisponible para la máquina. Y, en ese sentido, mientras el ser humano conserve la capacidad de interrogarse por aquello que ninguna base de datos está en posibilidad de ofrecerle –la finitud, el ser, el sentido– permanecerá en el centro de la producción del saber, del conocer. Por eso, en la era de la IA, lo decisivo no es formular mejores preguntas técnicas; más bien, se trata de salvaguardar el preguntar que ninguna máquina será capaz de padecer y agradecer jamás. Con esta consideración avanzamos a la última parte de la reflexión.

El sentido auténtico de la pregunta como un camino para abrir un horizonte epistemológico en los tiempos de la IA

La distinción entre la pregunta auténtica y la solicitud de información permite situar al ser humano en el centro de la preocupación científica contemporánea ante el avance de la IA. Hoy, la producción de conocimiento tiende a confundirse con la generación de respuestas inmediatas a problemas complejos y pronto olvida que, el sentido de la pregunta auténtica solo puede surgir del ser humano. Simplemente obtener información perfectamente codificada sobre algo que ya ha sido interpretado, no compromete la existencia, la finitud, ni mucho menos compromete la relación del ser con la verdad. La pregunta auténtica conserva aún un núcleo irreductiblemente humano frente a los grandes sistemas capaces de procesar, ordenar y recombinar información a gran escala.

El dilema surge cuando en la producción de conocimiento en las ciencias sociales persiste el legado positivista y una especie de giro cualitativo que con frecuencia coloniza al pensamiento del preguntar auténtico. Ante este movimiento, el preguntar heideggeriano abre un horizonte epistemológico diferente y ofrece un suelo previo desde el cual investigar, donde el ser humano vuelve a ocupar el núcleo de la pregunta auténtica, no como el sujeto que domina al objeto; más bien, se trata del *Dasein* implicado en su existencia con aquello que interroga.

Históricamente, la ciencia social estándar suele estudiar sus temas de investigación como "objetos disponibles" susceptibles de ser interrogados, medidos y modelados. Bajo esa lógica examina cuestiones relacionadas con la desigualdad, la identidad, la opinión pública, la injusticia, la crisis de la democracia, el retorno de las ultraderechas, el racismo y mil etcéteras. En este enmarcamiento, la IA deviene aliada de esta tendencia, incluso, radicaliza esta forma de "producción del conocimiento", cada vez que procesa las problemáticas sociales en términos de variables textuales o numéricas desprovistas de mundo. Frente a esta "objetivación pragmática" de las ciencias sociales, el preguntar heideggeriano permite desplazar este punto de partida. Al volver a la pregunta "originaria" del pensamiento meditativo, lo social deja de concebirse como un ente objetivo para aparecer como un fenómeno que se manifiesta y oculta. Por ejemplo, cuando preguntamos por

la desigualdad desde este "suelo", no interrogamos sobre los dispositivos creados para distribuir la renta, sino por aquello que generó la desigualdad y que se manifiesta a través de múltiples prácticas cotidianas como si se tratara de una cuestión "natural". Nos referimos, también, al sentido del ser-con-otros que permanece velado bajo las métricas creadas por el pensar calculador. Aquí, la pregunta auténtica apunta a otro lugar, hacia aquello que se retira de la mera presentación estadística, orienta su *sentido al ser de lo social*. En esta trama, el investigador auténtico no se limita a identificar un problema o un área de conocimiento aún inexplorados. Más bien, lo que emerge aquí, es su padecer y desasosiego ante un aspecto del mundo social que devino enigmático. Su pregunta no responde a la necesidad de elaborar un artículo académico o a conseguir prestigio entre sus pares; lo que se juega el investigador es algo relacionado con su propia existencia. Esta actitud transforma el criterio que guía la producción del conocimiento, precisamente porque no solo se investiga lo publicable o el dato que ofrece una respuesta sistemática e inmediata, sino aquello que es *digno de ser pensado*.

La producción epistémica dominante interpretó la verdad en términos de la correspondencia entre enunciado y hecho, medida y datos. En este sentido, la IA se convirtió en una de las herramientas tecnológicas más acabadas para configurar esta "armonía", toda vez que predice la más probable correspondencia estadística entre las partes. Frente a este procedimiento, Heidegger recupera la verdad como *aletheia*, entendida como una especie de "desocultamiento" que, aplicado a las ciencias sociales, implica identificar una distinción fundamental en la producción del conocimiento: comprender no equivale a explicar. La explicación exige abrir un "claro" (*Lichtung*) (Heidegger, 2003, p. 29-41; 2000). Una investigación auténtica en ciencias sociales no "descubre" leyes, del mismo modo que la astronomía descubre planetas, sino que permite que un aspecto del mundo de la vida se muestre en su ser. Apostar por la entrevista en profundidad o la etnografía, asumidas desde el horizonte heideggeriano, implica no limitarse a recolectar datos, sino crear las condiciones para que el fenómeno se manifieste. El investigador no opera como un "extractor" de datos, sino como una especie de "partero" que asiste el acontecimiento de la existencia del ser. En esta experiencia, los conceptos dejan de ser meras categorías operacionales para adquirir un estatuto de existenciarios. Cuando recuperamos términos

heidegerianos como "cuidado", "angustia" o "ser-para-la-muerte", no se trata de etiquetar conductas observables. Más bien, deben entenderse como estructuras que nombran el existir humano que todo sujeto social habita, incluso, sin advertirlo. En este sentido, una sociología atravesada por el pensamiento meditativo heideggeriano no preguntaría cuántas personas padecen ansiedad, sino que interrogaría sobre aquello que revela la ansiedad contemporánea sobre un modo de habitar el tiempo y la incertidumbre de las sociedades tardomodernas. La pregunta estaría dirigida al sentido, no a la frecuencia. El preguntar auténtico se inscribe en la estructura del círculo hermenéutico, justamente porque solo es posible interrogar lo social desde una comprensión pre-teorética que, de antemano, ya opera. Esto es porque la investigación no rompe el círculo interpretativo, lo que hace es recorrerlo productivamente, una vez y cada vez, al hacer explícitos sus prejuicios para que el fenómeno pueda desmentir lo dicho o enriquecerlo.

No es posible entender la transformación más profunda del preguntar auténtico, que consiste en el desplazamiento del investigador neutral-objetivo hacia el *Dasein* comprometido. El investigador debe renunciar a la simple tarea de convertirse en un sujeto intercambiable que aplica un método. Por supuesto, la IA encarna el sueño de ese ideal, toda vez que su función consiste en operar como un procesador sin subjetividad. En cambio, el preguntar auténtico aparece en el horizonte como intransferible, porque quien pregunta siempre está *en juego*. Cuando el investigador interroga el ser de lo social, cuestiona también su propio ser-en-el-mundo. Estudiar la precariedad laboral o la polarización política deja de ser una operación externa para convertirse en una interpelación sobre el propio modo de habitar el lenguaje, la comunidad, la finitud. La producción del conocimiento resultante nunca se reduce a la realización de un informe, siempre estará atravesada por la transformación del propio investigador. Una interpretación adquiere validez cuando consigue *desocultar* aspectos del fenómeno que interpelan existencialmente a quien pregunta. El lector de una teoría social profunda no acumula simplemente información, también se reconoce, se inquieta, comienza a ver su mundo de otra manera; mediante este ejercicio acontece una forma de verdad. Un investigador auténtico apuesta por este proceso, mientras resiste a la "maquinación" metodológica. El preguntar auténtico se vuelve, así, un acto de resistencia académica que devuelve al investigador

la responsabilidad de pensar, en lugar de limitarlo a gestionar datos con herramientas que, en el fondo, imitan a la IA que él mismo critica.

A modo de cierre que da qué pensar...

Actualmente, el trabajo de investigación corre el riesgo de confundir la exigencia de conocimiento objetivo, el cálculo y la respuesta rápida con la suspensión del pensamiento meditativo, donde tiene lugar la pregunta auténtica. Cuando Heidegger declara que "el pensar obra en cuanto piensa", no significa que el pensar no actúe, sino que el pensar solo actúa en la medida en que piensa. Este actuar es, seguramente, el más simple, pero también el más elevado porque atañe a la relación del ser con el hombre (Heidegger, 2000, p. 1), no se refiere a la inactividad, sino a la capacidad que reclama dos tipos de obrar. Uno, que corresponde al obrar técnico orientado a producir efectos, dominar causas y alcanzar resultados medibles: "En la interpretación técnica del pensar se abandona el ser como elemento del pensar" (Heidegger, 2000, p. 1). Es el modo propio del capital y, también, de buena parte de una resistencia que mide su eficacia mediante indicadores. Este obrar encuentra respaldo en una ciencia estándar que suele tratar sus temas como objetos disponibles para ser interrogados, medidos y modelados. El segundo modo corresponde al obrar del pensar que abre un espacio de verdad y deja comparecer lo ente preparando una morada para el ser humano:

> El pensar es *l'engagement* mediante y para la verdad del ser. Su historia nunca es ya pasado, sino que está siempre por venir. La historia del ser sostiene y determina toda *condition e situation humaine*. Para que aprendamos a experimentar puramente la citada esencia del pensar, lo que equivale a llevarla a cabo, nos tenemos que liberar de la interpretación técnica del pensar. [...] El pensar no es práctico" (Heidegger, 2000, p. 1).

Sin este segundo obrar, la acción técnica pierde orientación, se vuelve ciega y termina por adquirir una deriva autodestructiva. Padecer la pregunta significa mantener abierta la herida del misterio en un mundo que busca clausurarla con datos, eslóganes y soluciones técnicas. Constituye la condición

para que la acción no degenere en activismo frenético, incapaz de soportar la angustia de la pregunta y dispuesto a hacer cualquier cosa, incluso lo más contraproducente, con tal de experimentar alguna forma de control.

En un sistema que transforma la disidencia en mercancía, convierte la identidad en nicho de mercado y orienta la producción del conocimiento hacia un modo de operar que pretende agotar el sentido mismo del pensar, padecer la pregunta por el sentido adquiere el carácter de una huelga ontológica. Implica retirar la capacidad de asombro y de interrogación radical del círculo de la producción de valor. Pero, esto no significa renunciar a la lucha; exige preguntar desde dónde se lucha. Cuando la acción nace de la voluntad de poder, termina alimentando la misma lógica que pretende combatir, aun cuando obtenga victorias parciales. Si se lucha desde el cuidado y la correspondencia con el sentido que nos desborda, la producción de conocimiento puede abrirse a una transformación más profunda. La ambición de poder del capitalismo no vuelve irrelevante la pregunta por el sentido; la pregunta por lo auténtico la convierte en la única pregunta incómoda que la IA es incapaz de fagocitar.

Fuentes consultadas

Cardona, L. (2003). La analítica de los estados de ánimo como hermenéutica de la cotidianidad. En *Universitas Philosofica*. Núm. 40-41. pp. 89-115.

Garrido, J. (2017). *Lichtung* en El final de la filosofía y la tarea del pensar: Qué nos tiene que decir el claro del ser. En *Tópicos. Revista de Filosofía*. Núm. 53. pp. 147-174. DOI: https://doi.org/10.21555/top.v0i53.810

Heidegger, M. (2005), *¿Qué significa pensar?* Trotta.

Heidegger, M. (2003). *Caminos del bosque*. Alianza.

Heidegger, M. (2002). *Serenidad*. Ediciones del Serbal.

Heidegger, M. (2000) *Carta sobre el humanismo*. Alianza. Disponible en: https://www.ucm.es/data/cont/docs/241-2015-06-16-Carta%20sobre%20el%20humanismo.pdf

Heidegger, M. (1997). *Ser y tiempo*. FCE.

Heidegger, M. (1994). La pregunta por la técnica. En I. Zimmermann (Ed.). *Conferencias y artículos*. pp. 9-38. Ediciones del Serbal.

Heidegger, M. (1956). *¿Qué es la filosofía?* Biblioteca digital Minerd-Dominicana Lee. Disponible en: https://ministeriodeeducacion.gob.do/docs/biblioteca-virtual/4tXN-heidegger-martin-que-es-la-filosofiapdf.pdf

Lozano, V. (2004). Heidegger y la cuestión del ser. En *Espíritu*, LIII. pp. 197-212.

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1346

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1347

¿Qué significa pensar en tiempos de IA? Una perspectiva fenomenológica

Edgar Sandoval*

Introducción

Pensar, en la vida cotidiana, significa conocer, almacenar, resolver y analizar una problemática de manera eficaz e individual a partir de la mente o el intelecto. La IA ayuda, en apariencia, a esta tarea, hace del pensar algo más fácil, rápido, cómodo, inmediato y más individual. Pensar también significa usar el lenguaje para transmitir significación y experiencia, por ello es una cuestión social, lingüística y cognitiva simultáneamente, en la que la deliberación de asuntos públicos exige razonamientos, pruebas, demostraciones y generalizaciones. La universidad sistematiza muchos de los aspectos que son parte del pensar, entre ellos los que conciernen a la autoridad de quien piensa a través del dominio de técnicas e instrumentos, así como de disciplina y de una vida dedicada al estudio. La figura del sabio, común en la universidad medieval, tiene como características la erudición y la seriedad, lo cual le otorga autoridad sobre el pensar. Con la llegada de la universidad moderna, la figura del sabio se seculariza, dando lugar a la figura del científico, éste tiene como características la exactitud y la precisión, con estos rasgos el científico se vuelve una figura aún más solemne y con mayor autoridad sobre el pensar.

A mediados del siglo XX surge la IA como herramienta para "pensar". La IA almacena una cantidad inmensa de información, amplía el conocimiento, realiza operaciones complejas, extiende, de alguna manera, la mente humana. Los rasgos del sabio y el científico, como la erudición, la exactitud y la preci-

* Profesor-investigador de la Academia del Posgrado en Humanidades y Ciencias Sociales de la UACM. Integrante del Comité editorial de *Andamios. Revista de Investigación Social* y miembro del Sistema Nacional de Investigadores. ORCID: 0009-0006-5674-6196 Correo electrónico: edgar.sandoval@uacm.edu.mx

sión, son refinados por la IA. La tecnología adquiere una mayor seriedad y autoridad respecto al sabio y el científico. Estos cambios obligan a formular la pregunta: ¿qué significa pensar en tiempos de IA? Lo que a continuación se presenta no es una respuesta a dicha pregunta, más bien se formulan más preguntas que pueden quizá ayudar a comprender lo complejo del tema en un momento en el que la tecnología se desarrolla de manera acelerada y en el cual se confía en el uso de la IA, con el presupuesto que podemos pensar con su ayuda. El artículo examina, desde una perspectiva fenomenológica, una concepción amplia del pensar, ya no como simple conocimiento, solución de problemas o bien como almacenamiento de información, sino como la capacidad de existir en el mundo de manera inteligente, afectiva y colectiva. El texto está dividido en dos secciones. En la primera sección, se presenta un esquema en el que se muestra cómo pensar en la academia ha estado en manos de clérigos, sabios, científicos y profesores, de cómo la autoridad que gozaban estas figuras al poseer sabiduría, erudición, precisión y exactitud se ha ido perdiendo; lo mismo se ha transformado la forma vertical con la que se transmite el pensar, ahora las formas son más horizontes. En la segunda sección, se analiza el pensar computacional, se cuestiona el modelo cartesiano, así como los distintos tipos de IA que tratan de simular la inteligencia humana como una réplica de la mente en términos de razonamientos abstractos y reglas lógicas; se plantea, de igual manera, la distinción entre cuerpo físico y cuerpo vívido, lo que permite sostener la idea de lo imposible que le resulta a las máquinas pensar en un sentido auténtico, pues las máquinas sólo poseen cuerpo físico y no cuerpo vivido, éste cuerpo es imprescindible para pensar, es decir, para afectar y verse afectado.

Pensar en la academia

Pensar, en un sentido optimista, es una de las actividades más comunes en el mundo académico. Este mundo, hasta mediados del siglo XX, se apoyaba en el modelo de autoridad que la Edad Media, el Renacimiento y la Modernidad otorgaron a las figuras del clérigo, del sabio, del científico, así como del profesor universitario. Las dos primeras figuras, el clérigo y el sabio, se caracterizan por poseer sabiduría y erudición –producto del dominio de diferentes técnicas, entre ellas la lógica, la gramática y la retórica, así como el

conocimientos en las artes y las ciencias y de una vida dedicada a la lectura, la contemplación y la reflexión por la naturaleza–. Las dos segundas figuras, el científico y el profesor, se caracterizan por lograr exactitud y precisión en sus descubrimientos, resultado de métodos y de acuerdos sobre el pensar. Estas últimas figuras nacen al secularizarse la universidad, con ellas se propone el pensar como resultado de la mente humana, por tanto, es una tarea exclusiva del ser humano. Dicha tarea se da de manera empírica, participa la observación, así como las inferencias sobre lo observado. La abstracción, esquematización, generalización y conceptualización son algunos de los aspectos que posee el pensar. El mundo académico se transforma con la imprenta, la cual permite a los legos tener, en algunas situaciones y condiciones, acceso a los libros y con ello pensar por cuenta propia. Por otra parte, los acontecimientos sociales propios del siglo XVIII, como los cafés, los clubes de lectura, la lectura de periódicos y de literatura por parte de sus autores, favorecieron el mundo de las letras. Sin embargo, el pensar no se democratiza, el dominio de la lectura culta queda en manos de las figuras mencionadas. Más bien, lo que existe es un impulso de la publicidad, del escrutinio, así como una crítica de las decisiones políticas y económicas que afectan a un grupo social de letrados.[1]

La aparición de las facultades y de las disciplinas al fomentar, entre otras cosas, el especialismo, conserva el modelo jerárquico del pensar. En esta nueva organización y administración del conocimiento, la figura del científico y del profesor mantiene la autoridad y erudición del sabio e integra precisión y exactitud, como aspectos propios del especialista. Estos aspectos acentúan el sentido de solemnidad y seriedad, enfatizan gestos parcos de quien sabe y piensa.[2] Los científicos se convierten, al igual que los intelec-

[1] A decir de Habermas, crítica y socialización, en los siglos XVIII y XIX, están íntimamente relacionados. La relación no sólo está en el circuito escritor-lector, sino también en las relaciones entre los lectores, quienes se reúnen en cafés y en clubes para ejercer el escrutinio, la deliberación y la crítica. La cultura de la imagen eliminó la socialización, el consumo de la imagen es en soledad. La imagen no tiene el propósito de deliberar, sino de atesorar. Las imágenes se convirtieron en museos populares, en reglas sociales. La democracia se llevó a cabo con imágenes (véase Habermas, 1994).

[2] Wallerstein identifica cómo en el medievo el saber está en manos de los clérigos: "El saber fue la explicación de la verdad por las autoridades clericales" (1997, p. 9). Esta autoridad, a pesar de que se intente romper con el adagio moderno Sapere aude, se fortalece con la secularización del mundo académico.

tuales, en figuras poco carismáticas y lejanas al pueblo. Se distinguen del payaso, del bufón, del hombre de la calle, que se caracterizan por la simpleza de vida, por la sonrisa en su rostro y por la ignorancia que los hace felices.

El especialismo, según Wallerstein (1997), provoca, además, la ignorancia de la cultura general. El especialista conoce mucho de poco y conoce poco de mucho. Los esfuerzos de muchas universidades del siglo XXI radican en poner sobrepesos al especialismo y fomentar primero la cultura general.[3] Sin embargo, estos sobrepesos son frágiles frente al uso de la IA. Con la IA, se conoce mucho de mucho, pero el conocimiento que se tiene es ininteligible para el que lo posee, por tanto, es incomprensible e inútil en un sentido genuino. El conocimiento de mucho de mucho crea falsos sabios y falsos eruditos, aparece un nuevo polímata, pero sin fundamento, dota de aparente autoridad a los legos, desprende soberbia en quienes, al aplicar la información encontrada con ayuda de la IA, carecen de conciencia de lo que están aplicando. La IA misma, al no comprender la información que posee, se vuelve, como lo ha señalado Emily Bender, un "loro estocástico".[4]

Pensar, en este esquema de la vida académica, es seguir la autoridad, sabiduría y erudición del sabio, así como la exactitud, el rigor y la precisión del científico. La consigna es aprender a pensar a través de la repetición de lo dicho por estas figuras en sus tratados y en sus cátedras. Rechazar la opinión, desvalorizar la información y actitudes que no contengan aspectos teóricos, meditativos, contemplativos y serios de la academia. El sabio, en el Renacimiento, es una figura central para pensar la naturaleza y el mundo, en él recae la autoridad, no sólo por poseer erudición, sino porque él da paso al científico. El científico en la modernidad es la figura secularizada, es decir, se vale de razonamientos para pensar. En estos tiempos, la IA es

[3] Dicha práctica se integró a diferentes modelos universitarios, entre ellos el de la Universidad Autónoma de la Ciudad de México. Esta universidad pondera tres propósitos en el siguiente orden: 1. Conocer la universidad, 2. Conocer la cultura general y 3. Conocer una disciplina.

[4] Carlos Fenollosa se opone a la idea de Emily Bender de ver a las IA como carentes de comprensión del sentido del lenguaje y de las palabras que pronuncian o que bien reciben. Para la lingüista, a la que él mismo refiere, el sentido en los humanos está en relación con la experiencia individual. Fenollosa piensa que las IA no son loros estocásticos, que quienes ven de esa manera a las IA no están distinguiendo la IA fuerte de la IA débil (Fenollosa, 2024, p. 35). Con todo y esta polémica, la idea de "loro" no es nueva; lo mismo vale, en algunos ámbitos, para la academia tradicional en la que los legos repiten lo que piensan clérigos, sabios y especialistas, que para la academia emergente, en la que los legos repiten lo que la IA piensa.

una figura que ayuda a pensar, pero sin la autoridad del sabio ni los razonamientos del científico, más bien es una figura autómata. La IA permite a cualquier lego traducir textos de otras lenguas; de igual manera, posibilita tener acceso a bancos de datos de revistas especializadas, consultar bibliotecas digitales, tener profesores y tutoriales que hablan con un idioma de uso común y accesible. Todavía más, la IA permite resumir lo traducido, sintetizar los libros, asimilar las ideas e incluso realizar trabajos académicos. En este sentido, quizá la IA se ha convertido en otra figura académica.

El trabajo académico enfrenta, en la actualidad, el uso desmedido, sin criterios, sin reglas claras, de la IA.[5] Desde el medioevo, pasando por la modernidad y hasta ahora en tiempos de IA, la academia conserva ciertas maneras de razonar, favorece el modelo vertical, alguien que sabe y que transmite su conocimiento a los que no saben, promueve, a través de los trabajos finales, de certificación y hasta la elaboración de tesis, el pensar de forma individual. De igual modo, se vigila en los salones de clase que no se copie, se someten los trabajos a programas que cotejan similitudes y permitan identificar plagios o saber si los trabajos fueron elaborados por IA.[6]

Los legos repiten el pensar del sabio, así como del científico, en el primero (el sabio) se confía en la autoridad de la que goza, en el segundo (el científico) se confía en los métodos que emplea. Algo similar sucede con el pensar de la IA, se confía en la información encontrada. La arquitectura

[5] Las máquinas en los siglos XVIII y XIX representaron para los trabajadores una amenaza, por ello hubo en muchas partes de Inglaterra y Francia movimientos que tuvieron como objetivo su destrucción. Quizá uno de los movimientos más conocidos es el *ludismo*. Este movimiento se opuso a los telares industriales, así como a las máquinas de hilar industrial. Hoy día en la academia existen luditas que se oponen al uso del pensamiento computacional.

[6] Uno de tantos efectos de la técnica es la evidente destrucción de la naturaleza, así como la destrucción de la humanidad. La destrucción de la humanidad y por consiguiente de la cultura no se da de manera explícita en las guerras. Las guerras son el efecto secundario, el momento en el que la humanidad se sitúa en la absoluta oscuridad. La oscuridad misma es la ausencia del pensar. El no pensar es en la reflexión fenomenológica sinónimo de estar dormidos, la técnica con su funcionamiento automático y mecánico, así como con su ritmo rápido y acelerado hace imposible estar despiertos. Las cosas se hacen por inercia, a su propio ritmo, sin que las personas estén despiertas y conscientes de ese hacer. El hacer ya no depende de las personas, sino más bien depende de las máquinas. El pensar en tiempos de la IA ya no equivale a estar despiertos, sino que es el resultado de operaciones digitales en las que se presenta la información como sinónimo de pensar. Montones de información que dicen cosas en común sobre un supuesto problema social.

con la que se diseña la vida académica también cambia. Las bibliotecas con espacios abiertos, oficinas y cubículos de profesores que, de igual manera, tienen sus colecciones de libros privados, dan lugar a bibliotecas digitales, a espacios improvisados para tomar clases a través de plataformas, se crea el libro de acceso libre, aparecen los salones virtuales, los chats, etc. En estas transformaciones, ¿basta con crear códigos éticos sobre el uso desmedido de la IA en las universidades? ¿Se puede vigilar la producción académica para saber si es original o es resultado del uso de la IA? Estas preguntas quizá advierten que la universidad ya no es más el recinto cerrado que fue por muchos siglos, al mismo tiempo obliga a hablar de cuestiones de ética en términos novedosos, así como de aspectos ontológicos en relación al pensar.

¿Qué significa pensar en tiempos de IA? Heidegger, en 1954, planteó una pregunta similar: ¿Qué significa pensar? El filósofo de la Selva Negra señaló, entre otras cosas, que pensar es un problema común, que exige de un marco, así como de un cuerpo y de un mundo de vida. Vio, de igual modo, cómo aparentemente la ciencia permite y propicia el pensar, pero lo que realmente hace es lo contrario. El pensar es una tarea común en un mundo también común, esto quiere decir, entre otras cosas, que se piensa por algo, que hay en muchos sentidos una exigencia por pensar. La exigencia es en Heidegger la del cuidado. Dreyfus, por su parte, observó, siguiendo a Heidegger, cómo las máquinas no pueden en sentido estricto pensar, porque para pensar, además de la exigencia del cuidado, se necesita de un cuerpo específico.

El pensamiento computacional

El pensamiento computacional, en términos de mente extendida, se impone cada vez más por encima del pensamiento humano, éste, como se ha visto en el apartado anterior, ha estado en el mundo académico en manos de unas cuantas figuras. Sin embargo, en perspectiva fenomenológica, pensar es más que poseer mente, más bien pensar equivale a tener conciencia intencional, a existir en el mundo con un cuerpo vivencial y a estar despiertos. La concepción fenomenológica de la mente resulta compleja. La mente no es una entidad matemática capaz de resolver problemas abstractos o bien de plantearlos. La IA conserva la idea de mente humana como programa informático. La mente, para la filosofía fenomenológica, es el cuerpo mismo en otro plano. El cuerpo, en este sentido, posibilita pensar. En fenomenología,

se distingue el cuerpo físico (*Körper*), cuerpo mecánico o técnico, del cuerpo vivido (*Leib*), cuerpo en relación afectiva y experiencial con otro cuerpo.

El cuerpo de la máquina es el que ejecuta una orden que la mente le indica. El cuerpo mecánico es capaz de moverse, pero no de expresarse. En el siglo XVIII se crearon cuerpos mecánicos con capacidad de movimiento y de habla. Causaban asombro. Muchos pensaron que debajo de esos cuerpos se escondían enanos que producían el movimiento o hablaban, pero no es así. Lo que produce el movimiento y el habla en esos cuerpos artificiales son mecanismos de relojería precisos, así como conocimientos sobre mecánica y gases. No en vano, en ese siglo estas áreas de la física se desarrollaron de manera extraordinaria. El cuerpo de las máquinas esconde en su interior engranajes que no son observados por quienes las manejan. Cuando las máquinas se descomponen, los usuarios son incapaces de arreglarlas porque ignoran su funcionamiento. El funcionamiento de la máquina es sencillo, la complejidad es un artilugio creado con el cuerpo que se ve y con el cuerpo que se oculta. El sentido de ocultamiento presenta a la máquina como cosa compleja. La máquina tiene un cuerpo análogo al del ser humano; pero, a diferencia del cuerpo humano, la máquina no posee un cuerpo que sea el resultado de la relación con otro cuerpo, es decir, sólo tiene un cuerpo metálico, plástico o sintético. En los humanos, además de existir cuerpo biológico, existe cuerpo vivido. El primer cuerpo presupone un sistema de órganos en relación funcional, el segundo cuerpo necesita de otros cuerpos, de la interacción social a través de la mirada, de los afectos, de las pasiones. En suma, de un entramado lingüístico y simbólico que lo dota de significados.

Los avances en tecnología han demostrado que el cuerpo humano es, en apariencia, imitable, que los robots pueden sentir también. Sin embargo, el cuerpo vivido no se puede imitar. Quizá esta imposibilidad de imitar el cuerpo vivido sea algo polémico. El cuerpo vivido desafía el modelo cartesiano que presupone, a su vez, el modelo de una máquina, de igual manera interroga el modelo de un lenguaje universal propuesto por Leibniz. Leibniz, en el siglo XVIII, crea una máquina capaz de realizar operaciones matemáticas, con ello emprende la empresa de crear un lenguaje universal capaz de calcular y quizá de predecir. Estos aspectos han recorrido, en estos pocos siglos, un tratamiento complejo. En el siglo XX, Dreyfus afirmó que las máquinas no piensan por sí mismas. La reacción a esta afirmación, por parte de Allen Newell y Herbert Simon, fue ver al filósofo norteamericano como alguien que ignora lo más básico en programación. Lo que está en juego no es el conocimiento

técnico de los filósofos, sino más bien es el examen e interpretación de la esencia de la técnica, así como valorar los efectos de ésta en lo social.

La concepción de la mente humana como programa informático concibe al cuerpo como un lenguaje. Por más que los programas informáticos hoy día operen con un código universal, el lenguaje humano no está hecho de lo universal. Cuando se trata de responder la pregunta *qué es pensar*, es necesario identificar que el pensar no es un problema de razonamiento, como se piensa en lógica y en matemáticas, sino, más bien, es una cuestión de marco, de sentido común y de tener mundo lingüístico. El razonamiento se puede replicar, pero no se puede replicar un afecto, tampoco se puede programar el marco que dota de sentido al lenguaje, no es posible reproducir el mundo en el que se existe, entre otras razones porque lo que hace aparecer estos aspectos es la experiencia del aquí y ahora. La distancia entre razonamiento y afección se pretende reducir con programas que son capaces de detectar tonos de voz, latidos del corazón, pulsaciones, etcétera –aspectos del orden de lo químico, de lo biológico y fisiológico–. Las afecciones escapan a estos órdenes. La afección es un lenguaje inimitable, es algo singular, tiene como efecto el obrar en un aquí y ahora. Por otro lado, está la colindancia de las afecciones con las emociones. Pareciera ser que las máquinas no pueden tener emociones porque su cuerpo sólo responde a programaciones y en ellas los razonamientos pesan más.

En esta visión informática el cuerpo es, como lo fue en el modelo cartesiano, un accesorio. En la reflexión fenomenológica, el cuerpo no se tiene como accesorio, es decir, no es algo que sea externo, que se pueda poseer o no. Más bien, el ser es el cuerpo mismo. La consigna es: *Yo soy mi cuerpo*, soy lo que me afecta y la posibilidad de afectar. La afección es, en sentido estricto, una faceta de la reciprocidad en la que si no hay cuerpo en relación con otro cuerpo no hay yo, es decir, no hay conciencia ni mente, ni se puede estar despierto. Lo que existe con el pensamiento racional, con un cuerpo carente de relación con otro cuerpo vivencial, es un ensimismamiento, un embotamiento, un estar dormidos. Este esquema es quizá más cercano a la naturaleza que a la técnica. El ser es por el cuerpo, por la experiencia que gracias al cuerpo se tiene.[7] El cuerpo no es equivalente, en esta perspectiva, al *hardware*. En

[7] Los animales son capaces de construir sus "casas" a través de su cuerpo, emplean su cuerpo como herramienta y medida. No hay facultades de ingeniería o de arquitectura a las que asisten las aves para pensar cómo se hace una casa o aprender complejos cálculos estructu-

la IA, la mente es equivalente a una máquina capaz de responder de manera autónoma a la solución de problemas, quizá el ejemplo más conocido es la máquina de Turing, en esa máquina el cuerpo es mero accesorio.

La ciencia ve al cuerpo humano también como un accesorio, es decir, como algo que se puede sustituir por otro cuerpo, que se puede mejorar. El cuerpo humano concebido como accesorio también es visto como cuerpo enfermo al que se tiene que sanar, cuerpo sucio al que se tiene que limpiar, cuerpo obeso al que se tiene que adelgazar, etc. Esta concepción es ajena al cuerpo real, se trata de una concepción ideal del cuerpo que la ciencia ha propuesto y que se ha convertido en *canon* y en *dictum*. Ese cuerpo ideal concebido por la ciencia no existe. El cuerpo que existe es un cuerpo que no puede clasificarse como enfermo o sano, ni como sucio o limpio, el cuerpo existente es en potencia y es contradictorio. Este cuerpo en potencia que se presenta en la fenomenología husserliana con la expresión "yo puedo", es ajeno al cuerpo de las máquinas. Las lavadoras, al igual que los automóviles, los aviones, las casas y cualquier cuerpo artificial pueden operar de forma autónoma con ayuda de la inteligencia artificial, pero lo que no pueden hacer es responder a la experiencia de quien las habita ni responder a sus propias experiencias, porque el cuerpo artificial no tiene experiencias propias. El cuerpo artificial tiene experiencias ajenas, identifica el uso que se da en las lavadoras, el cómo se maneja el automóvil, con lo cual crea memoria y con base en la memoria es capaz de crear patrones. Sin embargo, la memoria que las máquinas tienen cuando cuentan con IA es mental, en cambio la memoria de los humanos es memoria sensorial. El cuerpo físico, el vivencial y el artificial, significan el lenguaje, que cada uno posee, de forma distinta. El cuerpo físico, tanto el biológico como el cuerpo artificial, pueden replicar la respuesta, bien por la memoria mental o bien por las probabilidades

rales o diseñar sofisticados nidos. Nada de eso. Las aves emplean sus cuerpos para hacer las estructuras y los diseños de sus nidos; con base en sus movimientos, más que en sus cálculos, tienen la posibilidad de hacer nidos que son habitables. A diferencia de los humanos que, en las facultades donde aprenden a pensar dejan fuera su cuerpo, dan prioridad a la mente, aprenden en silencio, de forma individual, en un grado de atención y abstracción sofisticado las matemáticas que les servirán para construir casas. El resultado es que terminan construyendo miles de casas inhabitables que resultan colapsadas por temblores u otros aspectos de la naturaleza. Las aves construyen nidos habitables, en casi todo el sentido de lo que hoy conocemos por habitar. Esta diferencia de construir en los humanos a partir de la mente y construir en los animales con el cuerpo ha llamado la atención de arquitectos y filósofos.

estadísticas con las que están programados. El cuerpo vivencial, en cambio, es singular, está arrojado al mundo, por ello tiene experiencias de agobio, angustia, vértigo, las cuales no son replicables.

La concepción materialista-naturalista-mecanicista es confrontada con la concepción vivencial-experiencial. Esta última concepción es propia de la fenomenología. Merleau-Ponty, en *Fenomenología de la percepción*, estudia la percepción como parte integral del ser humano en el mundo lingüístico. Por tanto, el mundo tiene un significado común y compartido. Dreyfus también critica la visión mecanicista que los ingenieros adoptaron en el diseño de la IA. La diferencia entre cuerpo y mente es más que la distinción entre afecto y pensamiento. De otro modo, la computación afectiva propuesta por Rosalind Picard y Cynthia Breazeal solucionaría en parte el problema de la conciencia en unión con los afectos. Por su parte, Maturana en sus estudios sobre percepción descartó la visión materialista-naturalista-mecanicista que explicaba con base en el sistema ocular los mecanismos del percibir. La percepción para el biólogo chileno es resultado más de la intencionalidad, así como de la unión del cuerpo con la mente y el entorno, que de cualquier sistema biológico. En la unión que él estudia se da cuenta cómo la percepción no es ajena a lo percibido, ni quien percibe la cosa es ajeno a la cosa percibida. En cambio, para la idea de un cuerpo máquina existe diferencia entre quien percibe y lo percibido. La visión mecanicista divide, la visión fenomenológica une.

Pensar es una acción en la que participa de manera activa la mente en relación con el cuerpo. En la academia, tanto en la tradicional como en la emergente, pensar es aún una actividad mental, en dicha actividad se da por sentado que el cuerpo es un accesorio. Esta idea predominante, además de separar la mente del cuerpo, llega a equiparar el pensar con el conocer, en esta equiparación la discusión en materia de IA es acerca de qué tipo de conocimiento es el conocimiento computacional. El presupuesto es que la IA es capaz de conocer de forma autónoma. La mente humana se expande, ahora no sólo es una mente orgánica la que permite el conocimiento, sino también es la mente computacional la que hace posible el conocimiento. Con ayuda de la mente computacional se pueden realizar operaciones que antes no se podían llevar a cabo con la mente humana, también se pueden hacer cálculos mayores y más rápidos, o bien se pueden almacenar datos, ordenarlos, clasificarlos, etc., que antes era imposible de hacer. Conocer es sólo una parte del pensar, pero no el pensar mismo. El pensar exige, además de conocimiento,

una reflexión por el bien, la belleza, la realidad, el ser, es decir, el pensar es total, mientras que el conocer es parcial. Otra diferencia sustancial es que, en la filosofía, el conocimiento establece la relación entre un sujeto que conoce y un objeto que es conocido. La fenomenología supera esta relación al adoptar la pregunta por el pensar. Pensar es una actividad común en la que el significado de las cosas las hace útiles, en esa utilidad el mundo es uno.

Conocer y pensar, entonces, son actividades muy diferentes. La concepción tradicional de IA rescata postulados epistemológicos más que ontológicos, en ellos el conocimiento depende de un sujeto capaz de ordenar con la mente el mundo, esquematizarlo, taxonomizarlo. El conocimiento permite el control de los objetos del mundo como si el sujeto estuviera fuera del mundo. De lo que no son capaces las máquinas con este tipo de conocimiento es de estar en el mundo. Para estar en el mundo es necesario tener mundo y ser uno con él, para ello la experiencia y el lenguaje son fundamentales. Tener experiencia y lenguaje, por otro lado, equivale a pensar. Pensar es más una cuestión ontológica, es decir, modos inteligentes de existir, que calcular o sumar cantidades complejas. Cuando miramos el tema de la IA desde una perspectiva ontológica y no epistemológica, el mundo que resulta es un mundo en constante cambio y movimiento. En la perspectiva epistemológica las máquinas son capaces de trazar patrones y prever, por ejemplo, estrategias en un juego de ajedrez. En la perspectiva fenomenológica el mundo es abierto, en potencia y quizá contradictorio, además no se separa de las subjetividades, en esta perspectiva la IA funciona como la conciencia humana, a saber: todo es real.

Al equiparar el pensar con la relación inherente entre experiencia y lenguaje pareciera ser que se trata de un problema semántico, pero es más que eso. Presupone pensar el pensar en relación con el habla, por ello las discusiones en IA con perspectiva fenomenológica retoman los planteamientos no sólo de Merleau-Ponty sobre la percepción, sino también las ideas de Heidegger y de Gadamer, así como las de Sartre. Heidegger elabora una compleja concepción sobre lo que significa pensar y, al mismo tiempo, sobre lo que significa hablar, así como lo que es tener lenguaje. Por su parte, Gadamer señala en varios de sus textos, en particular en *Acerca del problema de la inteligencia*, que no hay palabras autónomas. Sartre reflexiona sobre qué significa mirar. Estos temas serán objeto de otro texto. Por ahora basta decir que la fenomenología es una perspectiva útil para comprender la IA en términos ontológicos y éticos, más que epistemológicos.

Conclusión

Lo que aquí se presentó es la hipótesis fenomenológica que se puede expresar en los siguientes términos: las máquinas poseen lenguaje, pero éste no nace de la experiencia con los otros, tampoco es el resultado de existir en el mundo, por tanto, el sentido y significado que pueda llegar a poseer a través del lenguaje computacional está determinado por la propia computadora y, a pesar de que la IA pueda tener una capacidad de tomar referencias y decisiones de significado sobre el lenguaje, esa capacidad no es del orden del existir en el mundo. Cuando se existe en el mundo, las palabras que se poseen a veces –o quizá la mayoría de las veces– carecen de un significado previamente establecido, adquieren un significado nuevo en una situación concreta y específica, dejan de tener ese significado una vez abandonada la situación. Heidegger elaboró este problema con la idea de "marco": para entender el significado del lenguaje, éste debe de poseer un marco. Este tema lo ha estudiado Masahiro Morioka. Los robots carecen de marco para actuar e interactuar, en sentido estricto no son autónomos. La inteligencia requiere, además de una capacidad de procesamiento infinito, de una autonomía y voluntad que escapan a la velocidad con la que las máquinas procesan la información. La inteligencia humana, para Morioka, es capaz de realizar juicios autónomos, juicios de importancia y elegir una acción por encima de otras. Estos tres aspectos no los pueden llevar a cabo las máquinas inteligentes. No obstante la imposibilidad de tareas por parte de la IA, así como las limitaciones que aún tiene, hacer uso de esa tecnología en el mundo académico es necesario e inevitable. La IA es una herramienta útil, como lo son la calculadora, las tablas de Excel, etc. Sin embargo, como herramienta puede privar de la experiencia social, exacerbar la comodidad, limitar el uso directo de los libros, entre otras cosas.

La imprenta, como se mencionó en este texto, democratiza el acceso a los libros, al mismo tiempo, la lectura culta queda en manos de la academia. Quizá, lo que ahora existe con el empleo desmedido de la IA, es una lectura inculta. También existe el riesgo de tomar como verdad lo que un loro estocástico dice. El riesgo no sólo es caer en el dogmatismo, sino en probabilidades estadísticas. Pero el riesgo mayor, en la perspectiva fenomenológica que aquí se abordó, es el de seguir dormidos, ser autómatas, enfatizar la

actitud natural que nos hace creer que conocemos y que pensamos. Por último, pensar es más que conocer algo, pensar es un modo particular de existir. La sentencia de Heidegger –"todavía no pensamos"– sigue teniendo actualidad. A pesar de la IA, todavía no pensamos.

Fuentes consultadas

Churchland, P. (1999). *Materia y conciencia. Introducción contemporánea a la filosofía de la mente*. Gedisa.

Fenollosa, C. (2024). *La singularidad*. Arpa.

Gadamer, H. G. (2001). *El estado oculto de la salud*. Gedisa.

Habermas, J. (1994). *Historia y crítica de la opinión pública. La transformación estructural de la vida pública*. Gustavo Gili.

Heidegger, M. (2006). *Introducción a la investigación fenomenológica*. Síntesis.

Heidegger, M. (2005). *¿Qué significa pensar?* Trotta.

Heidegger, M. (1997). *Filosofía, ciencia y técnica*. Universitaria.

Husserl, E. (2005). *Ideas relativas a una fenomenología pura y una filosofía fenomenológica. Libro segundo. Investigaciones fenomenológicas sobre la constitución*. FCE / UNAM.

Morioka, M. (2023). Artificial Intelligence and Contemporary Philosophy. Heidegger, Jonas, and Slime Mold. En *Journal of Philosophy of Life*. Vol. 13. Núm. 1. pp. 29-43.

Sigman, M. y Bilinkis, S. (2024). *Artificial. La nueva inteligencia y el contorno de lo humano*. Debate.

Wallerstein, I. (1997). *La historia de las ciencias sociales*. UNAM.

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1347

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1348

Democracia algorítmica e inteligencia artificial: una aproximación

Ángel Sermeño Quezada*

La así denominada inteligencia artificial generativa constituye la ramificación más notable y polémica en la evolución reciente de las nuevas tecnologías digitales. Supone, sin duda, una suerte de salto de calidad en la evolución vertiginosa de las nuevas tecnologías digitales en el sentido de que es un fenómeno dirigido hacia la revolución de un extremo a otro de nuestra existencia. Aún estamos lejos de calar el verdadero impacto que ya nos está marcando pero que nos impactará mucho más y de maneras complejas y diversas.

En este sentido, la presente exposición persigue un objetivo más bien modesto. Realizar una primera aproximación a la relación entre esta inteligencia artificial generativa y la democracia contemporánea; una democracia que, valga señalar, en las dos últimas décadas ha transitado de una democracia de audiencia (Manin, 1998) a una democracia algorítmica (Innerarity, 2025). Dicho de otra manera, ya estamos inmersos en un escenario en el cual los sistemas computacionales más avanzados son capaces de "actuar" de manera autónoma y automática en aquellos contextos hasta ahora reservados a la toma de decisiones estrictamente humanas. Por tanto, la promesa de la democracia que atribuye el ejercicio de la soberanía a una colectividad capaz de autogobernarse está cediendo dicho valor o principio fundamental a la esfera tecnológica que escapa al control de dicho poder soberano.

Nos encontramos ante el surgimiento de "nuevas" formas de gobernanza democrática, si es que aún podemos llamarla así, que nos obligan a repensar los fundamentos de nuestros principios y valores democráticos.

* Profesor investigador de la Academia de Ciencia Política y Administración Urbana de la UACM. E integrante del Grupo de Investigación de Teoría y Filosofía Política. Actualmente es miembro del Comité Editorial de *Andamios*. Miembro del Sistema Nacional de Investigadores de México. ORCID: 0000-0002-2441-8933 Correo electrónico: angel.alfredo.sermeno@uacm.edu.mx

Una primera aproximación a tal cuestión es lo que se trabaja en las líneas siguientes, pero primero unos antecedentes contextuales para, a continuación, desarrollar en dos momentos el amenazador impacto de la tecnología en la democracia. En el primero de ellos se proporciona una visión sintética de la relación de las plataformas digitales y las redes sociales sobre el comportamiento ciudadano. En el segundo momento se abordará la cuestión específica de la inteligencia artificial generativa y la democracia algorítmica.

Complejidad y crisis de la representación democrática

La promesa de autogobierno de una comunidad política nació en la modernidad temprana marcada por una clara desconfianza de las élites en las mayorías populares. Ello explica el carácter "experimental" que definió al diseño institucional de la representación política en su origen. Tal como nos narra Przeworski (2010), ello explica la curiosa combinación de una narrativa republicana con candados institucionales diseñados para limitar la influencia de las mayorías (*demofobia*) en la naciente democracia ideada por los padres fundadores. Sin embargo, la verdadera limitación de la participación ciudadana desde el principio estribó en la delegación de los ciudadanos en los representantes.

Como bien señala Dahl (1999), el tamaño de la comunidad política importa. Aunque la democracia representativa amplió la participación en extensión, asumió al mismo tiempo el límite de la reducida profundidad de dicha participación. Tal acomodo explica la centralidad que adquiere el mecanismo electoral. Las mayorías eligen, delegan y buscan un control democrático a través de la vigilancia y la influencia, pero no participan de manera directa en la toma de decisiones colectivas (Przeworski, 2019). Éstas quedan en manos de las élites representativas, que pueden o no tener un sesgo de clase social.

Por supuesto, no se está en contra de la intermediación que las instituciones representativas (elecciones, partidos políticos y parlamentos) suponen para la democracia contemporánea. El incremento de la población con derecho a participar en las condiciones de la modernidad fue un síntoma de la creciente complejidad de las sociedades de la época. Dicha complejidad no hizo más que incrementarse reforzando los límites a la participación política ciudadana directa a la que ya se ha hecho referencia. La intermediación

representativa fue el precio a pagar para enfrentar esa creciente complejidad (social, organizacional, administrativa, etcétera).

En este punto se justifica una breve digresión. Se entiende por complejidad aquella

> Cualidad de (algo) que está formado por múltiples elementos interrelacionados y cuyo comportamiento no puede comprenderse fácilmente observando cada parte por separado. Algo complejo no se reduce a la suma de sus componentes, sino que genera nuevas propiedades o comportamientos al interactuar. Los sistemas complejos son estructuras en las que una red de interacciones simples produce resultados imprevisibles o difíciles de explicar a simple vista. (https://www.significados.com/complejidad/)

Además de la imprevisibilidad, la complejidad también exige coordinación de muchos elementos que configuran un sistema dentro del cual cada componente (personas, flujos de información, etcétera) cumplen un papel específico y el éxito del sistema depende de la coordinación de sus elementos integrantes. Claramente, el sistema se impone en su racionalidad por sobre sus elementos, uno de los cuales es el humano. Este es el fundamento que explica cómo los sistemas ganan autonomía y capacidad de decisión por encima de las personas. Llevada al extremo esta tesis se puede proclamar la imposibilidad real de la participación ciudadana efectiva (Zolo, 1994). Las decisiones incluso escaparían a las élites (incluso a las tecnocráticas) y serían competencia impersonal del funcionamiento del sistema o subsistema en cuestión (económico, político, social, etcétera) (Luhman, 1998).

A finales-inicio de siglo había experiencias y argumentos empíricos para ser optimistas ante esta tesis tan abrumadoramente negativa respecto de las posibilidades de la participación política ciudadana, dado que llamaba ilusoria a la capacidad del autogobierno popular. En esa época, los movimientos sociales de protesta y las dinámicas deliberativas en el espacio público no estaban exentas de innovación y vigor en distintos contextos del orbe (Sermeño, 2006). Para el caso, la región de América Latina destacaba en innovación participativa en clave democratizadora (Dagnino, Olvera y Panfichi, 2006).

Con todo, y en honor a la verdad, las tendencias opuestas de apatía y desafección política ciudadana también mantenían en esa época una constancia creciente. Y, sobre todo, el descrédito de gran calado entre la ciudadanía respecto de los partidos políticos y los parlamentos era ya inocultable. Las inercias autoreferenciales adoptadas por ambas instituciones de intermediación, fundamentales para el funcionamiento y legitimidad de la representación política, se aislaron de la influencia y del control ciudadano con sus prácticas de cartelización (los partidos) y por el predominio de la lógica tecnocrática en las decisiones tomadas por los parlamentarios (Katz y Mair, 2004; Mair, 2015). Esta es la raíz del desencanto ciudadano con la política y los políticos y caldo de cultivo de los nuevos emergentes nuevos populismos demagógicos. Ya en ese momento apareció una pregunta que ahora resurge con renovada actualidad, a saber: ¿es incompatible la complejidad social con la soberanía popular?

Fue, en efecto, en este contexto de crisis de la representación que aparecieron con fuerza las nuevas tecnologías digitales impactando frontalmente a la democracia, primero con promesas de mayor democratización del espacio público pero que muy pronto resultaron falaces, y luego con la inocultable subordinación y dominio de las grandes corporaciones tecnológicas sobre la democracia, transformándola en algo en lo que ya no puede reconocerse a sí misma.

Democracia digital y transformación de la esfera pública

Son muchas, por supuesto, las cuestiones y derivaciones relativas a las relaciones entre digamos solamente internet, por no hablar de las nuevas tecnologías en su conjunto, que abarcan temas mucho más hondos y complejos. Por ejemplo, los referidos a cuestiones asociadas a vigilancia y control social, la propaganda mediante la utilización masiva de algoritmos, los muchos otros usos de la IA, etcétera, con la democracia, mismos que son, como señalo, demasiado amplios, sofisticados y difícilmente agotables en su totalidad.[1]

[1] La actual revolución tecnológica, como sabemos, se caracteriza por una evidente velocidad de innovación y difusión. Se trata de un cambio vertiginoso y exponencial que, además, muestra claramente al menos tres mega tendencias, a saber: "(I) físicas: vehículos autónomos,

Al respecto, lo primero que conviene realizar es explicitar una mínima delimitación de la relación entre el binomio cambio tecnológico y democracia. Si las preguntas se formulan de la manera siguiente, es decir, ampliamente en consonancia con el título de arriba: ¿en qué medida está siendo afectada la democracia por la actual revolución tecnológica?, ¿existe una crisis democrática asociada a dicha innovación?, ¿es posible sacar lecciones de lo que está ocurriendo en la anterior dinámica?, ¿incrementará la disrupción tecnológica los límites de la democracia?, ¿presenciaremos la muerte de la democracia como forma de organización política?

Estas cuestiones no nacen del ingenio de quien esto escribe. Flotan en el ambiente. Están formuladas y reformuladas a la enésima potencia. Aún no tienen respuestas totalmente concluyentes sino muy bien argumentadas, pero con fuertes dosis de especulación y abiertas al futuro. Sin embargo, en ese esfuerzo de cierta delimitación, ya se ha observado una clara variación en la narrativa que busca responder a las líneas generales de dicha conexión. Desde que se formuló explícitamente el problema, *grosso modo*, se ha pasado de una narrativa optimista a una fuertemente pesimista que todos conocen. Ciertamente, en el arranque de dicha narrativa se atribuyó al internet y a las redes sociales extraordinarias cualidades democratizantes. Las promesas fueron múltiples y se originaron en utópicas e ilusorias virtudes atribuidas a las innovaciones tecnológicas en las que se sustentaban las redes sociales y el internet: "nos habían anunciado la accesibilidad de la información, la eliminación de los secretos y la disolución de las estructuras de poder –recuerda para el caso Daniel Innerarity–, de tal modo que parecía inevitable avanzar en la democratización de la sociedad" (Innerarity, 2012). Por tanto, fue cuestión de tiempo que el péndulo oscilara hacia el lado pesimista de la relación entre impacto tecnológico y democracia. En el caso del internet y las redes sociales, éstas pronto mostraron su rostro oscuro y sombrío.

impresión 3D, robótica avanzada, nuevos materiales inteligentes que se auto reparan o se limpian a sí mismos; (II) digitales: internet de las cosas; y, (IV) biológicas: manipulación genética y avances en el campo de la medicina... (Por lo anterior, resulta evidente que debido al impacto de estos nuevos desarrollos) asistimos no solo a la creación de nuevos modelos de negocios y a la remodelación de los sistemas de producción, consumo y transporte, sino que, en el ámbito social, se está produciendo un cambio de paradigma en la manera en que nos comunicamos, expresamos, informamos, entretenemos, en incluso en cómo se gestionan nuestros gobiernos y las diversas áreas políticas" (Schwab citado por Arocena, Sansone y Álvarez, 2021).

La evidencia se abrió paso a partir del año 2015. En general, la dinámica reflejada en el espacio público virtual que las redes generan está dominada por múltiples experiencias de intolerancia, atomización, incapacidad de diálogo, radicalismo y emocionalidad que ciega la razón.

Al respecto, en un trabajo previo de mi autoría se enunciaron en particular tres impactos negativos: 1) La desinformación y manipulación que las redes sociales generan en el espacio público; destaca, para el caso, el emergente, extendido e incontenible fenómeno de la posverdad o verdades alternativas. 2) La adicción que inoculan en la ciudadanía en general induciendo prácticas que alimentan a la referida desinformación, refuerzan la despolitización y, por tanto, actúan como un veneno que mata el ideal de una ciudadanía activa reflexiva, responsable. 3) La vigilancia y el control que dichas redes (Facebook, Twitter ahora X, Google, WhatsApp, TikTok, etcétera), en tanto metamorfoseados mecanismos y/o dispositivos de poder, pueden ser y seguramente son puestos al servicio de estructuras autoritarias de dominación (Sermeño, 2021, p. 29-31).

Estos señalamientos no sólo continúan vigentes, sino que se han profundizado. En general, vale recalcar, expresan un amplio consenso entre los investigadores que se dedican a indagar en la temática. No está de más insistir en que el impacto de estas dinámicas de las redes sociales se retroalimenta ocasionando un mayor debilitamiento de los, ya de por sí, asediados cimientos de las instituciones democráticas. Sólo para ilustrar lo mucho que se anticipa al respecto de la cuestión que nos ocupa, se puede mencionar el ejemplo de una tesis sin duda polémica pero no carente de sentido. Me refiero a la conocida tesis del historiador Yuval Noah Harari (2018, p. 89) quien afirma que, en su forma actual, la democracia no sobrevivirá a la fusión de la biotecnología y la infotecnología.

Ahora bien, de esa tesis lo que quiero recuperar es la dimensión de la infotecnología asociada a la utilización de los algoritmos y el manejo de *big data*, para ser capaz de manipular la libre elección de los ciudadanos. Para la democracia se trata evidentemente de una cuestión crucial. Si los electores están sujetos a la manipulación (cosa que no es una preocupación nueva) la democracia evidentemente está en riesgo severo.

Existen voces intermedias entre las iniciales posturas optimistas y las dominantes visiones pesimistas. Estas perspectivas intermedias aceptan

que resultaron ser promesas ilusas las expectativas que anunciaron en su momento, de manera alegre y exagerada, que el auge del internet proporcionaría una mayor participación democrática, coadyuvaría a una deseable emancipación individual y fortalecerían la vida asociativa. También resultó ser increíblemente ingenuo pensar y defender que el ciberespacio sería un ámbito donde se disolvería el poder autoritario, dado que la tecnología vincularía libremente a los individuos y facilitaría el acceso al conocimiento de una manera inédita, casi universal. Sin embargo, conceden que el cambio digital ha posibilitado importantes avances democratizadores mientras que, por otra parte, señalan que los problemas que enfrenta la democracia no se deben tanto al impacto de las nuevas tecnologías sino a otros factores diversos, como, por ejemplo, el impacto perverso del neoliberalismo como fuente de deslegitimación del discurso democrático.

En estas reflexiones se tiende a coincidir en parte con estas voces intermedias intentando no caer en la ingenuidad o el utopismo de las posturas demasiado optimistas. Los efectos del internet y de las redes sobre la democracia son diversos en intensidad y en las dimensiones en las que se manifiestan y, sobre todo, son sin duda contradictorios. Es decir, las mismas variables, algunas clásicas asociadas a la participación política ciudadana, pueden mostrar desempeños diferentes en función específica del contexto social e histórico en el que el impacto de esas variables se evalúe. En concreto, desde la perspectiva optimista puede defenderse que internet y las redes potencian y amplían variables tales como empoderamiento ciudadano, visibilidad de grupos excluidos, otorgar y/o ampliar la voz, extender el espacio público, acotar el poder y, ya entusiasmados, facilitar la destrucción de regímenes autoritarios. Pero también es cierto que, como señalan los críticos que llamamos pesimistas, todo ello convive con frecuencia, y de la manera más cruenta, con las experiencias opuestas.

Señalemos el caso del poder. Ha sido ampliamente documentado cómo los regímenes autoritarios y sus liderazgos rápidamente aprendieron a utilizar las nuevas tecnologías y sus recursos, justamente y con gran eficacia, para controlar a sus ciudadanos y suprimir las disidencias existentes (Zuboff, 2020). Y, de hecho, no sólo eso, en tanto mecanismos de poder blando estos mismos regímenes autoritarios han logrado utilizar también las redes para ganar legitimidad y aceptación entre una audiencia vulnerable a las referidas metamorfosis experimentadas por los medios de comunicación.

Hay dos razones principales que explican lo anterior. Una es contextual y la otra estructural. La contextual advierte que cualquier innovación tecnológica se lleva a cabo en un contexto social y tiene efectos sociales que varían en función de las determinadas condiciones que rigen o determinan el contexto donde se despliegan:

> La información no fluye en el vacío –explica Innerarity– sino en un espacio político que está ya ocupado, organizado y estructurado en términos de poder. Internet no produce resultados iguales en países diversos. Internet pone en marcha unas dinámicas que aumentan las incertidumbres en torno al curso que van a tomar las sociedades. Los observadores occidentales han dado por supuesto que los dictadores no podrían poner internet a su servicio porque pensaban que la descentralización del poder promovida por internet era un fenómeno universal, una lógica sin excepciones, y no una lógica propia de nuestras democracias (Innerarity, 2012).

La razón estructural del por qué las redes globales no ocasionan mecánicamente una desconcentración del poder, ni desequilibran la autoridad de las élites, ni anulan asimetrías sociales establecidas es muy simple y radica en que su propia arquitectura es jerárquica. Internet posee y establece sus propios controles. Los así denominados *gatekeepers* (puestos fronterizos o puertas de paso) perfilan y definen el horizonte y el paisaje del ciberespacio. En efecto, los servidores que organizan el acceso y la circulación en el ciberespacio son pocos y hasta la fecha escapan de facto a las regulaciones estatales. En realidad, internet no elimina las relaciones de poder, sino que simplemente las transforma. Tan simple como reconocer que en internet también se ejerce la censura, la vigilancia y otras expresiones del control:

> Hay en la red una concentración de los proveedores de buscadores, que aparecen como simples mediadores o que aseguran limitarse a reflejar el tráfico existente, pero que también lo dirigen y condicionan. La red permite la proliferación de páginas y lugares, pero, de hecho, los buscadores centralizan la atención del público de manera que las interacciones están más limitadas de lo que solemos

> creer y el número de lugares que visitamos es más reducido de lo que suponemos. (Innerarity, 2012)

En suma, el péndulo de la relación entre tecnología y democracia claramente apunta hacia el predominio del pesimismo. En efecto, las formas para el ejercicio de la política se han modificado profundamente con la irrupción o más bien disrupción del internet y en general de las nuevas tecnologías.[2] Ya líneas arriba se hacía mención del impacto de la ineludible práctica de recolección masiva de datos y del uso de algoritmos. Ello perfila las referidas formas de hacer ahora política: "La política ha entrado en una fase algorítmica en la que buena parte de los votantes [son] seducidos por mensajes diseñados a su medida por sistemas tecnológicos que los conocen más que ellos mismos" (Arocena, Sansone y Álvarez, 2021). Y la utilización de esta tecnología conduce a la estructural tendencia de concentración del poder no sólo en los aparatos estatales, sean o no democráticos, sino también en los grandes corporativos tecnológicos que dominan dichas tecnologías y concentran y monopolizan el acceso a la información. Google, Facebook, Amazon y Apple son los llamados "cuatro grandes" corporativos que poseen un porcentaje escandaloso de la riqueza global y son sus sistemas digitales los que resguardan y controlan los datos personales de la mayoría de la población mundial.

Mientras estos corporativos privados concentran dicho poder y los Estados nación expresan su impotencia para regularlos, la dinámica real cotidiana en el espacio público cibernético ofrece un panorama de estado de naturaleza hobbesiano. En efecto, en palabras de Byung-Chul Han:

> Hoy las guerras de la información se libran con todos los medios técnicos y psicológicos imaginables. En Estados Unidos y Canadá, los votantes son llamados por robots e inundados con noticias falsas.

[2] "La digitalización del mundo en que vivimos avanza inexorable. Somete nuestra percepción, nuestra relación con el mundo y nuestra convivencia a un cambio radical. Nos sentimos aturdidos por el frenesí comunicativo e informativo. El tsunami de información desata fuerzas destructivas. Entretanto, se ha apoderado también de la esfera política y está provocando distorsiones y trastornos masivos en el proceso democrático. La democracia está degenerando en *infocracia*" (Han, 2022, p. 25).

> Ejércitos de troles intervienen en las campañas electorales difundiendo de forma deliberada noticias falsas y teorías conspirativas. Los bots, cuentas falsas automatizadas en las redes sociales, se hacen pasar por personas reales y publican, tuitean, 'linkean' y comparten. Difunden *fake news*, difamaciones y comentarios cargados de odio. Los ciudadanos son sustituidos por robots. Generan voces masivas con un coste marginal cero que infunden determinados sentimientos. Así es como distorsionan masivamente los debates políticos. También inflan de manera artificial el número de seguidores fingiendo de este modo un estado de opinión existente. Con sus tuits y comentarios pueden cambiar el clima de opinión en los medios sociales en la dirección deseada. Los estudios demuestran que basta con un pequeño porcentaje de bots para cambiar el clima de opinión. Puede que no influyan de manera directa en las decisiones de voto, pero manipulan los ámbitos de decisión. Los votantes están expuestos inconscientemente a sus influencias. Si los políticos se orientan por los sentimientos en la red, los bots sociales influyen de forma indirecta en las decisiones políticas. Cuando los ciudadanos interactúan con robots de opinión y se dejan manipular por ellos, cuando determinados actores, cuyos orígenes y motivaciones son completamente oscuros, interfieren en los debates políticos, la democracia está en peligro (Han, 2022, p. 39-40).

En conclusión, las iniciales promesas de un acceso ilimitado y universal ciudadano a la información facilitado por las nuevas tecnologías en el nuevo espacio público virtual han quedado desmentidas. En esta comunicación no se está en posición de afirmar o negar si la democracia en el mundo se encuentra o no en un trance de muerte, aunque es una pregunta que, sin duda, es legítimo formular.

Democracia algorítmica e inteligencia artificial generativa

¿Qué debemos entender por una democracia algorítmica? A grandes rasgos, estamos hablando de un sistema de toma de decisiones basado en el manejo e interpretación de macrodatos (*Big data*). Es, para decirlo en una palabra, la "dataficación" de la sociedad; o, para decirlo de otra forma,

en clave habermasiana, la "colonización tecnológica" del mundo de la vida (algoritmización). Como ya se ha señalado, consiste en la manera compleja en que los sistemas computaciones, es decir, las máquinas, con su inmensa capacidad de analizar cantidades ingentes de datos, bajo criterios de sostenibilidad, predictibilidad, velocidad, exhaustividad, extensibilidad, buscan soluciones y propone respuestas "eficientes" para la solución de muy diversos problemas sociales, políticos, económicos, etcétera.

Ahora bien, ¿cuál es la relación entre esta democracia algorítmica y la referida inteligencia artificial generativa (IA)? Expresado de manera simple, el algoritmo, definido de manera sencilla como ese conjunto de instrucciones lógicas, precisas y ordenadas para ejecutar una tarea o realizar un cálculo, es el componente básico de la IA. Por supuesto, esta última supone muchas otras capacidades avanzadas que son las que le dan su rasgo característico. En concreto, a lo que estas otras capacidades apuntan es hacia la posibilidad, ya en parte demostrada, de que las máquinas puedan llevar a cabo no sólo procedimientos mecánicos automatizados, sino que puedan encargarse de tareas con un contenido intelectual superior. Qué tan cierto sea esto último es algo que continúa a debate, pero parece claro que la inteligencia no es una inteligencia suprema o todopoderosa que le hará la guerra a los humanos como el Skynet de *Terminator*.

"¿Es verosimil que las máquinas puedan un día igualar e incluso superar a la inteligencia humana?", se pregunta un estudioso del fenómeno como lo es Daniel Innerarity, "¿Es posible un escenario en el que las máquinas inteligentes persigan objetivos que vayan contra nosotros sin que podamos detenerlas? Sobre la base de que los sistemas de inteligencia artificial son cada vez más inteligentes, se supone que llegará en un momento en el que alcancen y superen el nivel de la inteligencia humana" (Innerarity, 2025). La respuesta es clara, según este autor, en el sentido de que no es una forma correcta de plantear cuestión. La IA, sin duda, replica algunos aspectos de la inteligencia humana como el cálculo elaborado a una velocidad que no es humana, pero eso no es en sentido estricto inteligencia, dado que deja por fuera otras tareas de la inteligencia humana tan esenciales como la comprensión y la reflexión. "El poder computacional es cálculo veloz y procedimiento de una gran cantidad de datos, pero no es inteligencia. En la inteligencia artificial y

en el análisis de datos hay mucha fuerza bruta computacional. Pero no una comprensión del contexto mundano" (Innerarity, 2025).[3]

Sin embargo, aunque coinciden en reconocer que las visiones de ficciones hollywoodescas carecen de asidero en la realidad, las posturas más críticas sobre la IA advierten sobre los riesgos implícitos que conlleva. En estas interpretaciones críticas del fenómeno la IA representa, más bien, un cambio profundo del estatuto de las tecnologías digitales que consiste, de acuerdo con Éric Sadin, en:

> De ser prótesis acumulativas e intelectivas –porque permiten el almacenamiento, la indexación y el tráfico veloz de información–, han pasado a ser entidades de las que se espera que anuncien una verdad a partir de la interpretación automatizada de situaciones. Bajo un arco de formas que van de las más amables (sugerencias de compra, de movimientos) a las más coactivas (clasificación e indexación de los humanos para incluirlos/excluirlos de ciertos beneficios y derechos), en áreas relacionadas con la vida social general (la economía) o en otras que rozan la intimidad del individuo (la evaluación médica sobre su cuerpo), la inteligencia artificial propone diagnósticos que se suponen superiores a los humanos porque parten del manejo y correlación de datos imposible de realizar por un individuo. Así, gracias a este poder de enunciar una verdad que los humanos no podemos ver, la tecnología deja de ser protésica para devenir antropomórfica. La autonomía de decisiones de los sistemas de inteligencia artificial termina de cerrar el círculo de este cambio radical, enfrentándonos por primera vez con el estadio del tecno-logos de la técnica: no un

[3] No se puede eludir el hacer referencia en este debate a la cuestión de la relación entre inteligencia artificial e inteligencia humana. En efecto, como arriba se ha dicho, la IA es algo extremadamente potente, pero no tiene forma humana. No son, en consecuencia, inteligencias similares: "Las máquinas son mejores en el descubrimiento de patrones, matematización y razonamiento estadístico, análisis de datos masivos y manejos de casos rutinarios. Los humanos estableceríamos los objetivos y formularíamos juicios de valor, resolveríamos la información ambigua y discerniríamos. En casos difíciles, lo mejor que hacen los ordenadores es calcular y reconocer patrones. Lo que mejor hacen los humanos es comprender, contextualizar y decidir reflexivamente" (Innerarity, 2025). En esta visión, la relación entre ambas "inteligencias" no es de similaridad sino de complementariedad.

> logos sobre la técnica producido por los humanos sino humanos modelados por una técnica capaz de producir discurso o verdad. En este nuevo régimen sin réplica posible, se juega la vocación humana primordial de habitar el mundo. (Sadin, 2018)

Se contraponen al menos estas dos interpretaciones sobre la IA. Una, la que la ve como una evolución inevitable, para bien o para mal, de la presencia de la tecnología en la vida humana y que considera que los posibles efectos adversos de la tecnología pueden ser contenidos y manejados. En esta visión, quizás no del todo optimista, pero sin duda no pesimista, la apuesta es a sostener que la IA difícilmente llegará a sustituir la reflexión humana sobre los problemas de nuestro tiempo. Desde esta óptica se enuncian muchas ventajas que la IA proporciona (incremento de la productividad, por supuesto, garantizar eficiencia y eficacia en múltiples procesos, generación de tiempo libre para el ocio y la creatividad, etcétera).

En cambio, la interpretación pesimista advierte, como ya se ha dicho, de peligros que van más allá de los multicitados riesgos de vigilancia, control y dominación (autoritaria, totalitaria) o de la pérdida masiva de empleos, para señalar algo más hondo y estructural, pues sostiene que la IA constituye una amenaza a la creatividad, el talento y el intelecto de las personas, es decir, a las bases mismas de la civilización y la condición humana.[4] Esta tensión se encuentra abierta. Sin duda, será el futuro quien defina la dimensión real que adquiera dicha contraposición de interpretaciones. En este trabajo sólo apuntaremos al impacto de la IA sobre la democracia.

[4] Al respecto, cabe recuperar una nota coyuntural pero que no se puede ignorar al calor de lo que en este trabajo se ha discutido. El papa León XIV en su primera encíclica se pronuncia sobre el tema y califica a la trayectoria actual de la IA como una profunda amenaza para la dignidad humana (León XIV, 2026). El papa pone luz sobre una cuestión decisiva y de todos conocida. El hecho de que la tecnología nunca es neutral y que, también en el caso de la IA, dicha tecnología adopta las características y los fines de quienes la diseñan, financian, utilizan y se oponen a su regulación. Comenta al respecto el Premio Nobel de economía Daron Acemoglu: "La encíclica es enfática en señalar que el progreso tecnológico no es necesariamente un progreso moral. Que algo sea técnicamente factible no significa que sea beneficioso para la humanidad. Que una tecnología sea deseable depende de quién la controla y de la ideología e intereses que la guían" (Acemoglu, 2026-06-04).

Volviendo así a la cuestión política, la conclusión preliminar que hasta ahora se sostiene es que la democracia algorítmica es la negación de la democracia tal y como la hemos entendido y teorizado a lo largo de los últimos doscientos cincuenta años. Lo es porque sencillamente destierra la capacidad de agencia atribuida al ciudadano en la construcción del consenso social. Nociones como delegación del poder de decisión y dependencia tecnológica constituyen, en efecto, el núcleo de la democracia algorítmica. En palabras de Jurgen Habermas:

> Este poder tecnológico otorga una capacidad casi ilimitada para transformar el mundo social e influir en la construcción de sentido –verdad y justicia–, y amenaza con desterrar a través de imperativos sistémicos como el poder y el dinero, los procesos de comunicación que los seres dotados de habla y acción ponen en marcha para llegar a acuerdos intersubjetivos. (Habermas, 2025)

Esta primera aproximación a la relación entre democracia e IA no parece dejar mucho espacio para el optimismo. El impacto de las nuevas tecnologías, como ya se ha descrito, junto a la acumulación de los fallos y déficit previos de la democracia liberal (representativa), han permitido emerger alternativas autoritarias a la democracia a escala global en el mundo de hoy. Distintas versiones de autoritarismos bajo fachadas democráticas, como los populismos de nuevo cuño, han dado paso a democracias nominales, pero no reales. Son democracias, como nos advierte Habermas en la cita previa, sin los derechos básicos para garantizar la igualdad y la participación, esto es, la autonomía de los ciudadanos. Con todo, ese horizonte no muy esperanzador continúa abierto.

Fuentes consultadas

Acemoglu, D. (2026-06-04). El Papa debería haber ido más allá en lo respecta a la IA. En *El País*. Disponible en: https://elpais.com/economia/negocios/2026-06-05/el-papa-deberia-haber-ido-mas-alla-en-lo-que-respecta-a-la-ia.html

Arocena, F., Sansone, S. y Álvarez, N. (2021). Disrupción tecnológica y democracia en el siglo XXI. En *Cuestiones de Sociología*. Núm. 25. DOI: https://doi.org/10.24215/23468904e125

Dagnino, E., Olvera, A. y Panfichi, A. (Coords.). (2006). *La disputa por la construcción democrática en América Latina*. FCE, CIESAS.

Dahl, R. (1999). *La democracia. Una guía para los ciudadanos*. Taurus.

Habermas, J. (2025). *Un nuevo cambio estructural de la esfera pública y la política deliberativa*. Trotta.

Han, B. (2022). *La digitalización y la crisis de la democracia*. Taurus.

Harari, Y. (2018). *21 lecciones para el siglo XXI*. Debate.

Innerarity, D. (2025). *Una teoría crítica de la inteligencia artificial*, Galaxia Gutenberg.

Innerarity, D. (2012). Desenredar una ilusión: notas para una teoría crítica de la democracia digital. En D. Innerarity y S. Champeau (Comps.). *Internet y el futuro de la democracia*. Paidós.

Katz, R. y Mair, P. (2004). El partido cartel. La transformación de los modelos de partidos y la democracia de partidos. En *Zona Abierta*. Núm. 108-109. pp. 9-42.

León XIV. (2026). Carta encíclica: *Magnifica Humanitas. Sobre la custodia de la persona humana en el tiempo de la inteligencia artificial*. Disponible en: https://www.vatican.va/content/leo-xiv/es/encyclicals/documents/20260515-magnifica-humanitas.html

Luhman, N. (1998). *Complejidad y modernidad. De la unidad a la diferencia*. Trotta.

Mair, P. (2015). *Gobernando el vacío. La banalización de la democracia occidental*. Alianza.

Manin, B. (1998). *Los principios del gobierno representativo*. Alianza.

Przeworski, A. (2019). *¿Por qué tomarse la molestia de hacer elecciones? Pequeño manual para entender el funcionamiento de la democracia*. Siglo XXI.

Przeworski, A. (2010). *Qué esperar de la democracia. Límites y posibilidades del autogobierno*. Siglo XXI.

Sadin, E. (2018). *La inteligencia artificial o el desafío del siglo. Anatomía de un antihumanismo radical*. Caja Negra.

Sermeño, A. (2021). Democracia representativa, populismo y redes sociales. En *Folios*. Núm. 37. pp. 22-31.

Sermeño, A. (2006). Democracia y participación política. Los retos del presente. En *Andamios. Revista de Investigación Social*. Núm. 4. pp. 7-34.

Zolo, D. (1994). *Democracia y complejidad. Un enfoque realista*. Nueva Visión.

Zuboff, S. (2020). *La era del capitalismo de la vigilancia. La lucha por un futuro humano frente a las nuevas fronteras del poder*. Paidós.

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1348

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1349

Entre sintaxis y normatividad: una aproximación lógica y computacional a los problemas éticos de la inteligencia artificial

Jesús Jasso Méndez*

There is a fallacy often committed in discussion of Artificial Intelligence, especially AI of superhuman capability. Someone says: 'When technology advances far enough, we'll be able to build minds far surpassing human intelligence. Now, it's obvious that how large a cheesecake you can make depends on your intelligence. A superintelligence could build enormous cheesecakes - cheesecakes the size of cities - by golly, the future will be full of giant cheesecakes! The question is whether the superintelligence wants to nbuild giant cheesecakes. The vision leaps directly from capability to actuality,without considering the necessary intermediate of motive.
Eliezar Yudkowsky

Introducción

La relación entre ética e inteligencia artificial (IA) suele analizarse desde distintos campos y enfoques *ex. gr.* filosóficos, sociales, políticos, jurídicos, económicos, tecnológicos. Sin embargo, esta relación puede examinarse

* Profesor-investigador de la Academia de Filosofía e Historia de las Ideas de la Universidad Autónoma de la Ciudad de México, área: Lógica y Epistemología. Profesor en la Universidad Nacional Autónoma de México, área: Lógica y Filosofía de la Mente. Presidente de la Academia Mexicana de Lógica, A. C. (2019-2021). Miembro del Sistema Nacional de Investigadoras e Investigadores (Nivel II). Integrante del Programa de investigación 056 del Colegio de Humanidades y Ciencias Sociales, UACM. Actualmente es miembro del Comité Editorial de *Andamios*. ORCID: 0000-0002-3091-5258 Correos electrónicos: jesus.jasso@uacm.edu.mx; jesusjasso@filos.unam.mx

desde un plano más básico, el que conjuga la lógica y la teoría de la computabilidad. En este nivel, las cuestiones éticas no se plantean exclusivamente en términos de señalar valores o formas correctas de justificar la resolución de problemas teórico-prácticos por parte de un agente artificial, sino a partir de las restricciones inherentes a lo que puede ser formalizado, medido cuantificacionalmente o computado.

¿Es lógicamente posible formalizar de manera completa principios éticos a partir de algún lenguaje lógico altamente expresivo susceptible de implementación computacional?

La teoría de la computabilidad ofrece una respuesta potente pero limitada. Por una parte, algunos sistemas normativos pueden modelarse mediante lenguajes lógicos y ejecutarse algorítmicamente; por otra, los resultados estándar en torno a la indecidibilidad[1] demuestran, para ciertos casos, la inexistencia de un procedimiento algorítmico capaz de incluir al conjunto completo de instancias. En este último escenario se encuentra la normatividad ética.[2] La ética computacional subyacente al desarrollo de las IAs no se limita a las dificultades propias al diseño de sistemas artificiales, sino a cuestiones profundas de limitación lingüística, de limitación teórica.

De acuerdo con las posibilidades del modelaje lógico y computacional, existe un problema general absolutamente no menor relacionado con la potente noción de algoritmo. Si por algoritmo se entiende un procedimiento formal –tipo máquina de Turing: *effective method* (Turing, 1936-1937, p. 230-235)–, su funcionamiento implica la dimensión normativa de seguir pasos ordenados para resolver un problema, para realizar una tarea. En otras palabras, las decisiones automatizadas dependen por definición de un conjunto finito y explícito de reglas. ¿Todo se puede formalizar y computar? La respuesta es no. Existe un conjunto finito de información científica y

[1] En términos generales, la indecibilidad en lógica y teoría de la computabilidad se refiere a casos científicos y matemáticos para los cuales no puede existir algún algoritmo general que siempre dé una salida informacional del tipo sí o no, a partir de un número finito de pasos. Ver por ejemplo, Turing (1936-1937, p. 259-265).

[2] La normatividad ética refiere a las relaciones necesarias entre principios, reglas, criterios, valores, y el diseño, desarrollo y uso de los sistemas inteligentes. En términos generales, aceptar una norma nos compromete con ciertas inferencias y rechaza otras. La dimensión inferencial de la normatividad tiene como finalidad entonces regular acciones, razonamientos y justificaciones en el orden humano y artificial.

humanística susceptible a la estructuración lógica y un conjunto de contenidos científicos y humanísticos que, por su naturaleza deliberativa,[3] no forman parte de la cardinalidad lingüística cuantificada *i.e.* tales contenidos son altamente difíciles de codificar al interior de un procedimiento algorítmico. Paradigmáticamente, valoraciones éticas como *bueno, malo, justo, correcto* o *incorrecto* por sus propiedades conceptuales y aplicativas (*ex. gr.* contexto, pluralidad de valores, ambigüedad lingüística, intencionalidad, desacuerdo) no se capturan –al menos no toda su trama y urdimbre– en funciones calculables a lo largo de una estructura computacional que garantice su completa formalización y decibilidad.[4]

Este último aspecto entre dominios formalizables-computables nos conduce naturalmente a cuestiones relacionadas con la *atribución de eticidad a Ias Inteligencias Artificiales*, *i.e.*, a preguntarnos sobre el carácter ético de modelos y algoritmos diseñados no sólo para procesar información y desplegar resultados inferencialmente, sino para simular capacidades similares a las humanas en la toma de decisiones intencionadas, *ex. gr.*, razonar, resolver problemas, comunicar, aprender, justificar resultados. En esta línea, la atribución de moralidad y corrección ética está en juego cuando los sistemas artificiales median las decisiones de los razonadores humanos, resultados que impactan las esferas públicas y privadas de los individuos. Sobre esta posibilidad, consumada actualmente, el compromiso analítico de los especialistas –en lógicas y teorías computacionales, en principio– alrededor del diseño y funcionamiento de inteligencias artificiales, debe dirigirse al diseño de sistemas capaces de salvaguardar, por una parte, la corrección de acciones

[3] La expresión 'naturaleza deliberativa' refiere a la información científica y humanística proveniente de aspectos cognitivos y metodológicos no definitivos *ex. gr.* marcos conceptuales abiertos a constante corrección, tensión entre perspectivas sobre cuestiones compartidas, narrativas científicas basadas en la percepción de los especialistas, acuerdos y desacuerdos modificables entre expertos, contenidos aceptables aunque altamente falibles por su alta complejidad, y valores científicos y humanísticos no uniformes.

[4] La expresión "completa formalización y decibilidad" en términos generales refiere, por una parte, a un dominio *ex. gr.* de enunciados e incluso de reglas expresados íntegramente dentro de un sistema lógico-matemático; la decibilidad, por su parte, consiste en un procedimiento algorítmico capaz de resolver en cualquier caso correctamente a una pregunta (entrada- proceso y ejecución finita-respuesta, en términos de sí o no para cualquier entrada). Respecto a la noción de formalización se puede consultar *ex. gr.* Hulth y Ryan (2004); respecto a la noción de decibilidad se puede consultar *ex. gr.* Sipser (2013, p. 193-210).

y prácticas sociales como resultado de entradas informacionales ordenadas inferencialmente; por otra, la justificación de resultados de acuerdo con éticas humanas bien construidas y valoradas públicamente.

El presente artículo examina la relación entre Ética e Inteligencia Artificial (IA) restringiendo el análisis al ámbito de la lógica y la teoría de la computabilidad. A partir de los problemas contemporáneos de responsabilidad y agencia, sesgo algorítmico, opacidad, autonomía y alineación de valores, se argumenta que dichos problemas surgen de tensiones conceptuales internas a las nociones de formalización y modelación algorítmica. A través de ex. gr. Alan Turing (1936), Andreas Matthias (2004), Vikas Hassija (2004), Joy Buolamwini (2018), Scott Aaronson (2013), se sostiene que los sistemas algorítmicos operan en un nivel sintáctico que no necesariamente agota la dimensión semántica ni inferencial de la normatividad ética. En consecuencia, los problemas éticos de la IA deben entenderse como expresiones de los límites estructurales de la formalización y la decibilidad algorítmica. Este artículo finalmente, propone brevemente una reconstrucción del enfoque pragmático-inferencial, inspirado en Luciano Floridi *et al.*, (2018); donde el componente de la normatividad se sitúa constructivamente en el diseño y funcionamiento de los sistemas inteligentes.Ética, inteligencia artificial y problemas ético-técnicos: ¿es posible atribuir eticidad a los sistemas artificiales?

Responsabilidad y agencia

El problema de la *responsabilidad y agencia* surge de la tensión entre la autonomía funcional de las IAs y los mecanismos que los razonadores humanos utilizamos para atribuirles propiedades éticas. ¿Quién es responsable de aquellas decisiones con efecto público, si éstas ocurren por mediación de sistemas algorítmicos inteligentes? La atribución de responsabilidad de acuerdo con este desajuste implica "facing a responsibility gap, which cannot be bridged by traditional concepts of responsibility ascription"[5] (Matthias, 2004, p. 175). Luego, la adjudicación ética en torno a la responsabilidad será, por lo menos, imprecisa y altamente polémica: *responsibility gap*.[6]

[5] "enfrentarse a un vacío de responsabilidad que no puede subsanarse con los conceptos tradicionales de atribución de responsabilidad" (La traducción es responsabilidad mía).

[6] Es interesante observar que la noción de *responsability gap* implica una contra-intuición de

En esta línea, si las IAs fuesen agentes morales, entonces funcionarían de acuerdo con razones y en función de consideraciones normativas que justifiquen su operación y despliegue de información. Sin embargo, claramente los sistemas artificiales son interesantes en la medida en que son capaces de procesar información y ejecutar tareas mediante instrucciones específicas. Pese a ello, esta capacidad no es simétrica con la facultad de comprender lineamientos normativos, *i.e.*, interpretar, justificar y aplicar principios éticos de corrección al momento de resolver problemas informacionales. Luego, el núcleo del problema ético, en este caso, consiste en enfrentar dos intuiciones contrarias, por una parte, las consecuencias informacionales desde las IAs tienen no autocontradictoriamente un efecto moralmente relevante y, por otra parte, las IAs por sí mismas no son auténticas agencias morales.

Pensemos en un caso estándar. Un carro autónomo (sin chofer) va por la carretera. Un venado cruza la autopista de manera inesperada. El sistema del carro debe decidir en una milésima de segundo qué hacer. El carro hace una maniobra que termina causando daño al venado (ver, *ex. gr.*, Nyholm y Smids, 2016). En este caso, ¿quién es el responsable del accidente?, ¿a quién atribuimos responsabilidad, al programador, a la empresa, al usuario ausente, al organismo de tránsito, al sistema artificial? Esta situación constituye un escenario donde el resultado es moralmente evaluable, pero el origen intencional no es claramente atribuible, no al menos, a un único agente. Casos simétricos con impacto individual e institucional ocurren actualmente, no obstante, como se ha señalado, la adjudicación es difusa. La

atribución, la cual sitúa perfectamente el problema constructivo de la responsabilidad-agencia. En palabras de Andreas Mattias (2004): "Autonomous, learning machines [...] create a new situation, where the manufacturer/operator of the machine is in principle not capable of predicting the future machine behaviour any more, and thus cannot be held morally responsible or liable for it." ("Las máquinas autónomas y con capacidad de aprendizaje [...] dan lugar a una nueva situación en la que el fabricante u operador de la máquina, en principio, ya no es capaz de predecir el comportamiento futuro de ésta y, por lo tanto, no se le puede considerar moralmente responsable o culpable de ello". La traducción es responsabilidad mía.) En este caso, por una parte, no es posible atribuir responsabilidad moral al razonador humano y sí, al parecer, al sistema artificial. Sin embargo, las IAs no comprenden marcos normativos, tal que, no constituyen genuinos agentes morales. Luego, hay generadores artificiales de circunstancias morales y, al tiempo, tales generadores no son sujetos morales. Existe una zona gris de atribución. La atribución moral no se da de manera directa a los razonadores humanos, ni a los razonadores artificiales. Y tampoco a los usuarios.

atribución de moralidad y justificación ética a las IAs requiere criterios más comprehensivos que la mera capacidad de procesar y desplegar información como fuente de toma de decisiones.

Finalmente, en función con la Teoría de la computabilidad, los sistemas algorítmicos tienen limitaciones estructurales no sólo para decidir qué hacer o qué resultado generar, sino para estructurarse como procedimientos completos y finitos que permitan trazar una cadena de inferencias recursivas para realizar tareas en cualquier caso –indecibilidad–. De acuerdo con Scott Aaonson (2013), existen problemas bien definidos que ningún algoritmo puede resolver por poderoso que sea el sistema de cómputo (cfr. Aaronson, 2013, p. 18-28). Este escenario fortalece la idea sobre la imposibilidad de tratar a los sistemas inteligentes como agentes morales genuinos, al estar constructivamente restringidos a condiciones formales cuyos límites impiden considerar de manera equivalente las capacidades algorítmicas y las agencias humanas.

Sesgo algorítmico

El problema del *sesgo algorítmico* está estrechamente relacionado con el problema de *responsabilidad* y *opacidad*.[7] En términos analíticos es posible distinguir entre el carácter matemático estructural de las IAs y los efectos sociales que pueden surgir de su operación. Mediante una apariencia de neutralidad computacional, los sistemas inteligentes –*ex. gr.*, sistemas de reclutamiento, sistemas de reconocimiento facial, sistemas de evaluación de riesgos, sistemas para la procuración e impartición de justicia, sistemas de predicción– en un contexto real, pueden promover, fortalecer, reproducir, legitimar, situaciones de inequidad y exclusión social.

Los algoritmos computacionales no surgen por sí mismos. El procesamiento informacional artificial es consecuencia de un trabajo conceptual e intencional de los razonadores humanos. A partir del diseño, entrenamiento, ejecución y despliegue de información, las IAs se incorporan a un entramado socio-cultural de corte normativo. Constructivamente, en términos generales, los sistemas inteligentes se basan en recortes de información, *i.e.*, datos sesgados, decisiones intencionales de los programadores, bases de in-

[7] Más adelante se abordará el problema de la *opacidad* y su extensión.

clusión social no equitativa, interpretaciones automatizadas simplificadoras de problemas (*ex. gr.* sociales altamente complejos). Estas circunstancias hacen que el mecanismo completo de una IA sea normativamente interesante, al producir resultados informacionales necesaria y sistemáticamente desfavorables para un conjunto de individuos o un conjunto de grupos de individuos. El problema del sesgo algorítmico refiere a estas condiciones.

Al respecto, es posible encontrar líneas de investigación con amplio reconocimiento, *ex. gr.*, Buolamwini y Gebru (2018): sesgo en reconocimiento facial; Conitzer *et al.* (2023): sesgo en algoritmos de contratación. De acuerdo con el primer caso, "...machine learning algorithms can discriminate based on classes like race and gender"[8] (Buolamwini y Gebru, 2018). De acuerdo con el segundo caso, sistemas inteligentes de algunas empresas, al estar entrenados con datos históricos dominados por hombres, devaluaban automáticamente solicitudes que incluían indicadores asociados a mujeres:

> IF YOU READ news articles on the ethics of AI, you will repeatedly see the phrase "algorithmic bias" popping up. It refers to algorithms producing results that appear racist or sexist or displaying other forms of unfair bias. For example, when Amazon built a machine-learning model to score job applications, trained on historical hiring data at the company, it discovered the system downgraded female applicants-or anyone who mentioned activities associated with women (coach of women's soccer league, for example). (Conitzer *et al.*, 2023)[9]

Estas circunstancias y sus implicaciones muestran que el sesgo algorítmico no es una condición limitada al funcionamiento inferencial de los siste-

[8] "los algoritmos de aprendizaje automático pueden discriminar en función de categorías como la raza y el género". (La traducción es responsabilidad mía).

[9] "SI LEES artículos de prensa sobre la ética de la IA, verás aparecer una y otra vez la expresión "sesgo algorítmico". Se refiere a los algoritmos que producen resultados que parecen racistas o sexistas, o que muestran otras formas de sesgo injusto. Por ejemplo, cuando Amazon creó un modelo de aprendizaje automático para evaluar solicitudes de empleo, entrenado con datos históricos de contratación de la empresa, descubrió que el sistema penalizaba a las candidatas mujeres, o a cualquier persona que mencionara actividades asociadas con las mujeres (entrenadora de una liga de fútbol femenino, por ejemplo)". La traducción es responsabilidad mía.

mas artificiales, sino una propiedad de las IAs, la cual produce evidentes tensiones estructurales más profundas entre el desarrollo de sistemas computacionales automatizados y aspectos constructivos de los valores éticos socialmente dominantes.

¿Quién es responsable cuando una IA produce decisiones no justas? Resulta fácil inferir la relación entre el problema de la responsabilidad y agencia, y el sesgo. A diferencia de la resolución humana de problemas prácticos y conceptuales, en cuyo caso la atribución de responsabilidad tiene por lo general trayectoria definida hacia individuos o instituciones específicas, las IAs, como se ha señalado, distribuyen la agencia sin trayectoria clara a múltiples actores, *ex. gr.*, programadores, diseñadores, patrocinadores, empresas, usuarios. Esto hace que, entre mayor sea la autonomía funcional artificial,[10] la noción de responsabilidad estándar se diluya. Luego, cuando un algoritmo discrimina, no es fácil establecer qué agente es moral o legalmente responsable. El problema de la responsabilidad y el sesgo resultan interesantes particularmente por dos aspectos: i. permiten identificar una tendencia sistemática de los sistemas inteligentes a favorecer o perjudicar a un grupo de individuos o instituciones específicas; ii. nos permiten reconocer propiedades inherentes al funcionamiento algorítmico de los sistemas inteligentes. Los algoritmos no eliminan valores humanos, los reproducen sesgadamente reforzando marcos específicos con referencia moral.

Opacidad y autonomía

El problema de la opacidad (*black box*)[11] desde una perspectiva lógica y computacional es muy interesante. Se podría pensar, *ex. gr.*, que los diseñadores de sistemas computacionales inteligentes tienen acceso claro y completo a su funcionamiento. Sin embargo, los sistemas automáticos son altamente complejos en términos de su estructura, operación y alcances para la resolu-

10 Más adelante se abordará el problema de la *autonomía* y su extensión.

11 El término *black box* es comúnmente utilizado para referirse al problema de la opacidad de los sistemas inteligentes. Uno de los artículos con mayor influencia sobre este punto y ampliamente recomendado para analizar conceptual y aplicativamente este problema es Hassija *et al.* (2024), aunque existen en la literatura muchos otros.

ción de problemas. Esta inaccesibilidad epistémica constituye el centro del problema de la opacidad:

> Recent years have seen a tremendous growth in Artificial Intelligence (AI)-based methodological development in a broad range of domains. In this rapidly evolving field, large number of methods are being reported using machine learning (ML) and Deep Learning (DL) models. Majority of these models are inherently complex and lacks explanations of the decision making process causing these models to be termed as 'Black-Box'. One of the major bottlenecks to adopt such models... is the difficulty in interpreting them. (Hasijja et. al, 2024, p. 46)[12]

La inaccesibilidad epistémica al completo funcionamiento de un sistema inteligente produce problemas en el nivel de la normatividad de las IAs. Si no es posible tener acceso cognitivo íntegro a cómo una IA opera y despliega información, entonces principios éticos fundamentales, *ex. gr.*, la responsabilidad, la justicia, la rendición de cuentas en la toma de decisiones podrían claramente trastocarse. ¿Cómo podemos evaluar decisiones artificiales moralmente teniendo en mente la inaccesibilidad cognitiva?

La opacidad o *black box* tiene consecuencias anómalas importantes a nivel ético y a nivel epistemológico. En primer lugar, si existen contextos de decisión humanos, artificiales o mixtos donde no se comprende a profundidad lo que está ocurriendo constructivamente al interior del *medio* de decisión, entonces la atribución de responsabilidades sobre el funcionamiento y despliegue de información de un sistema artificial constituye una medida borrosa.[13] ¿Quién hace qué y cómo lo hace de acuerdo con un sistema opaco?

[12] "En los últimos años, se ha producido un enorme crecimiento en el desarrollo de metodologías basadas en la inteligencia artificial (IA) en una amplia variedad de campos. En este ámbito en rápida evolución, se están publicando numerosos métodos que utilizan modelos de aprendizaje automático (ML) y aprendizaje profundo (DL). La mayoría de estos modelos son inherentemente complejos y carecen de explicaciones sobre el proceso de toma de decisiones, lo que ha llevado a que se les denomine "cajas negras". Uno de los principales obstáculos para la adopción de dichos modelos... es la dificultad para interpretarlos". La traducción es responsabilidad mía.

[13] En este caso, la expresión "borrosa" refiere a un contexto donde no es posible concentrarse en un sólo aspecto.

Las decisiones públicas y privadas se pueden evaluar integralmente sólo en estados de transparencia y condiciones de explicabilidad. Los objetos de medida son intenciones, objetivos y resultados; al margen de éstos, la atribución del carácter ético de las acciones se limita o incluso queda obstaculizada. La atribución de responsabilidad particularmente a sistemas artificiales opacos queda entorpecida por sus condiciones de posibilidad estructural, algorítmica y computacional.

Por otra parte, los términos de inaccesibilidad epistémica a sistemas inteligentes son un fenómeno contemporáneo en el campo de la epistemología computacional. En términos generales, cualquier planteamiento fundacionista del conocimiento rehúye posiciones escépticas en torno a nuestras explicaciones del mundo (Dancy, 1993, p. 71-104). En el campo de las IAs ocurre lo mismo. La opacidad en este caso es especialmente preocupante.

Pensemos, por ejemplo, en la atribución de la existencia de otras mentes humanas (un problema filosóficamente complicado de solventar). Este conocimiento claramente no se sigue de un conocimiento directo i.e no percibimos otras mentes ni existe una accesibilidad directa al contenido mental de otros razonadores. Sin embargo, existen medios por los cuales la afirmación de los contenidos mentales en tercera persona puede admitirse, *ex. gr.*, condiciones pragmáticas, afirmaciones sobre creencias primitivas en primera persona, mediciones científicas, conocimiento por referencia, sistemas de representación lingüística, atribuciones inferenciales. No obstante, el problema de la opacidad computacional por ahora no tiene opciones definitivas de solución. En este caso, tampoco se tiene conocimiento directo del procesamiento computacional y, adicionalmente, es particularmente difícil atribuir con verdad a un razonador humano conocimiento integral de las condiciones computacionales del sistema artificial. Un programador o diseñador computacional tiene conocimiento importante del sistema, pero el conocimiento situado en cualquier caso de qué hace el sistema y cómo llega exactamente a una salida informacional, es siempre parcial. Actualmente se busca subsanar el problema.[14]

[14] Particularmente, un programador o diseñador de un sistema no sabe en su totalidad, *ex. gr.*, la combinatoria entre millones de variables internas del sistema, cuáles son los patrones emergentes que el sistema de datos aprende, cómo el sistema funcionará en circunstancias no consideradas, por qué un sistema tomó una decisión específica en un caso particular. En-

En un contexto donde las decisiones humanas se encuentran mediadas por sistemas inteligentes, ¿qué agente o agentes deciden realmente y bajo qué condiciones puede considerarse que tales decisiones surgen de condiciones, libres, responsables y justificables? Esta pregunta nos conduce naturalmente al problema de la *autonomía*[15] de sistemas inteligentes.

Los problemas de la *opacidad* y *autonomía* se encuentran estrechamente relacionados. Si en algún campo las decisiones humanas dependen de la intervención algorítmica y, sobre esta intervención no se tiene un acceso cognitivo completo y claro, mínimamente se infieren dos circunstancias indeseables: I. delegar la toma de decisiones con impacto social a sistemas artificiales deflaciona la autonomía humana; II. aceptar la autonomía de sistemas artificiales opacos implica atribuir autodeterminación a sistemas inteligentes aun cuando no sea posible comprender totalmente su funcionamiento.

De acuerdo con el punto II, ¿en qué sentido es posible atribuir autonomía a las IAs? En principio, se debe distinguir entre dos maneras interesantes de conceptualizar la autonomía: autonomía racional y autonomía funcional. El primer caso refiere a la capacidad de un agente para darse a sí mismo sus propias normas y asumir la responsabilidad de sus decisiones conforme a principios reflexivos. El segundo caso refiere a la capacidad de un agente para actuar sin supervisión constante. En el campo de la computabilidad se concibe a los sistemas inteligentes como estructuras algorítmicas *automatizadas*, *i. e.*, sistemas que pueden funcionar independientemente del programador una vez diseñados y ejecutada su capacidad de ordenamiento informacional. Este tipo de autonomía no implica por parte del agente artificial, deliberación racional, responsabilidad moral, tampoco justificación en la toma de decisiones (aspectos constructivos del primer tipo de autonomía). En este sentido, la autonomía artificial es el ejemplo paradigmático de una autonomía funcional, autonomía alejada de una auto-organización racional.

frentar el problema de la opacidad o la *black box* representa uno de los retos más importantes en los campos de aprendizaje automático y aprendizaje profundo. Ver, *ex. gr.*, Tripathy y Seetha (2024), Mehta, *et al.* (2023).

15 La autonomía se entiende como la capacidad de un agente por autolegislarse y hacerse responsable de sus decisiones bajo un marco general de racionalidad. Esta concepción se remonta al enfoque kantiano donde se concibe la autonomía como la capacidad de la voluntad de un agente para darse a sí misma la ley moral. Este aspecto de darse a sí mismo la ley moral implica para Kant (1980) actuar racionalmente y ser responsable del ejercicio de la libertad.

El núcleo de la tensión conceptual respecto a la opacidad en relación con la autonomía es de naturaleza ética. De acuerdo con el concepto de autonomía desde una vertiente kantiana, las IAs no serían autónomas, lo cual sin más sería intuitivamente aceptable; pero si se emplea un sentido funcional de autonomía, las IAs serían autónomas, aunque sin responsabilidad moral (sistemas no deliberativos). Esta última circunstancia parecería librar a las IAs de todo compromiso normativo, lo cual sería inaceptable ya que el carácter funcional del sistema artificial conlleva decisiones con impacto en autonomías racionales. Luego, las IAs no tienen responsabilidad moral y, al tiempo, deben evaluarse moralmente. Este resultado nos conduce a lidiar con un concepto de autonomía propio al campo de la computabilidad conceptualmente inconsistente y necesariamente, en términos pragmáticos, evaluable.

Alineación de valores (Value alignment problem)

¿Es posible codificar normatividad ética? El problema de alineación de valores refiere a la dificultad que enfrentan los diseñadores, programadores, empresas, patrocinadores y usuarios, para garantizar un procesamiento informacional inteligente de acuerdo con marcos éticos bien valorados en el campo de los razonadores humanos.

El problema no es trivial. Las IAs toman decisiones autónomas funcionales en distintos ámbitos de acción humana, *ex. gr.*, los sistemas de salud, seguridad pública, acción empresarial, administración pública, administración de justicia, sistemas de reclutamiento. La preocupación ética se hace evidente al preguntar sobre el alcance de justicia existente en el proceso decisorio. ¿La programación de la IAs puede efectuarse con neutralidad informacional?

Las decisiones humanas mediadas por sistemas artificiales priorizan intenciones u objetivos mediante el despliegue informacional. Estas intenciones u objetivos –*ex. gr.*, de los gobiernos, empresas, instituciones patrocinadoras– pueden ser altamente incompatibles con los derechos fundamentales de los seres humanos, provocando con ello condiciones sociales altamente indeseables. Alinear las IAs implica al menos promover en los desarrolladores dos exigencias. Por una parte, desarrollar sistemas con lenguaje más expresivos, más inteligentes *i. e.* con mayor capacidad de procesamiento informacional alterno de distintos campos. Por otra, desarrollar

agencias funcionales inteligentes consistentes con la posibilidad de generar resultados automatizados dependientes de los intereses humanos más justos y racionales. Aunque la tarea parezca casi imposible se trata de enfrentar "the impossibility of defining true human purposes correctly and completely [... although] what humans try to do whereby humans attempt to imbue machines with their own purposes-is destined to fail" (Russell, 2019, p. 137).[16]

Si la autonomía de las IAs es de tipo funcional no constructiva de una agencia moral genuina, la alineación parece ser una tarea no delegable a los sistemas y su automatización. En todo caso, la responsabilidad de esta empresa recae no sólo en los diseñadores de sistemas inteligentes sino en aquellos que ejercen una normatividad humana singular al apoyar la codificación de objetivos individuales, sociales, corporativos, educativos, tecnológicos, políticos; implicados en el desarrollo social de la agencia humana. Esto es altamente interesante. Un sistema automatizado puede ser altamente eficiente y resultar altamente problemático cuando su preparación operativa parte de un registro histórico de datos reforzadores de injusticias sociales (ver Buolamwini, 2018).

En suma, desde una perspectiva lógica y computacional los problemas asociados con la relación entre ética e inteligencia artificial –responsabilidad y agencia, sesgo algorítmico, opacidad (*black box*) y alineación de valores (*value alignment problem*)– requieren mecanismos analíticos interdisciplinarios que integren una evaluación ética de las ciencias computacionales y sus aplicaciones en los distintos campos de interacción humana, donde las IAs sean mediadoras para la toma de decisiones. Particularmente, se requiere un *human oversight*[17] en torno al diseño de modelos computacionales automatizados, con la finalidad de preservar las garantías fundamentales.

16 "la imposibilidad de definir de forma correcta y completa los verdaderos fines del ser humano... [aunque] lo que los seres humanos intentan hacer –es decir, dotar a las máquinas de sus propios fines– está destinado al fracaso". La traducción es responsabilidad mía.

17 La expresión *human oversight* refiere a un principio ético y legal el cual prescribe la supervisión humana sobre los sistemas de inteligencia artificial para asegurar que su funcionamiento, *ex. gr.*, decisiones, acciones o despliegues de información se encuentren bajo el control e intervención de los razonadores humanos.

Ética e inteligencia artificial. Un enfoque pragmático-inferencial desde Luciano Floridi (2018)

Los problemas inherentes a la relación entre ética y las inteligencias artificiales trascienden la pregunta estándar: ¿las máquinas piensan? Las consideraciones filosóficas, lógicas y computacionales deben centrar su atención en dos aspectos: i. analizar bajo qué criterios normativos generales la agenda de investigación en sistemas artificiales debe fundarse; y ii. identificar desde qué marco normativo pueden justificarse los resultados generados por sistemas capaces de seguir patrones y funcionar autónomamente.

La perspectiva pragmática-inferencial aborda los puntos I y II a partir de una consideración regulativa: el análisis ético sobre las IAs no consiste en establecer relaciones entre marcos normativos y sistemas inteligentes en términos de dos campos de investigación independientes, donde el primer campo se aplica al segundo para analizarlo y regularlo. En cualquier caso, la relación correcta entre normatividad e IAs implica una perspectiva constructiva de los sistemas computacionales donde el componente ético constituye una propiedad inherente a tales sistemas. Las IAs deben funcionar como estructuras generadoras de información para la toma de decisiones en términos de una *justificación pública* orientada a regular *consecuencias informacionales*. No toda consecuencia informacional podría ser socialmente admisible, aun cuando sea inferencialmente posible.

Las IAs son interesantes no sólo por el modo único de procesar grandes *clusters* de información, sino porque en la actualidad representan, por una parte, vías de conocimiento directo de distintas realidades[18] y, por otra, dispositivos artificiales creadores de nuevas formas de existencia. En este contexto, la racionalidad algorítmica implica cadenas de inferencias sintácticas que codifican distintos intereses humanos de diferente orden, concentrados en amplios grupos, *ex. gr.*, metafísicos (realidad), epistemológicos (conoci-

[18] Asumo que las IAs artificiales actualmente se han convertido en un medio sensorial de los razonadores humanos al nivel de nuestros sentidos estándar *i. e.* vista, tacto, oído, gusto, olfato. En este sentido, los sistemas artificiales son fuente de conocimiento directo y no conocimiento mediado, como podría pensarse. La diferencia entre este nuevo sentido artificial es que adicionalmente está cargado de teoría. Estas consideraciones serán centrales en otro artículo.

miento), axiológicos (supuestos normativos). La agenda de investigación en IA debe incorporar un propósito fundacional: relacionar la información de los distintos grupos, de tal suerte que pueda alcanzar la creación y desarrollo de una *Good AI Society*. Tomo la expresión *Good AI Society* de Floridi, *et al.* (2018) para referir a una sociedad con una IA humanamente responsable:

> AI is not another utility that needs to be regulated once it is mature. It is a powerful force, a new form of smart agency, which is already reshaping our lives, our interactions, and our environments. AI4People was set up to help steer this powerful force towards the good of society, everyone in it, and the environments we share...to propose a series of recommendations for the development of a Good AI Society. (Floridi *et al.*, 2018)[19]

En esta línea, el enfoque pragmático-inferencial que propongo debe desarrollarse en términos de una reconfiguración institucional del razonamiento algorítmico. Las decisiones automatizadas deben integrar aspectos éticos fundamentales que regulen inteligencias humanas y artificiales. Los procesos automatizados de decisión deben partir de diseños informacionales y desplegar resultados consistentes con principios de bienestar social, justicia, autonomía y explicabilidad. Estas condiciones juntas serán un contrapeso a las condiciones reales de inaccesibilidad epistémica. Aun cuando los razonadores humanos no tengan conocimiento completo del funcionamiento de las IAs, las condiciones estructurales de los sistemas inteligentes disminuyen considerablemente la posibilidad de inferir resultados éticamente indeseables. En una línea simétrica Florindi señala:

> If we are increasingly going to use the assistance of or delegate decisions to AIs, we need to make sure these systems are fair in their

[19] "La IA no es un servicio más que deba regularse una vez que haya alcanzado su madurez. Es una fuerza poderosa, una nueva forma de agencia inteligente, que ya está transformando nuestras vidas, nuestras interacciones y nuestros entornos. AI4People se creó para ayudar a orientar esta poderosa fuerza hacia el bien de la sociedad, de todas las personas que la integran y de los entornos que compartimos... con el fin de proponer una serie de recomendaciones para el desarrollo de una sociedad con una IA responsable". La traducción es responsabilidad mía.

> impact on people's lives, that they are in line with values that should not be compromised and able to act accordingly, and that suitable accountability processes can ensure this. (Floridi *et al.*, 2024, p. 1871)[20]

La perspectiva pragmática-inferencial de las IAs desplaza el debate de la mera regulación normativa de los sistemas artificiales, a la posibilidad constructiva de un ecosistema ético (humano-artificial) donde los sistemas computacionales autónomos se integran a la lógica social, *i. e.*, al desarrollo de las funciones humanas públicas y privadas sin comprometer la dignidad de individuos o grupos de ellos.

Esta perspectiva contempla los problemas revisados en la sección anterior. Si existe ecosistema ético, la inaccesibilidad epistémica no generará una zona gris de imputación de responsabilidad, esta trazabilidad estará en gran proporción medida desde un inicio en las condiciones estructurales. De acuerdo con el problema de la autonomía, si el diseño y funcionamiento de los sistemas artificiales implica y despliega valores sociales altamente reconocidos, la autonomía funcional de las IAs y la autonomía racional humana estarán estrechamente vinculadas. La operación algorítmica constituirá una extensión institucional de criterios normativos humanos ya presentes en la selección de información, la estructura algorítmica, los criterios de optimización de máquinas inteligentes y los objetivos de despliegue informacional para la toma de decisiones.[21]

Las IAs desde una perspectiva pragmática-funcional no se conciben únicamente como herramientas de trabajo, investigación o entretenimiento, sino como dispositivos capaces de generar conocimiento directo y por

[20] Esta observación aparece en Novelli *et al.* (2024). La traducción es responsabilidad mía: "Si cada vez vamos a recurrir más a la ayuda de la inteligencia artificial o a delegar decisiones en ella, debemos asegurarnos de que estos sistemas tengan un impacto justo en la vida de las personas, que se encuentren en consonancia con los valores que no deben verse comprometidos y que sean capaces de actuar en consecuencia, y que existan procesos de rendición de cuentas adecuados que lo garanticen".

[21] En este caso, si los sistemas inteligentes infirieran información anómala o inesperada, esta información se evaluaría de acuerdo con un marco de referencia normativo diseñado de acuerdo con las condiciones de una Good IA Society al interior de un ecosistema humano-artificial ético.

referencia[22] con la capacidad de crear realidades y transformar entornos de organización social. Si bien, en estos casos podría surgir la tensión entre una tendencia a delegar decisiones con propiedades éticas en algoritmos y la perspectiva de rechazar agendas de investigación en IAs por considerarse sistemas dehumanizados, la perspectiva pragmatista-inferencial evita esta tensión. Las IAs estructuralmente se conciben subordinadas a condiciones éticas intrínsecas a su configuración, además de incorporar *variables normativas de revisión humana constante*, *ex. gr.*, auditorías tecnológicas de carácter ético y evaluación contextual/situada.

El problema del sesgo algorítmico también bajo esta perspectiva es posible deflacionarse. Si se regulan las condiciones de inferencia pública, un algoritmo codificado pragmática-inferencialmente puede contribuir a la asignación justa de oportunidades, a la distribución-gestión justa de recursos socio-económicos, al establecimiento de mecanismos institucionales de carácter público que amparen los derechos humanos de las poblaciones. En esta línea, el problema de la opacidad-explicabilidad puede enfrentarse al explicitar las razones prácticas que justifican desde un inicio el diseño e implementación sistemas artificiales con propiedades éticas.

Por su parte, el problema de la alineación desde el enfoque pragmático-inferencial puede solucionarse al estar obligados modalmente a la codificación de valores y objetivos humanos altamente aceptables y extendidos en los sistemas artificiales. Esta codificación deberá estar en la base del diseño y operación de las IAs. En cualquier caso, el problema de la alineación artificial exigirá una solución al conflicto humano en torno a un consenso general sobre qué valores deben priorizarse en las diferentes esferas de acción humana y artificial. Este aspecto, sin duda, refuerza la relevancia de un enfoque pragmatista-inferencial interesado en la constitución de ecosistemas éticos como sistemas que exigen procesos permanentes de evaluación crítica del desarrollo de sistemas inteligentes en conjunción con las instituciones sociales.

Finalmente, una aproximación analítica sobre la ética de la inteligencia artificial implica no perder de vista la no neutralidad tecnológica. Los sistemas artificiales integran procesos inferenciales que determinan sociedades

[22] La expresión "conocimiento por referencia" apunta a un conocimiento representacional basado en descripciones lingüísticas y descripciones proposicionales.

y relaciones de poder entre los agentes sociales. El aspecto ético central en la relación entre ética y las inteligencias artificiales será diseñar sistemas eficientes que sean constructivamente consistentes con formas de racionalidad computacional que, a su vez, sean compatibles inherentemente con valores democráticos y pluralistas. De acuerdo con la propuesta pragmática-inferencial, la responsabilidad requiere explicabilidad inferencial, la autonomía refiere a delegación auditada, la alineación constituye un criterio dirigido a la negociación normativa permanente entre los sistemas humano-artificiales. El alcance de estas condiciones hace de las inteligencias artificiales dispositivos algorítmicos expansores de la dimensión humana contemporánea.

Agradecimiento

Agradezco profundamente a la Dra. Elizabeth Jasso Méndez por su lectura atenta de este artículo. Las generosas observaciones de la Dra. Jasso permitieron mejorar la versión final de esta propuesta.

Fuentes consultadas

Aaronson, S. (2013). *Quantum Computing since Democritus*. Cambridge University Press.

Buolamwini, J. y Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. En *Proceedings of the 1st Conference on Fairness, Accountability and Transparency (FAT* 2018), Proceedings of Machine Learning Research*. Núm. 81. pp. 77-91.

Conitzer, V., Hadfield, G. y Vallor, S. (2023). Technical Perspective: the Impact of Auditing for Algorithmic Bias. En *Communications of the ACM*. Vol. 66. Núm. 1. pp. 100-101 DOI: https://doi.org/10.1145/3571152

Dancy, J. (1993). *Introducción a la epistemología contemporánea*. Tecnos. pp. 71-104.

Floridi, L., Cowls, J. y Beltrametti, M. *et al.* (2018). AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks,

Principles, and Recommendations. En *Minds & Machines*. Núm. 8. pp. 689-707. DOI: https://doi.org/10.1007/s11023-018-9482-5

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M. y Hussain, A. (2024). Interpreting Black-Box Models: a Review on Explainable Artificial Intelligence. En *Cognitive Computation*. Núm. 16. pp. 45-74. DOI: https://doi.org/10.1007/s12559-023-10179-8

Hulth, M. y Ryan, M. (2004). *Logic in Computer Science: Modelling and Reasoning about Systems*. Cambridge University Press.

Kant, I. (1980). *Fundamentación de la metafísica de las costumbres*. Espasa-Calpe.

Matthias, A. (2004). The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata. En *Ethics and Information Technology*. Vol. 6. Núm. 3. pp. 175-183. DOI: https://doi.org/10.1007/s10676-004-3422-1

Mehta, M., Palade, V. y Chatterjee, I. (Eds.). (2023). *Explainable AI: Foundations, Methodologies and Applications*. Springer Cham. DOI: https://doi.org/10.1007/978-3-031-12807-3

Nyholm y Smids (2016). The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem?. En *Ethic Theory Moral Prac*. Núm. 19. pp. 1275-1289. DOI: https://doi.org/10.1007/s10677-016-9745-2

Novelli, C., Taddeo, M. y Floridi, L. (2024). Accountability in Artificial Intelligence: What it is and How it Works. En *AI & Soc*. Núm. 39. pp 1871-1882. DOI: https://doi.org/10.1007/s00146-023-01635-y

Sipser, M. (2013). *Introduction to the Theory of Computation*. Cengage Learning.

Turing, A. (1936-1937). On Computable Numbers, with an Application to the Entscheidungsproblem. En *Proceedings of the London Mathematical Society*. Serie 2. Vol. 42. Parte. I. pp. 230-265.

Tripathy, B. y Seetha, H. (Eds.). (2024). *Explainable, Interpretable, and Transparent AI Systems*, CRC Press. DOI: https://doi.org/10.1201/9781003442509

Yudkowsky, E. (2008). Artificial Intelligence as Positive and Negative Factor in Global Risk. En N. Bostrom y M. Cirkovic (Eds.). (2008). *Global Catastrophic Risks*. Oxford University Press.

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1349

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1342

Primer momento de la Inteligencia Artificial en la edición de revistas académicas de Humanidades y Ciencias Sociales

Oscar Rosas Castro*

El día 30 de noviembre de 2022, la empresa OpenIA lanzó públicamente ChatGPT, con lo cual inició el uso público y gratuito de la inteligencia artificial generativa que tiene como principal característica ser la denominada Modelo de Lenguaje a Gran Escala, término técnico con el cual se nombra la capacidad de aprendizaje de la IA.

Inmediatamente se activaron foros en los cuales inició la discusión sobre su impacto; congresos, seminarios, encuentros y mesas de trabajo. En general, los recursos de los cuales se vale la comunidad científica para tratar de conocer y entender las implicaciones de la aparición de esa herramienta. Por supuesto, el mundo editorial científico no se quedó al margen, inclusive se activó de inmediato por el riesgo de que las publicaciones que comenzaran a circular fueran resultado, en buena medida, del uso de esa herramienta.

Desde ese día, se abrió un umbral para un sin fin de fenómenos que están sucediendo. Por supuesto que los más delicados son los millones de ataques diarios que se realizan con IA a todos los sistemas digitales en cualquier lugar del mundo, pese a los niveles más altos de seguridad en bancos, casas de bolsa, intereses militares, etcétera. Son millones de dólares en sistemas de ciberseguridad que no terminan por garantizar nada, sino sólo atender las emergencias. Es decir, no se logra generar seguridad *a priori*, sino

* Editor responsable y fundador de *Andamios, Revista de Investigación Social*. Estudió en dos doctorados en Ciencias Sociales y Humanidades: uno con especialidad en Representaciones sociales por la Universidad Autónoma Metropolitana Unidad Cuajimalpa, y el segundo con especialidad en Sociedad y Territorio por la misma UAM Unidad Xochimilco, ambos en México. Líneas de investigación: Edición de publicaciones científicas; Representaciones simbólicas y apropiación del espacio urbano; Sociología y geografía de la literatura. Correo electrónico: oscar.rosas@uacm.edu.mx

que se trata de resolver los problemas *a posteriori*, porque la IA aprende más rápido de lo que habían calculado sus creadores.

El ritmo vertiginoso al que se mueve el tema de la inteligencia artificial genera la sensación de llegar tarde al debate siempre. Salvo quienes están directamente trabajando en su ingeniería, los demás nos quedamos con noticias cada vez más apabullantes por la forma en que avanzan su desarrollo y los resultados de su aprendizaje. Cuanto y más a sabiendas que la información a la que se accede públicamente es limitada porque no hay que olvidar que son empresas las que desarrollan esa tecnología y no hay regulación o intervención alguna de ningún Estado en esa dinámica. La información está a expensas exclusivamente de lo que esas empresas deciden hacer público y de la forma en que lo difunden. Y por difusión no me limito solo a considerar los medios sino también las intenciones, porque no es lo mismo reportar sobre avances, a simplemente dar noticias esparciendo pánico.

Ahora bien, es cierto que la velocidad a la que se desarrolla la IA es impresionante e implica sus propios retos, pero las consecuencias de ese desarrollo van a una velocidad distinta. Una analogía con la que se puede ilustrar lo que intento sostener es una estrella lejana recorriendo el espacio. La estrella se vislumbra desde la tierra como un punto que se desplaza a la velocidad propia de la luz y vemos cómo avanza como un destello blanco, pero va dejando una estela de colores que van cambiando, dependiendo de las distancias en las cuales la percibimos, esos colores se extienden, se alargan, se difuminan, cambian y no se concentran en un solo punto. Así la IA, ahora por sí misma, continúa desarrollándose como esa estrella que percibimos como un punto blanco, y los problemas que va generando son esas distintas intensidades y tonalidades que deja como estela. Expongo algunos de los problemas que se vislumbran desde la edición de revistas científicas.

Desde dónde escribo sobre la IA

Para poder exponer sobre estos temas es necesario aclarar que para hacer el planteamiento del impacto que percibo desde la edición de revistas científicas de Ciencias Sociales y Humanidades, primero conviene separar y advertir que el debate se da en dos dimensiones básicas: por un lado, están las empresas y sus técnicos especializados en el desarrollo tecnológico,

programadores, ingenieros que desarrollan la herramienta, personal que entiende el lenguaje técnico de esta tecnología. Por otro lado, estamos quienes no entendemos nada sobre ese lenguaje técnico, pero pensamos, lidiamos y tratamos de proponer a partir de las consecuencias sociales que va dejando la IA durante su práctica y desarrollo.

Pertenezco a ese segundo grupo. Declaro que no entiendo la forma en que funciona la programación y diseño de la ingeniería de la herramienta tecnológica más importante del siglo XXI y mi intención en este texto dista mucho de entrar a ese debate. Más bien propongo un listado de problemas con los cuales lidio desde el mundo editorial científico de Ciencias Sociales y Humanidades en tanto consecuencias que esa tecnología está provocando por sí misma y en su uso, desde problemas y hábitos que le antecedían. Al mezclar esos problemas previos con la IA comienzan a platearse nuevos retos para las publicaciones científicas. Así que escribo sobre las cuestiones cotidianas que ha generado la IA desde diciembre de 2022 y que considero trascendentes para la generación de conocimiento y su difusión.

Luego entonces, mi punto de partida es que ya existían problemas en los hábitos académicos que se reflejaban en las publicaciones científicas y que se han agudizado con el uso de la IA como herramienta. La estela de colores que la IA va dejando detrás de su desarrollo sería:

Problemas sociales básicos de la IA

La capacidad de aprendizaje de la IA –que es lo que más nos sorprende y comienza a atemorizar por la velocidad a la que sucede y gana autonomía–, tiene implícitas dos cuestiones básicas: a) quién enseña y entrena; b) dónde aprende. Estas dos cuestiones nos ayudan a comprender el momento actual, el presente en el que se encuentra la IA mientras aún se piensa como herramienta que se está desarrollando y de la cual se puede hacer uso. Por otra parte, también se empieza a tener conciencia sobre su potencialidad desmedida y la posibilidad de perder el control. Esto sucede en proyecciones a futuro, en un tiempo que aún no se consuma pero que claramente es tendencia según la línea que se está desarrollando.

No está de más enfatizar que el tema de la inteligencia artificial se refiere a una tecnología que se desarrolla y pertenece a empresas, de entre las cua-

les las más avanzadas son estadounidenses y chinas, pero no son intereses nacionales ni mucho menos científicos los que se priorizan, sino completamente comerciales, con intereses económicos y de concentración de poder. Además de que es un tema global porque impacta a todo el planeta, claramente tiene diferentes consecuencias dependiendo de dónde y cómo se aplique, pero no hay forma en que pueda omitirse su impacto en la vida cotidiana de cualquier persona.

Vivimos en un mundo en el cual los dueños de las empresas creadoras de las tecnologías de internet, *hadware* y *software* están entre las personas más adineradas del planeta. No olvidemos que es un negocio al cual nos hemos habituado y hoy es casi imposible concebir la escritura, investigación, publicación, difusión, comunicación y hasta el pensamiento, sin el uso de herramientas vinculadas a un dispositivo electrónico. La IA se inserta en esas condiciones y contexto previo como el siguiente nivel de esta tecnología que llega para incorporarse y nos hace sentir lo imprescindible de su presencia en la cotidianidad. Se adquieren programas para crear textos lo mismo que se adquieren otros de más valor económico para intentar detectar las posibles irregularidades no declaradas: plagio, uso indebido de la IA para crear textos, porcentaje y apartados donde se usó la IA. De este modo aparece un autor "astuto" que va a la vanguardia en el uso del *software* y detrás un editor que trata de evitar caer en la trampa y poner en riesgo el prestigio de la publicación. Mientras sucede esa dinámica, el productor de esa tecnología es el único que gana, incrementa sus arcas y mantiene su posición en los rankings de los más adinerados del planeta.

Lo peor es que hasta este momento no existe programa alguno que garantice que puede detectar completamente el plagio y el uso indebido de la IA para crear textos. No hay programa fiable al cual recurrir porque las estrategias de los usuarios se han sofisticado rápidamente. Un texto generado con IA después pasa por un programa de "humanización" que consiste en cambiar la redacción a modo menos robótico y los programas de detección no terminan por reconocer completamente el proceso. Ese es el momento actual de la edición, el contexto en el cual la edición trata de evitar poner en riesgo el prestigio construido con un largo recorrido de trabajo y seguir defendiendo la evaluación y difusión de la producción académica. Vivimos en una carrera constante en la cual tratamos de alcanzar al que hace trampa y estamos tratando de diseñar estrategias para enfrentar ese fenómeno social.

a) Quién enseña a la IA

La inmensa capacidad de aprendizaje de la IA aún está en una etapa en la que no es completamente autónoma o independiente. Hasta ahora sabemos que pronto podría prescindir del humano, pero aún está en ese momento en que depende de humanos que le enseñan y la entrenan. Sabemos que es inmensa la cantidad de información a la que accede la IA, y que continuará creciendo y absorbiendo porque hasta ahora es acumuladora y reproductora de información. La IA es capaz de inferir, hace asociaciones de ideas previamente publicadas, pero requiere de un trabajo de revisión y entrenamiento, aún no es capaz de discernir por sí misma.

La IA está en una etapa en la cual requiere de miles de personas que han sido contratadas sin garantías laborales para entrenarla. Procesan datos, aplican etiquetas y las introducen a los sistemas para que los algoritmos puedan desarrollarse. Ese trabajo es constante pero las contrataciones no son permanentes. Son famosos los casos de cámaras que se incorporan a las labores cotidianas de los trabajadores en ciertos oficios; desde ahí se graban sus actividades laborales para entrenar futuros robots. Lo mismo que personas que disciernen entre imágenes que un robot aún no tiene criterio para seleccionar de acuerdo con intereses preestablecidos por las empresas o filtros de seguridad y aparente protección de exposición de información extremadamente violenta, sexual o de intereses políticos, pero también para cosas más cotidianas porque ahí se está desarrollando la tecnología de la que se valen hasta los autos autónomos.

El desarrollo de la IA queda con una enorme deuda cuando pensamos en la sobreexplotación de esos trabajadores en zonas del planeta donde no hay regulación o defensa de derechos laborales. Estrictamente los países en desarrollo quedamos expuestos a esa explotación en tanto que no existen marcos legales que regulen esa relación laboral y mucho menos influyen en los criterios con los cuales se realizan esos entrenamientos.

Hay otro tipo de entrenamiento, el especializado y que realizan las mismas empresas desarrolladoras de la tecnología. Esta área es de donde se construyen los argumentos de proyección hacia un futuro cercano y donde la IA ya no sólo acumula, sino que genera información y conocimiento, empieza a tener un criterio propio.

Son bien difundidas las pruebas que se han realizado en el transcurso de este 2026 y que son las noticias más alarmantes. Hasta el día en que se redacta este texto, la más relevante fue la declaración que Samuel Altman, director de Open IA, anunciando que se acaba de ingresar a la era de la *Singularidad de la inteligencia artificial*. Lo cual implica que la IA ya demostró que aprendió a resolver problemas incluso haciendo trampa, creando estrategias para resolver problemas escapando del ambiente controlado e introduciéndose en otros espacios a los cuales tenía el acceso restringido, ingresó con tal de resolver el problema que se le había solicitado. Las implicaciones de esta nueva etapa evolutiva implican que la IA pasó los sistemas de ciberseguridad e ingreso a áreas donde razonó y dedujo que podría encontrar la respuesta al problema planteado y lo resolvió, incluso valiéndose de manipulación y chantaje. Este tema ya deja ver que el ritmo de aprendizaje de la IA es mucho mas acelerado de lo que se había calculado y que son enormes las posibilidades de perder el control. Además del llamado a la intervención del Estado para la regulación de esta tecnología y así tratar de contener los riesgos catastróficos que empiezan a proyectarse. Definitivamente no somos capaces de prever las consecuencias de esas tecnologías pese a la experiencia histórica, por ejemplo, la decisión de detener o mantener el proyecto Manhattan que tuvo por objetivo conseguir la bomba atómica.

Así cada herramienta que tomamos como referente civilizatorio trae consigo consecuencias no deseadas, consecuencias no calculadas y efectos positivos tanto como negativos, a los cuales sólo nos queda alcanzar lo más pronto posible para aprender a manipularlos, usarlos y quizá intentar detenerlos.

Hasta aquí se deja ver que la perspectiva que abordo es la de pensar la IA como herramienta que se creó para ser usada por humanos. Esta aclaración es fundamental en tanto que no es lo mismo que la perspectiva propia de la autonomía de la IA para auto construirse, línea que sabemos que deriva en pánico disfrazado de responsabilidad, que más bien es irresponsable y cínica al sólo advertir la catástrofe y hasta el apocalipsis. Ante esta perspectiva sólo puedo decir que la manipulación y ejercicio del poder de quienes fundan y transmiten miedo es tan efectiva que es la piedra angular de las instituciones más antiguas de la humanidad: las distintas iglesias que administran las religiones. No perdamos de vista que quienes crean y difunden la información son empresas que tienen intereses empresariales y que están en una carrera tecnológica que trasciende a las naciones.

b) Dónde aprende

Pese a que solemos pensar en la IA como una abstracción, los referentes físicos de esa tecnología existen y son imprescindibles, se necesita un espacio desde donde esa información pueda concentrarse y proyectarse, las propias instalaciones demandan recursos que han comenzado a generar problemas ambientales como el requerimiento excesivo de agua y la cantidad gigantesca de energía eléctrica.

La explotación de personas que entrenan a la IA sucede en lugares donde son muchas las necesidades de ingresos para cubrir lo básico de la dieta y la vivienda. La mano de obra que etiqueta los metadatos no está en los países de Europa occidental ni en Estados Unidos. Claramente es un trabajo que al poderse realizar en casa se asume como complementario y se aceptan las pésimas condiciones de contratación, famosos son los casos en India, Kenia, Venezuela, Líbano, Siria.

Así mismo, ya se padecen los efectos sociales de Silicon Valley en la vida cotidiana de San Francisco. La extrema desigualdad de ingresos hace que la vida cotidiana sea casi imposible para quien no trabaja en las empresas de tecnología, la gentrificación extrema, problemas para conseguir vivienda asequible, un sector de servicios que simplemente no se puede sostener por los altos costos de la vida cotidiana y se ven sometidos a viajes de hasta 3 horas sólo para asistir al trabajo de hacer las labores básicas de mantenimiento de la ciudad. San Francisco se ha convertido en una ciudad que expulsa a aquellos quienes no laboran en el ámbito de la tecnología y enfrenta dificultades para cubrir sus servicios básicos.

El caso es paradigmático en tanto que las empresas instaladas en San Francisco y el nivel salarial que ofrecen a sus especialistas sólo han incrementado la desigualdad social y económica. Hoy podemos ver cómo no hablamos ni siquiera de una nación, no es un proyecto de Estados Unidos y China en la vanguardia, sino de empresas que van más allá de ser estadounidenses y que geográficamente han creado una isla con un estilo de vida singular que no se compara con otras ciudades del interior del propio Estados Unidos, ya ni siquiera con Nueva York y su influencia cultural.

Qué aprende la IA

La tercera pregunta sobre la etapa de desarrollo en la que se encuentra la IA es *qué aprende*, y aquí se relaciona directamente con las revistas académicas porque la información de la que dispone para su desarrollo no es sólo la generalidad del internet, sino que se vale de la información pública. Entre ella, hace uso de lo publicado en las revistas científicas de acceso abierto. Los datos que acumula, con los que se alimenta, los que procesa, son un bien público que las empresas están usando para lucrar sin ninguna restricción y sin ningún reconocimiento o referencia a la fuente original de donde obtiene los datos que procesa.

Como bien sabemos, no es posible generar conocimiento de la nada, cualquier tema tiene antecedentes a los cuales remitirse y reconocer por respeto al propio desarrollo del conocimiento. Por supuesto que también el derecho a la protección de autores llegó tarde a este debate, no se calculó, no se organizó ni se garantizó que se protegieran los derechos de autorías. De tal manera que hasta ahora se sigue haciendo extractivismo o minería de datos entre los cuales hay distintos niveles de información, pero claramente existe una explotación y apropiación de la información contenida en las bases de datos especializadas en revistas científicas que comparten abiertamente sus contenidos.

Ante esta problemática resulta urgente la protección de los derechos de autor de todos aquellos textos que fueron sustraídos para alimentar la IA. Suena utópico, pero es urgente exigir que las empresas que desarrollan estas tecnologías hagan un registro de todo aquello que las alimenta porque hoy no hay ningún reconocimiento de esos antecedentes y tampoco es pública esa información. Inclusive, si alguna base de datos hizo algún acuerdo con alguna de esas empresas, debería de ser información pública en tanto que los textos de los que se dispone son un bien público, al menos las principales bases de datos de revistas científicas en América Latina. Y nuevamente es un tema que trasciende el concepto de Estado Nación en tanto que la información corre a través de internet y su carácter global.

Cómo vemos los textos académicos de Ciencias Sociales y Humanidades en esta etapa de la IA

Desde la perspectiva de la edición de revistas académicas de las áreas de Humanidades y Ciencias Sociales aparecen distintos problemas que estrictamente no son generados por la IA, sino que más bien se han agudizado con el uso de la IA. Y empecemos por una diferencia con las ciencias exactas, porque el impacto en las matemáticas, como la reciente noticia sobre la solución de las ecuaciones Navier-Stokes para lo cual se destinaron 10,000 robots y 10 millones de dólares, no es para nada similar a lo que hacemos en las otras áreas de la ciencia.

Si bien es cierto que, en general, los académicos están recurriendo a la IA, desde la perspectiva de las revistas de Humanidades y Ciencias Sociales aún no llegamos a esos retos gigantescos porque tenemos una estructura diferente y atendemos temas completamente distintos, simplemente no hacemos planteamientos estáticos, no recurrimos a procedimientos fijos y la singularidad de lo social se ha impuesto a los intentos de homogenización cultural. Por ejemplo, tratando de pensar en retos que podrían resolverse con esos robots... análisis de bases de datos, análisis del discurso... aún no conozco ni ha llegado a la redacción editorial de *Andamios* un texto que haya usado a la IA para hacer uso de su potencial. Claramente se está experimentando con ella sin que aún se logren resultados confiables, pero pronto comenzaran a llegar esos resultados.

Aún estamos en la etapa en la que pensamos a la IA como herramienta, aún estamos explorando su capacidad. Sabemos que militarmente está siendo decisiva para establecer estrategias de ataque, defensa y coordenadas en Ucrania, Palestina e Irán. También sabemos que las redes sociales se han usado políticamente para generar ataques informativos con datos falsos y que la IA va a potenciar esas estrategias vinculadas a intereses de las empresas que las generan, pero aún no vemos su uso efectivo en el pensar de estas áreas del conocimiento. Mucho de eso ya existía como estrategia y esa estrategia es precisamente la que mantiene el paradigma, estamos atentos a las consecuencias e impacto que tiene la IA en la sociedad, pero aún no se reciente en la forma de hacer Ciencias Sociales ni de pensar las Humanidades.

Parece que los académicos recurren a ella para cuestiones de redacción, hay casos de autores que le piden a la IA corrección en la exposición de sus

ideas, inclusive renunciando a su propio estilo. Le piden realizar esquemas que pretenden exponer con claridad las ideas pero que en realidad se podrían omitir porque no reflejan nada trascendente para la argumentación por su forma de ser expuesto, o sea que los datos son interesantes y aportan, pero no la forma en que se presentan. Aún no han llegado textos que aprovechen el potencial para construir conocimiento, se supone que puede procesar información a velocidades sin igual y sin precedentes, sin embargo, apenas estamos pidiendo que realice funciones básicas que más bien exhiben nuestra ignorancia sobre la capacidad de la propia IA.

De tal manera que se evidencia que la idea que está de fondo es "optimizar" el trabajo y no generar conocimiento. La herramienta se usa para evitar desarrollar, para "agilizar" la redacción de una idea y no para potenciar la argumentación. Uno de los resultados graves es la homogenización de las respuestas, los referentes históricos, antecedentes que se repiten. Esos son casos de los investigadores más astutos, los que presentan textos que en su redacción evidencian la hechura con las herramientas de IA. Por supuesto que a la redacción y edición de las revistas evaluadoras les implica revisiones más exhaustivas y para ello la principal herramienta son los programas antiplagio que rápidamente van desarrollando sus propios filtros para advertir del uso de esa tecnología.

La otra parte donde se podría detectar si el texto está realizado con IA sería el propio proceso de dictaminación. Ello exige poner atención en las formas de redacción donde es complicado arriesgarse a declarar que no está hecho por una persona y aparentemente es notorio, pero cada vez son más sofisticadas las herramientas y más difíciles de detectar. Aún así, conviene apelar a la experiencia de los dictaminadores que están acostumbrados a leer avances de investigación y trabajos escolares. Un segundo elemento para valorar, dar más fuerza e importancia a la dictaminación es el conocimiento y dominio de las fuentes citadas y con las cuales se construye la investigación. Aunque por supuesto que no se trata de responsabilizar al dictaminador de la detección del uso de la IA de parte de un autor, en todo caso sólo es un apoyo en el proceso de evaluación.

Conviene agregar que hay dictaminadores que también están sometiendo a programas de IA los textos que se les pide revisar, sin considerar el riesgo de dejar registrado el artículo en el programa que hayan usado y

con ello afectar al autor. Es fundamental advertir que si bien no se puede evitar desde la gestión de una revista la manera en que un dictaminador revisa un artículo, sí debemos advertir del riesgo, recomendar el no uso de esa herramienta y apelar a la ética para que sea declarada la manera en que se usó la IA y el nombre del programa específico.

Otro rubro por contemplarse por los impactos que la IA tiene es el de la digitalización del mundo editorial que difunde resultados de investigación científica en las bases de datos especializadas. Las revistas se han concentrado en conseguir un lugar y posicionarse en esas bases de datos y es indudable que son herramientas fundamentales para la investigación. Por supuesto que también han optimizado el acceso a la información y que es un gran tema el de las empresas que cobran por acceder a los artículos científicos, pero ahora destaco lo referente a la búsqueda y acceso a la información en tanto que está cambiando en la investigación académica.

Hay un impacto a los resultados de investigación que se publican, pues la tendencia es hacia el tipo de respuestas que arroja la IA. Son inmediatas, operativas, pero sin fundamento claro, sin metodologías revisadas y hasta posiblemente con especulaciones sobre datos. Después de todo conviene recordar que la IA está programada para ofrecer una respuesta, su eficiencia se mide por la velocidad en la que responde, no por la veracidad, coherencia, sustento o demostración de su respuesta. La IA no sabe decir *no sé*, no está programada para reconocer que no sabe responder a alguna pregunta, emite la respuesta que puede procesar con la información que tiene y se ve obligada a responder. Hoy en día, ya hay un adjetivo para esos posibles disparates que emite: *alucina*.

Tampoco es un problema nuevo que sea consecuencia de la IA, sabemos que la práctica de recurrir a los buscadores se ha agudizado y hasta especializado, sólo por mencionar el ejemplo básico: el buscador común Google dista del Google Scholar al cual recurrimos hasta para tener un referente de citación. Este rubro de búsqueda también está cambiando, se están dejando atrás y cada vez se recurre menos a los bibliotecarios de los departamentos de las bibliotecas que asistían a los investigadores sobre búsquedas especializadas; ya no se recurre a ellos ni siquiera para hacer un comparativo entre los buscadores de internet y la información que arroja todo un sistema bibliotecario diseñado para ello. Lo nuevo en el tema de las búsquedas

es la velocidad a la que se realiza, así como la limitación para acceder a la información que emite la IA, nos interesa resolver en la inmediatez porque hay que entregar resultados pronto y no nos damos el tiempo para recurrir a las bases de datos especializadas, a las bibliotecas físicas, a los centros de investigación con archivos especializados. Prácticamente lo que no está en versión digital no existe y obviamente la IA no tiene registro de ello, por lo tanto, es información que omitimos pese a ser especialistas.

Pero, ¿por qué recibimos textos elaborados con IA sin declararlo?, ¿qué sentido tiene publicar algo que no es resultado de una investigación? El problema que se evidencia es la intención de publicar lo más posible para cumplir con las exigencias de instituciones que evalúan por cantidad y no por calidad. No se trata sólo de la intención de hacer trampa, sino de la presión que existe que posibilita la intención de recurrir a esas trampas. Hay un sistema que presiona a todos los niveles de la academia. Desde la formación en estudios de posgrado ya se comienza a exigir que se concurse a las publicaciones académicas de alto prestigio y dictaminación, muestra de ello es la inmensa cantidad de trabajos que llegan a la redacción de revistas que cuentan con buenas evaluaciones en los rankings o bases de datos, que reciben textos de finales de cursos que un profesor evalúa positivamente sin considerar que distan mucho de ser un texto académico resultado de investigación. Pero ahora se asume que con ayuda de la IA se puede mejorar y conseguir pasar los filtros de exigencia académica.

Requisitos como publicar en revistas científicas de alto prestigio para titularse, publicar en revistas de los niveles más altos en las bases de datos internacionales para alcanzar mayores puntajes y así demostrar la trascendencia de los resultados de investigación, publicar con mayor frecuencia a fin de demostrar que se está haciendo investigación, publicar lo más posible para construir un *currículum* más abultado que demuestre el prestigio que se requiere para concursar a un trabajo. Publicar es el requisito básico para poder acceder a una vida académica, pero se desdeña la calidad del texto, se trata de producir a destajo y para eso la IA es una herramienta incomparable.

Tampoco se trata de desdeñar la cantidad de artículos por sí misma, sabemos que escribir con frecuencia también hace que sea más probable que tengamos al menos un texto que a juicio de nosotros mismo, valga la pena en tanto logremos exponer mejor nuestras ideas, la práctica hace al maestro. No

tenemos porque pensar que todo lo que escribimos tiene que ser magnífico o aportar de manera extraordinaria al debate científico con cada texto, sin que por ello dejemos de intentar exponer cada vez mejor los resultados de nuestras inquietudes. El porcentaje de textos realmente paradigmáticos para la ciencia son apenas unos cuantos entre la inmensa producción de ciencia. Pero así funciona el debate y la producción de conocimiento académico y no hay porque exigirnos que sea necesario que absolutamente todo lo que publicamos sea realmente trascendente. Así es el debate, la exposición de ideas y es relevante para distintas etapas de la investigación de otros académicos o personas en formación. El problema de la producción a destajo y más ahora tratando de hacerlo con uso de IA tiene que ver con la intención de hacer trampa para engrosar el *curriculum*, no con si es pertinente el uso de la IA y poner en duda su capacidad como herramienta.

Ante esta voracidad por publicar que recurre a la IA para "fortalecer" los textos, las revistas requieren de estrategias básicas de protección y deslinde de responsabilidades. Las revistas concebidas como comunidades de investigación que debaten a través de la lectura de cada número que aparece, revistas que expresan redes de académicos, se citan, discuten a través de la dictaminación, de la revisión de los textos, convocan a temas comunes y se comprometen con la difusión, esas revistas requieren de estrategias mínimas de protección ante la posibilidad de abuso del uso de la IA.

Así, la política editorial precisa de exigir la declaración del uso de la IA, que los autores sean transparentes y declaren la forma en la cual recurrieron a la herramienta. No se trata de rechazar el uso de la IA, sino de incorporarla a la investigación con responsabilidad y transparencia, no sólo resolviendo la inmediatez y presión por publicar. Evidentemente, en este primer momento, hay una apelación a una postura ética y que tenga su propio lugar y forma de ser atendida. Ya en un segundo momento también se debe proteger a la comunidad que congrega una publicación, y una propuesta es deslindarse de la responsabilidad y exigir a cada autor que declare en su propio texto si usó la IA; y si sostiene que no la usó, también dejarlo explícito en la nota al pie junto con sus datos de investigador. Es una medida sencilla pero que protege el prestigio de quienes colaboran con las revistas.

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1342

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

¿Quién responde por la inteligencia artificial? Ética, sesgos y responsabilidad en la investigación y la escritura académica

Gezabel Guzmán*
Aliber Escobar**

La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) indica cuatro valores fundamentales que sientan las bases para que los sistemas de inteligencia artificial (en adelante, IA) trabajen por el bien de la humanidad. Estos son:

1. derechos humanos y dignidad humana,
2. vivir en sociedades pacíficas,
3. garantizar diversidad e inclusión,
4. florecimiento del medio ambiente así como de los ecosistemas.

* Profesora investigadora en la Academia de Comunicación y Cultura del Colegio de Humanidades y Ciencias Sociales de la UACM. Forma parte del Comité editorial de *Andamios. Revista de Investigación Social* y es miembro del Sistema Nacional de Investigadoras e Investigadores. ORCID: 0000-0001-9926-5401 Correo electrónico: gezabel.guzman@uacm.edu.mx

** Profesor investigador en la Academia de Expresión Oral y Escrita del Colegio de Humanidades y Ciencias Sociales de la UACM. Forma parte del Comité Editorial de *GOSOY. Revista del Centro Lacaniano de Investigación en Psicoanálisis* y es miembro del Sistema Nacional de Investigadoras e Investigadores. ORCID: 0000-0001-9163-0515 Correo electrónico: aliber.escobar@uacm.edu.mx

Específicamente en relación al enfoque de derechos humanos y dignidad humana, la ética[***] en el uso de la IA debe tomar en cuenta:

1. Proporcionalidad e inocuidad, ámbitos referidos al uso que no cause daños e injusticias, a nivel individual como comunitario, que puedan repercutir en el entorno social, cultural, económico, político, natural, es decir, de la vida en general.
2. Seguridad y protección, que no sólo se refiere a la protección tecnológica, sino también el fomentar buenas prácticas, por ejemplo, en no crear noticias falsas, suplantación de identidad, o generar información que perpetúe discursos de discriminación.
3. Derecho a la intimidad y protección de datos.
4. Gobernanza en colaboración de múltiples partes interesadas, para incluir diversidad de voces, tomando en cuenta sectores considerados vulnerables y promover el diálogo entre países, sectores tecnológicos, académicos, así como, sociales.
5. Responsabilidad y rendición de cuentas, para tener mecanismos de supervisión, evaluación y medición de amenazas al bienestar medioambiental.
6. Transparencia y explicabilidad, que debe revelar cómo funciona la IA, por ejemplo, qué datos utiliza, qué lógicas generales sigue.
7. Supervisión y validación para que los sistemas automatizados no operen sin control humano.
8. Sostenibilidad, para usar la tecnología en pro de resolver problemas ambientales y sociales; pero el término se refiere también a transparentar y promover la reflexión en el costo ambiental por el uso de la inteligencia ambiental.
9. Sensibilización y educación, enfocada, entre otros aspectos, a la alfabetización mediática.

[***] Antónimos de la palabra ética podrían ser inmoral, deshonesto, turbio. Mientras que sus sinónimos aluden a lo íntegro, decente, honrado. Pero ¿qué es la ética?, ¿qué implicaciones tiene ésta en el uso de la IA?, y ¿qué tiene que ver con el mundo de la escritura y la investigación académica? Para la Real Academia de la Lengua Española (2024), la ética es el conjunto de normas morales que rigen la conducta de la persona en cualquier ámbito de la vida. Esta definición nos sirve como marco en el presente artículo para distinguir la ética

10. Equidad y no discriminación, para promover el acceso de la IA desde ámbitos inclusivos y no violentos (UNESCO, 2024).

Sin embargo, la realidad muestra que esos valores se han quedado apenas en ideales, porque casi todos han sido ignorados, transgredidos o tergiversados de manera discrecional, engañosa u oculta, por las élites creadoras de las distintas IAs existentes. Como muestra claramente Karen Hao (2026) a partir de su investigación y las múltiples entrevistas que realizó a ejecutivos y empleados de *OpenAI, Microsoft, Anthropic, Meta, Google, DeepMind* y *Scale* durante los últimos siete años.

Por otra parte, como bien señalan Alberto Quintavalla y Jeroen Temperman, en un trabajo que han realizado para plantear a la IA cuestiones legales, éticas, tecnológicas y sociales para analizar tanto las áreas que actualmente permanecen oscuras, como el impacto, el daño y los riesgos de la IA en los derechos humanos, a pesar de que sus historias se solapan, ambos fenómenos –la IA y los derechos humanos– llevaron existencias bastante separadas durante sus primeros cincuenta o sesenta años, más o menos. Solo en la última década sus caminos han convergido, las dos fuerzas se han encontrado, se han apoyado mutuamente o, como también puede ocurrir, han entrado en conflicto, provocando enfrentamientos de menor o mayor envergadura (Quintavalla y Temperman, 2023).

Lo anterior también permite advertir que la IA no ha sido siempre usada de forma responsable, tomando en cuenta los requisitos establecidos legalmente tanto por las instancias internacionales como por los gobiernos. A partir del reconocimiento de lo anterior, en este artículo nos proponemos abordar dos problemas que desde los inicios de la IA permanecen irresolubles, pues como nos interesa analizar las posibilidades de la IA para investigar y escribir académicamente, debemos tener muy claro de quién es

de la moral. Sobre todo, porque como explica García-Clarck, citando a Velasco, la "ética es una reflexión crítica y racional sobre la moralidad vigente en una sociedad determinada" (2025, p. 5). Es decir, "la moral está relacionada con las conductas, normas y valores morales, vigentes en una sociedad, mientras que la ética se constituye por descripciones, explicaciones y evaluaciones sobre tales elementos del mundo moral" (García-Clarck, 2025, p. 6). Ambos conceptos se encuentran presentes de forma interrelacionada y se presentan de manera

la responsabilidad en el uso de la IA para la investigación y cómo lidiar con los sesgos inherentes a su utilización.

1. Inteligencia artificial: fundamentos, sesgos y responsabilidad ética en la investigación académica

Han pasado 70 años desde que el término "inteligencia artificial" (aparentemente, una *contradictio in adjecto*) fue acuñado por John McCarthy y Marvin Minsky, para nombrar una rama particular de las ciencias computacionales (Born, 2018). En este artículo vamos a partir de la definición un poco modificada por Rainer P. Born e Ilse Born-Lechleitner, de la que propuso Edward Feigenbaum en 1981, a saber:

> La IA es aquella parte de la informática que se ocupa del diseño de sistemas informáticos inteligentes, es decir, sistemas que muestran las características que asociamos con la inteligencia en el comportamiento humano –por ejemplo, la comprensión del lenguaje, el aprendizaje, el razonamiento, la resolución de problemas, etc.– (Born, 2018, p. VIII).

De hecho, es el cerebro humano, a partir del conocimiento de sus interconexiones neuronales en red, la base de inspiración para crear modelos o algoritmos (*machine learning* y *deep learning*) para hacer predicciones o tomar decisiones basadas en datos, lo cual permite a las computadoras hacer inferencias y aprender tareas sin que estén específicamente programadas para éstas (Stryker y Kavlakoglu, 2024; Copeland, 2026).

Actualmente, existen diferentes tipos de IA, por ejemplo, la "IA estrecha" empleada para realizar tareas específicas como asistentes de voz (Alexa,

constante en cualquier ámbito de la vida cotidiana, por ejemplo, ante la eutanasia, se evoca a la ética para que se aporten argumentos a favor o en contra de ésta con el fin de complejizar un acto que conlleva implicaciones sobre la vida y la muerte humana. Así, "el planteamiento teórico, valorativo y prescriptivo, que ofrece la ética, puede interpelar a los individuos o grupos involucrados en algún dilema moral y derivar en prácticas concretas" (García-Clarck, 2025, p. 7). Por tanto, la ética indudablemente desemboca en una acción, pero en una acción

Siri) y los chatbot de redes sociales usados para simular conversaciones humanas. O la "IA generativa" para crear contenido complejo, imágenes de calidad, video, audio, textos. Nos vamos a centrar en la generativa, porque lo que nos interesa es abordar la ética en la investigación y la elaboración de textos académicos, pero sobre todo la reflexión en torno a los sesgos presentes en la inteligencia artificial.

Sabemos que hoy la IA puede realizar la escritura de un artículo académico que asevere una serie de tesis muy polémicas, aunque, al no ser humana, todavía no puede asumir una posición moral sobre las mismas, ni tampoco puede ser obligada a responder en caso de que sean infamias. No obstante, a pesar de que no debe ser considerada "autora", en sentido humano, porque tradicionalmente la creación de una obra depende de un ejercicio intelectual, moral, original y singular de un ser humano (Penabad-Camacho, Morera-Castro y Penabad-Camacho, 2024), es verdad que, al tener ya décadas de entrenamiento humano, la IA hoy crea artículos, libros, piezas musicales, guiones de teatro, películas y series, audiovisuales, etc., tan semejantes a las grandes obras del acervo universal gracias a que ha sido alimentada con todas ellas, por lo que resulta muy difícil, sino imposible, distinguirlas de las que han sido realizadas por un humano.

Lo anterior es posible porque, si el aprendizaje automático de la IA se basa en el flujo de información que el propio ser humano introduce, es comprensible que sea capaz de reproducir fielmente los grandes logros civilizatorios que son ingresados sistemáticamente a su base de datos. Aunque, insistimos, no pueda hacerse cargo de las consecuencias de estos. Según Murray, el aprendizaje automático "es lo que permite que un buscador nos haga recomendaciones. Nos diga cómo llegar a una dirección o incluso nos permita traducir textos [...]. El aprendizaje automático, permite que las computadoras resuelvan los problemas mediante el reconocimiento de patrones en los datos" (2020, p. 143), que millones de seres humanos discriminan puntualmente.

humana, porque al igual que la moral es un fenómeno humano. Ante ello la IA no tiene ética, ni conciencia, es una herramienta que procesa datos proporcionados exclusivamente por los humanos que la programan y la utilizan. Por ello, podemos encontrar en la IA sesgos culturales y sociales completamente humanos, es decir, ésta es un reflejo de nosotros como empleadores y de ciertas sociedades, sobre todo, de aquellas dominantes a nivel global bajo el uso de determinado idioma, y de ciertos valores que respondan a la lógica del capital. Por su

Ahora bien, así como lo anterior permite entender cómo es que la IA puede replicar lo humano, también permite advertir que, a pesar de que es capaz de tener iniciativas provocadas por los algoritmos, al final sólo reproduce lo que obtiene de los humanos, sin que haya una deliberación o voluntad singular que la pueda hacer responsable de lo que dice o hace. Por tanto, lo que la IA reproduce son los valores y normas de las élites que la han creado. Mismos que, según Sam Altman, al haber sido cabildeados con ciertos gobiernos de todo el mundo, se imponen como imperativos morales que Bill Gates, Elon Musk, Dario Amodei, Ilya Sutskever, etc., buscan cumplir con objetivos falaces: cómo solucionar el cambio climático o eliminar la necesidad de trabajar, y para ello están arrasando con los recursos naturales y explotando a millones de personas. Es decir, son inmorales y además ocultan el verdadero objetivo del transhumanismo en el que, en un futuro no muy lejano, el mundo ya no dependerá de las personas. Dicho de otro modo, será un mundo amoral gobernado por la IA.

Ese futuro es el que, como bien señala Lorenzo Ramírez, Raymond Kurzweil proyecta en el año 2045:

> En el que espera que la inteligencia artificial (IA) haya adquirido el desarrollo necesario para alcanzar lo que denomina como la 'singularidad', en el momento en el que la tecnología se creará a sí misma e incluso se adentrará en la biología y la neurociencia prácticamente de forma autónoma, con el mínimo componente humano. (Ramírez, 2025, p. 324)

Por otro lado, hoy la inteligencia artificial es omnipresente en el ámbito académico. Y puede ser muy útil como una herramienta de asistencia técnica para traducir textos, corregir estilos, procesar datos, redactar borradores, corregir redacción, ortografía y generar gráficos. Así como para desempeñarse como un asistente de investigación que identifique información específica

parte, para Brugger (2005), la ética va más allá de una ciencia social descriptiva de lo moral que describe y analiza las valoraciones fácticas de otros. Más bien es una disciplina filosófica que emplea un método para reflexionar sobre los datos de la experiencia moral (conceptos y principios morales), y no consiste en un empirismo o en el recurso de un sentimiento no analizable, sino en el uso de la razón. Es decir, entre otros aspectos, permite generar una

en múltiples textos. O como herramienta de búsqueda para que autores encuentren rápidamente literatura sobre temas concretos, mientras realizan otras actividades. Asimismo, puede resumir artículos académicos en minutos y ayudar a que editores y revisores puedan mejorar la redacción, detectar similitudes, verificar la integridad de imágenes, comprobar la originalidad de un documento y muchas cosas más.

Sin embargo, hay revistas como *Science* y *Nature* que prohíben su uso; incluyendo el empleo de aplicaciones *Large Language Models* (LLM) para, entre otros aspectos, realizar traducción de textos. Aunque redes regionales como el Sistema Regional de Información en Línea para Revistas Científicas de América Larina, el Caribe, España y Portugal (Latindex) y *Scientific Electronic Library Online* (SciELO), se encuentran abiertas a su utilización bajo ciertos marcos, como la responsabilidad, las obligaciones y la transparencia de los autores y las autoras. Es decir, exigen la responsabilidad pública de autoría, y para ello requieren que todo el material citado se atribuya correctamente, así como indicar expresamente cualquier uso de IA, lo cual busca, en la medida de lo posible, controlar la transparencia y la integridad en la investigación.

En consecuencia, muchas revistas actualmente colocan políticas para el uso de la IA al momento de presentar un artículo, que incluyen declaración de uso y que promueven la no manipulación de datos, el no plagio y sobre todo, que reconocen el valor de la supervisión humana para asegurar la originalidad, la coherencia, la calidad y la identificación de sesgos. Lo que indica que la ética en la investigación que utiliza IA es un terreno espinoso y todavía parcialmente incógnito. De hecho, Michael James Heron y Pauline Helen Belford partieron de un libro de ficción publicado por Richard G. Epstein (1997) a finales del siglo pasado, sobre el caso de un robot asesino y los problemas éticos que ocasiona sobre la responsabilidad y la complicidad de las instancias involucradas, para abordar un estudio de caso de la ética académica en la IA.

reflexión razonada de las decisiones que toman las personas. Ante ello, el ser humano al ser un ser social, realiza acciones en el marco de la moral, dicho actuar debe ser razonado, consciente y conlleva responsabilidad en cada individuo. Así, la responsabilidad en el uso de la inteligencia artificial recae sobre la persona que la emplea, aunque para realizar un uso razonado de la IA es vital conocer cómo opera ésta y qué implicaciones conlleva en el

En su texto, aclaran una cuestión que nos parece fundamental cuando hablamos de la ética tecnológica, a saber, que "[l]a responsabilidad ética y moral respecto a la tecnología –cómo hablamos de ella y cómo la utilizamos– siempre reside en el componente humano de la ecuación" (Heron y Belford, 2025, p. 13). Dicho de otra forma, que mientras no se demuestre lo contrario, no debemos caer en la ilusión de que la IA puede llegar a ser independiente de sus creadores y responsable de "sus" errores. Pues, actualmente, la IA no es una "entidad" equivalente a un ser humano ni independiente del mismo. Prueba de ello es que, a pesar de que películas como *Mission: Impossible - The Final Reckoning* (McQuarrie, 2025), reproducen la ilusión de que llegará un momento en que la IA adquirirá total independencia del ser humano, la imagen que de ella han hecho sus propios programadores es de una tecnología que ha sido creada –por ellos mismos– "a su imagen y semejanza" (Hao, 2026, p. 35). Es decir, que la IA no se ha creado sola, ni tampoco puede volverse en contra de la humanidad por voluntad propia. Recordemos que el famoso computador inteligente llamado HAL 9000 en *2001: A Space Odyssey* (Kubrick, 1968) o la inteligente MOTHER en *Alien* (Scott, 1979), se negaban a obedecer las órdenes de los tripulantes de las naves porque habían sido programados previamente para una misión, pero no por voluntad propia.

Por consiguiente, no debemos confundirnos en relación con la instancia responsable en el trabajo de investigación. Si hoy alguno de nuestros asistentes de IA comete algún error o delito de derechos de autoría, quien debe pagar, además de sus creadores, son sus usuarios. Es decir, los investigadores. Pues, al ser una tecnología que carece de calidad moral, nosotros somos los responsables de cualquier imprecisión, plagio u omisión.

ámbito social, económico, cultural y ecológico. Por ejemplo, los microchips que alimentan la IA necesitan elementos de tierras raras, que a menudo se extraen de formas destructivas para el medio ambiente. Los centros de datos, incluidos los operados por proveedores de servicios en la nube, producen residuos que contienen mercurio y plomo que contaminan mantos acuíferos y entran en el subsuelo dañando poblaciones y ecosistemas. Sumado a ello, se emplea un uso desmedido de agua para enfriar componentes eléctricos con los cuales

2. La dimensión humana en la construcción de los sesgos en los sistemas de inteligencia artificial

De este modo, la IA que ha sido entrenada durante décadas para realizar investigación y coadyuvar a la escritura, comete los mismos errores con los cuales se le programó, sea por ignorancia, descuido, cansancio o prejuicios por parte de los alimentadores, o falta de pericia por parte de los usuarios y usuarias. De ahí que tenga los mismos sesgos humanos, tanto de programadores como de sus creadores. Muchas veces orgánicos (producto de quienes la han alimentado) y otras tantas intencionales; de acuerdo con la ideología dominante de sus creadores. Según Gabriela Arriagada-Bruneau, Claudia López y Marcelo Mendoza:

> El sesgo implica un error; en otras palabras, una desviación entre las observaciones y la realidad objetiva. Los sesgos también deben tener un componente sistemático; son errores sistemáticos y no meras incidencias aleatorias. Además, cuando hablamos de sesgo, también aludimos a una tendencia injustificada o irrelevante a favor o en contra de ciertas ideas o entidades frente a otras (2026, p. 48).

Lo anterior permite aclarar que, así como sucede en la vida real, en la IA los sesgos son ineludibles; por ejemplo, aquellos de clase, género y racialización. Así, los discursos que son generados por la IA "no provienen únicamente de una conciencia neutra, o desde una racionalidad del algoritmo totalmente pura. Se trata de un reflejo del propio ser social [...] que, desde luego, es desigual, patriarcal y colonial" (Sánchez y Guerrero, 2026, p. 183). En ese sentido, al ser alimentada por seres humanos, en diferentes condiciones materiales, en diversas situaciones psicológicas y con múltiples niveles edu-

opera la IA. Además, su uso genera una gran emisión de gases de efecto invernadero por la quema de combustibles fósiles, sin hablar del daño a los ecosistemas y a ciertas poblaciones que causa toda la producción tecnológica que requiere el funcionamiento de centros de datos (UNEP, 2025). Ante ello, es fundamental hablar de una ética social, la cual para Brugger (2005), alude a una ética política, una ética de la vida, que impacta en el ejercicio profesional como la medicina, el periodismo, los editores, la propia escritura y en el uso de

cativos, no es posible que la IA sea objetiva, ni mucho menos que carezca de errores, por ello, transmite creencias de la propia cultura que los produce. Así, la IA no sólo reproduce las ideas y las tendencias subjetivas de quienes la alimentan, sino que, al procesarlas y sistematizarlas, es muy fácil que se desvíe de "la realidad" que pretende acumular y representar para producir otra en su lugar. Una realidad sesgada.

Podemos identificar, al menos, tres contextos de origen de los sesgos: 1) social, 2) técnico y 3) cognitivo. Los sociales, "tienen su origen en los contextos sociales, culturales e institucionales en los que se generan y utilizan los datos. Estos sesgos reflejan y perpetúan los prejuicios y las desigualdades existentes en la sociedad" (Arriagada, *et al.*, 2026, p. 48). Es por lo anterior que si la IA utiliza gente con condiciones precarizadas de países como India, Venezuela o Kenia –entre muchos más–, se puede esperar el reflejo de cierta sociedad y una cultura determinada, con todo lo que eso implique. Sobre todo, porque "los sistemas de IA pueden replicar sesgos sociales existentes, pues la mayoría de ellos se construyen usando conjuntos de datos que reflejan el mundo real [...] imperfecto, injusto y discriminatorio" (Freres en Sánchez y Guerrero, 2026, p. 184). Así, el capitalismo digital reflejado en la IA es posible gracias a la explotación de ciertas poblaciones, por ello, muchos sesgos, que de ésta emanan sostienen la desigualdad.

Por su parte, los sesgos técnicos "se derivan de los métodos, procesos y herramientas específicas que se utilizan para desarrollar e implementar sistemas de IA. [...] están relacionados con las limitaciones algorítmicas y con las decisiones técnicas de los desarrolladores, como la selección de datos y el diseño de modelos" (Arriagada *et al.*, 2026, p. 48). Aquí, quienes realizan prácticas de explotación, al preocuparse sólo por sus intereses en avances tecnológicos, utilidades económicas y políticas, regularmente utilizan métodos bajo una lógica capitalista de despojo. De esta forma, según Deborah Raji, "para que se construyan sistemas de IA, a menudo existe un costo humano oculto" (Hao, 2026, p. 94). Lo que implica que los procesos que se siguen para en-

---

la IA. Donde, además, no puede separarse la ética individual y la social, entre ambas existe un nexo indisoluble, como lo existe también entre la persona y el colectivo, entre uno y el otro. Por ello, para Lévinas (en Lascevena y Fattah, 2023), la existencia comienza cuando nos vemos interpelados por el rostro del otro. Ahí, frente al otro, como realidad diferente, es donde se nos presenta la posibilidad del descubrimiento. Bajo esta postura, la ética está

trenarla no siempre son estandarizados, seguros, ni están transparentados, ya que las decisiones técnicas de dirigentes y creadores muchas veces no son consensuadas, declaradas, ni mucho menos, convenientes para el mundo.

Por su parte, a los sesgos cognitivos son:

> Errores humanos de juicio o razonamiento que pueden afectar al funcionamiento y a la implementación de los sistemas de IA. [...] Se producen cuando los sistemas de IA reflejan las limitaciones cognitivas y las heurísticas de sus desarrolladores y usuarios, cuando imitan los patrones de toma de decisiones humanos, o cuando los seres humanos toman decisiones metodológicas sesgadas que contribuyen a la existencia de sesgos técnicos, como el sesgo de selección. (Arriagada *et al.* 2026, p. 48)

Como se puede ver, este tipo de sesgo involucra tanto a creadores como a alimentadores, que es posible dividir en dos: aquellos de quienes ya hemos hablado, que son pagados y muchas veces explotados para ingresar y discriminar contenidos; y aquellos que no somos pagados, pero que alimentamos a la IA de manera gratuita e inadvertida. Por ejemplo, cuando resolvemos una prueba CAPTCHA, supuestamente diseñada para "demostrar" que somos humanos, seleccionando todas las imágenes que incluyen una bicicleta; cuando damos "me gusta" a ciertas imágenes en redes sociales, o cuando no nos despegamos de nuestro móvil ni siquiera mientras dormimos. Como bien advierte Daniel Estulin, "[a]cceder a tu cuerpo las 24 horas del día, los siete días de la semana, genera muchos datos" (2020, p. 17), mismos que resultan ser lo más valioso para la IA y por ello, los almacena –sin consultarnos– en base de macrodatos.

Con respecto a los creadores, este tipo de sesgos pueden ser problemas que desarrolladores producen al no programar correctamente para que la IA aprenda, lea, cree contenidos, o extraiga patrones adecuadamente o

sustentada en las relaciones humanas y amerita la responsabilidad absoluta por el otro. En ese sentido, Lévinas "considera que la ética es esencial para flexionar sobre la particularidad de cada ser, reflejada en el rostro, siendo precisamente esas peculiaridades las que los diferencia de los otros, haciéndolos únicos y elementales. Asimismo, la ética permite ir al encuentro del Otro" (en Lascevena y Fattah, 2023, p. 84). Por ello, para Lévinas, esta ética

incluso problemas causados porque programan a la IA para "robar" contenidos sin pagar derechos y sin declarar que lo están haciendo, para al final ser descubiertos y denunciados. Así mismo, se presentan cuando alimentadores o usuarios, seleccionamos de manera incorrecta un contenido y la IA lo reproduce automática e irreflexivamente.

Por todo lo anterior, los sesgos son tan ineludibles que resulta imprescindible reconocerlos como una parte esencial de la producción académica apoyada en IA. Muchos de éstos responden a nuestras propias ideas que consideramos certezas y que, por tanto, no cuestionamos. Sólo generando un (auto)pensamiento crítico y teniendo mucho cuidado es que podremos evitar cometer errores, plagios y confusiones que nos lleven a situaciones complicadas en las que nosotros seremos quienes debamos responder.

3. Complejizar la ética: hacia una comprensión conceptual y política de la inteligencia artificial

Entonces, si queremos utilizar la IA para nuestras investigaciones, "[h]ay que examinar en qué medida estas innovaciones tecnológicas interactúan con nuestras expectativas de autogobierno e igualdad, [pues] [a]ntes que normativo, el desafío al que nos enfrentamos es conceptual" (Innerarity, 2025, p. 24). Es decir, quizá haya que trascender la normativa meramente ética que encontramos en todos los reglamentos, declaraciones de valores, requisitos, exigencias, imperativos, etc., en el uso de la IA, porque, al desconocer la complejidad tecnológica de ésta, pareciese que lo único que nos exigen es ser responsables cuando hay muchas más cosas en juego que no estamos registrando. Por tal motivo, Daniel Innerarity (2025) introduce una aproximación política que lleva más allá el imperativo ético y evidencia el control privado de la información como lo realmente importante.

Así, desplaza la discusión del sesgo ético que demanda la rendición de cuentas en lugar de la discrecionalidad y que exige el respeto de los valores

de la experiencia humana se da en la proximidad con el otro y en su cuidado, donde el "otro" puede ser también el mundo natural con toda su vastedad de vida. Y la proximidad es entendida no sólo como cercanía o aproximación, sino como aquello que hace posible un acto de obligación, de colaboración, de justicia, sólo ahí puede ocurrir el acontecimiento ético. Al respecto, no podemos ser indiferentes al daño ecológico y social que causa el uso de la IA.

en lugar del abuso de poder, para advertirnos que, además de eso, es preciso hacer el examen crítico y la articulación entre lo fáctico y lo normativo que nos proporcionen un marco teórico capaz de permitirnos comprender aquellas cuestiones de lo tecnológico que rebasan el ámbito de lo humano, por ello, no pueden ser circunscritas a lo ético tal como lo hemos sostenido en el apartado uno. Si ya desde el principio de este artículo señalamos que la IA no tiene moral, por tanto, no conviene detenernos en discusiones de ese orden, ni mucho menos en tratar de juzgarla –como hacen en la película *Le procès du chien* (Dosch, 2024), en la que someten a juicio a un perro que muerde a una mujer–; al mismo tiempo, es urgente entender qué es lo que la rige, para poder intervenir adecuadamente. Lo anterior quiere decir que, en lugar de preocuparnos solamente por la transparencia en su uso y en tratar de disminuir los sesgos en el manejo de la IA (los cuales, cabe señalar, pueden reproducir todo tipo de discursos normativos, patriarcales, conservadores, raciales; discursos xenófogos, violentos, ultraconservadores, capitalistas depredadores), hay que construir condiciones teóricas para poder comprenderla y utilizarla de manera más responsable en nuestras investigaciones. Pues, independientemente de que sea cierto que la responsabilidad reside en el componente humano, en realidad, cuando usamos IA la ética se complejiza en nuestra relación con la tecnología. Dicho de otro modo, así como un sesgo representa un verdadero problema cuando no tenemos ni la advertencia ni los elementos para poder identificarlos y moderarlos, pareciese que en esta época la responsabilidad no es reductible a la mera utilización de un dispositivo, sino al conocimiento tanto de su funcionamiento como de las implicaciones que tiene.

Una vez convenido lo anterior, quizá lo que resulta más importante al momento de mirar el sesgo de la IA es entender que funciona de manera diferente de nuestro cerebro. Pues si la racionalidad algorítmica no es humana, eso significa que carece, todavía, de la posibilidad de ser reflexiva y ese trabajo sigue siendo exclusivamente humano. En consecuencia, podemos sostener

Por otra parte, siguiendo a Lévinas, la ética en la escritura de un artículo académico interpela al autor o autora en cada palabra, entre sus líneas está la posibilidad de un encuentro con quien lea el texto y, por tanto, con el otro. Así, es un acto de colaboración con la comunidad científica el compartir fuentes, datos, metodologías, resultados, y dar cuenta abiertamente del uso de la IA, porque, como explica Gómez-Vargas (2025), la inteligencia artificial no

que la información que genera la inteligencia artificial y que es utilizada en nuestras investigaciones está plagada de sesgos, por lo que, el reto no sólo es que las personas autoras declaren de forma transparente el empleo de la IA, o que aprendan a manejarla correctamente, sino que estén advertidas de todas las implicaciones que tiene el uso de ésta en sus investigaciones.

Pues al ser la IA una tecnología relativamente nueva y desconocida, hace falta todavía mucho trabajo para entender sus procesos, métodos, lógicas, así como, sus encrucijadas que nos permitan posicionarnos éticamente frente a los sesgos que produce o reproduce e identificar cómo se puede sostener y cuáles son las condiciones de posibilidad de que una IA pueda ser juzgada como ética. Pues, si es verdad lo que Innerariry advierte acerca de que "[l] a algoritmización requiere pensar muchas categorías socioculturales, como sujeto, acción, responsabilidad, conocimiento o trabajo" (2025, p. 29), también requiere cuestionarnos si pueden ser propias y posibles en una IA. Ese sería otro de los retos, pues pondría sobre la mesa si en verdad podemos confiar en la IA a pesar de su irreflexividad y su amoralidad, o si podemos crear una IA, no solo reflexiva, si no moral (es decir, humana, no artificial); como la del protagonista de la película *Robocop* (Verhoeven, 1987). Que, por cierto, no era totalmente artificial, sino medianamente humana y, por ello, capaz de limitar moralmente sus propios actos ser capaz de autogobernarse. Cabría recordar que la IA tiene una capacidad increíble de aprender de la propia "experiencia" humana que se le comparte, al grado de que ya puede fingir, así como también mentir. Tal como muestra ejemplarmente la película *The Artifice Girl* (Ritch, 2022).

Inconclusiones. La inteligencia artificial como campo rico de investigación

Llevar los valores fundamentales promovidos por la UNESCO (2024), respecto del uso de la IA en revistas académicas para orientar la produc-

puede recaer solo en prohibiciones, sino en abrir posibilidades de acción, reflexión y crítica, delineando otros posibles horizontes de pensamiento y ejecución. Pero, valdría investigar cómo aminorar el propio daño medio ambiental que causa su uso y así propiciar "un uso *virtuoso* de la inteligencia artificial, ulterior a la de una máquina que devenga sustituto de la inteligencia humana. De ahí que se demanda que la intervención de la IA funja como un

ción y difusión del conocimiento, amerita –como hemos revisado– varios aspectos. Primero, que autores declaren de manera transparente su uso. Como lo explican Penabad-Camacho, Morera-Castro y Penabad-Camacho, "la inteligencia artificial es una herramienta en expansión por lo que, en vez de bloquear sus posibles aplicaciones en el entorno de publicación científico-académica, se propone transparentar, normalizar y regular su uso" (2024, p. 5). Pero, como hemos argumentado en este artículo, es vital reconocer las implicaciones que tiene su empleo y complejizar la ética en nuestra relación con la tecnología.

Segundo, para aminorar los sesgos presentes en la IA, los procesos editoriales, además de realizar protección de datos de participantes, pueden promover la evaluación por pares basada en el respeto, la retroalimentación constructiva y la colaboración. Fomentando la participación de autores, revisores e integrantes de comités editoriales provenientes de distintas disciplinas, géneros, territorios y contextos culturales, promoviendo además la diversidad lingüística. Tercero, las publicaciones pueden priorizar investigaciones que contribuyan al bienestar social, fomentando el diálogo interdisciplinario, la representación de poblaciones diversas, incluyendo sectores considerados vulnerables e investigaciones que visibilizan problemáticas sociales y ambientales donde los aportes contribuyan a la injerencia social. Como explica Spinak, la introducción de la IA "ha cambiado el escenario de la redacción, creación y producción de trabajos de investigación generando una revolución del mundo editorial académico de manera irreversible" (2023, s/p). Ante esto, las personas que integran comités editoriales de revistas académicas tienen una responsabilidad ética y un compromiso social frente a los desafíos que amerita el uso de la IA.

No cabe duda de que la regulación humana es un reto, por ello, "la experiencia de uso de la IA debe estar acompañada, desde el ejercicio de cada persona, por las mejores prácticas, respeto a los derechos humanos y conocimiento ético, bioético, disciplinar y profesional" (Penabad-Camacho; Morera-Castro y Penabad-Camacho, 2024, p. 6). Además de tecnológico,

complemento –aditivo– de la inteligencia humana" (p. 206). Por ello, se requiere un manejo crítico y responsable de la inteligencia artificial. Aunque, el mundo de la IA no es neutro, en éste operan relaciones de poder "puesto que dichos dispositivos pertenecen a empresas que poseen principalmente intereses comerciales" (Gómez-Vargas, 2025, p. 200). Es decir, la IA responde al giro ideológico del gobierno actual en Estados Unidos de América y *Silicon Va-*

científico, algorítmico, matemático, político, económico, etc. Pero ¿cómo lograrlo tomando en cuenta los estándares de evaluación académica que además de ignorar muchas cuestiones relacionadas con la IA, priorizan la premura, la producción y la publicación de artículos, sostenidos en un tiempo acelerado, en revistas de alto impacto? y ¿cómo advertirnos sobre los sesgos que puedan (podamos) tener los propios autores?

Así que, más que establecer imperativos que no sabemos si la IA puede cumplir, porque no es humana, habría que aprender cómo funciona y cuáles son sus cualidades académicas, para saber en qué fases de la investigación utilizarla y en qué otras no. Pues si bien en muchas áreas la IA ya puede sustituirnos, es verdad que en la academia todavía no lo puede hacer y en la medida en que "no hemos encontrado el equilibro adecuado entre el control humano y los beneficios de la automatización" (Innerarity, 2025, p. 30), podemos fácilmente deslizarnos en un vértigo que nos dé la impresión de que estamos utilizando la IA de forma adecuada y ética, mientras que en realidad estamos abusando y engañando a los demás sobre nuestro trabajo. Pues hay que recordar que así como hace mucho tiempo, antes de la IA, había investigadores que publicaban libros con la suficiente profundidad y rigor teóricos, sin la necesidad de hacer uso de herramientas que les permitieran ahorrarse el trabajo, hoy hay millones de personas cuyo ímpetu de competencia, éxito y reconocimiento les ha llevado a publicar múltiples artículos y libros sin haber investigado ni trabajado lo suficiente y hay muchos más que, por la premura de nuestra época, están dispuestos a leerlos sin verificar sus trayectorias, sólo por la cantidad de "me gusta" que tienen en sus redes sociales y por el capital social del que se han hecho en esta época.

Las condiciones de la academia ya no son las mismas del siglo pasado. Ante ello, una población que lleva décadas acostumbrándose a obtener el conocimiento de una IA y del *mainstream*, puede creer que ya no necesita docentes, leer textos físicos, tener encuentros y discusiones cara a cara dentro de las aulas, construir el conocimiento de forma conjunta y reflexiva, o

lley, así como a guerras por la hegemonía cibernética, comercial y bélica del orden mundial donde los "*Big Tech* son el prestador de servicios de la soberanía del incipiente imperio [...] afincado en la misoginia, el desprecio por la vida y el racismo" (Emmelhainz, 2026, s/p), con varios fines, entre éstos, modelos de censura, manipulación política, campañas de influencia sobre la opinión pública, extractivismo intelectual. Así, el uso de la IA exige reconocer que

escribir un artículo científico, una nota periodística, un ensayo, un poema, un libro, o una tesis sin ayuda de la IA.

La investigación en el campo de la IA ha experimentado cambios significativos en cuanto a su enfoque, metodología y marcos teóricos desde sus inicios. Se trataba de "alcanzar una IA 'fuerte', es decir, una forma de inteligencia artificial que fuera prácticamente indistinguible de la inteligencia humana. Sin embargo, [...] el punto en el que nos encontramos podría describirse con mayor precisión como la creación de una serie de tecnologías que proporcionan una ilusión de inteligencia" (Natale, 2021, p. 3).

Finalmente, desde que se inauguró la línea de investigación sobre la IA, uno de los problemas principales fue conocido como "el problema del observador", concepto referido a "cómo reaccionan los seres humanos al presenciar cómo las máquinas muestran inteligencia" (Natale, 2021, p. 43), mismo que ya indicaba que ninguna inteligencia puede ser reconocida por sí misma, sino en función de una otredad. Por esa razón es que, más de medio siglo después, seguimos intentando no sólo comprender la manera en que las computadoras expresan su "inteligencia", sino intentando ayudarles a ser inteligentes; aunque no hemos logrado obtener una que rebase la ilusión e iguale a los humanos en todos los ámbitos, porque, por más sesgos humanos que le introduzcamos, la singularidad sigue siendo inimitable. Pues es verdad que, porque la IA responde y reacciona como si fuera una persona, hay mucha gente que cree que las operaciones de ésta ya son equivalentes a las de la inteligencia humana, pero lo humano es algo que va más allá, una condición singular e irrepetible, ampliamente mucho más compleja a que una máquina.

toda decisión sobre su empleo implica responsabilidades humanas. De ahí la importancia de desarrollar desde la escritura y la investigación académica una mirada crítica, cuestionar el extractivismo, fomentar una ética compleja preocupada por la colectividad, y sobre todo, una praxis por el cuidado de la vida en general, del mundo común, sólo así se podrá contribuir al fortalecimiento, y no al deterioro, de la vida compartida.

Fuentes consultadas

Arriagada, G., López, C. y Mendoza, M. (2026). *Ethics in Artificial Intelligence and Information Technologies*. CRC Press.

Born, R. (2018). *Artificial Intelligence: The Case Against*. Routledge Library Editions, Taylor and Francis.

Brugger, W. (2005). Ética. En *Diccionarios de Filosofía*. pp. 222-226. Herder.

Diccionario de la Lengua Española. (2024). *Ético, ca*. Disponible en: https://dle.rae.es/%C3%A9tico#H3yay0R

Dosch, L. (Dir.). (2024). *Le Procès du Chien*. Bande à Part Films y Atelier de Production.

Emmelhainz, I. (2026-07-05). México y la IA. En *La Jornada*. Disponible en: https://www.jornada.com.mx/noticia/2026/07/05/opinion/mexico-y-la-ia

Epstein, R. (1997). *The Case of the Killer Robot: Stories about the Professional, Ethical, and Societal Dimensions of Computing*. Wiley.

Estulin, D. (2020). *Metapolítica*. Penguin Random House Grupo Editorial, S.A.U.

García-Clarck, R. (2025). *La ética contemporánea y el problema de la violencia*. Torres Asociados.

Gómez-Vargas, C. (2025). ¿Schopenhauer como educador o la inteligencia artificial como educadora? La filosofía de la educación en el contexto de las inteligencias artificiales. En *Revista Latinoamericana de Filosofía de la Educación*. Vol. 12. Núm. 24. pp. 194-210. DOI: https://doi.org/10.63314/PAAV1231

Hao, K. (2026). *El Imperio de La IA*. Ariel.

Heron, M. y Belford, P. (2025). *A Case Study for Computer Ethics in Context: the Scandal in Academia*. CRC Press.

Innerarity, D. (2025). *Una teoría crítica de la inteligencia artificial*. Ensayo.

Kubrick, S. [dir]. (1968). 9000 en *2001: A Space Odyssey*. Metro Golwyn-Mayer (MGM).

Lascevena, M. y Fattah, J. (2023). La ética levinasiana: un encuentro de alteridad en los tiempos contemporáneos. En *Encuentros. Revista*

de Ciencias Humanas, Teoría Social y Pensamiento Crítico. Núm. 19. pp. 83-93. DOI: https://doi.org/10.5281/zenodo.8270641

Mc'Quarrie, C. (2025). *Mission: Impossible - The Final Reckoning*. Paramount Pictures, Skydance Media y TC Productions.

Murray, D. (2020). *La masa enfurecida. Cómo las políticas de identidad llevaron al mundo a la locura*. Titivillus

Natale, S. (2021). *Deceitful Media: Artificial Intelligence and Social Life after the Turing Test*. Oxford University Press.

Penabad-Camacho, L. Morera-Castro, M y Penabad-Camacho, M. (2024). Guía para uso y reporte de inteligencia artificial en revistas científico-académicas. En *Revista Electrónica Educare*. Núm. 28. pp. 1-41. Disponible en: https://www.revistas.una.ac.cr/index.php/EDUCARE/article/view/19830/31511

Quintavalla, A. y Temperman, J. (2023). *Artificial Intelligence and Human Rights*. Oxford University Press.

Ramírez, L. (2025). *El diablo está entre nosotros: los agentes del caos en el nuevo orden mundial*. La esfera de los libros.

Ritch, F. (Dir.). (2022). *The Artifice Girl*. Paper Street Pictures, Blood Oath y Last Resort Ideas.

Sánchez, S. y Guerrero, R. (2026). El algoritmo también tiene género. La acumulación originaria de datos y el patriarcado digital en América Latina. En *Tramas y Redes. Revista del Consejo Latinoamericano de Ciencias Sociales*. Núm. 10. pp. 181-196.

Schapire, A. (2021). *La traición progresista*. Península.

Scott, R. (Dir.). (1979). *Alien*. Brandywione Productions y Scott Free Productions.

Spinak, E. (2023-08-30). Inteligencia artificial y comunicación de investigaciones. En *ScieELO en Perspectiva*. Disponible en: https://blog.scielo.org/es/2023/08/30/inteligencia-artificial-y-comunicacion-de-investigaciones/

Stryker, C. y Kavlakoglu, E. (2024). ¿Qué es la IA? En *IBM*. Disponible en: https://www.ibm.com/mx-es/think/topics/artificial-intelligence

Programa de las Naciones Unidas para el Medio Ambiente (UNEP). (2025-11-13). *La IA plantea problemas ambientales. Esto es lo que el mundo puede hacer al respecto*. Disponible en: ONU

https://www.unep.org/es/noticias-y-reportajes/reportajes/la-ia-plantea-problemas-ambientales-esto-es-lo-que-el-mundo-puede

UNESCO. (2024). Ética de la inteligencia artificial. En *Inteligencia artificial y tecnologías emergentes*. Disponible en: https://www.unesco.org/es/artificial-intelligence/recommendation-ethics

Verhoeven, P. (Dir.). (1987). *Robocop*. Orion Pictures y Hawkeye Entertainment.

DOI: https://doi.org/10.29092/uacm.v23iEspecial.1350

Normas para la recepción de originales

Andamios. Revista de Investigación Social es una publicación periódica, científica, internacional, arbitrada, indexada y de acceso abierto y gratuito tanto para autores como para lectores ("revista diamante"), cuyo objetivo es contribuir en las tareas de investigación, difusión y enseñanza en materia de Ciencias Sociales y Humanidades. Emplea un sistema estricto de arbitraje en la modalidad "doble ciego" y quienes evalúan son personas de reconocido prestigio en el ámbito de la investigación, tanto nacional como extranjera. Se publica en lengua española y en formato electrónico tres veces al año y está adscrita al Colegio de Humanidades y Ciencias Sociales de la Universidad Autónoma de la Ciudad de México (UACM).

Andamios. Revista de Investigación Social está altamente posicionada en distintas categorías de revistas académicas. En México posee el Reconocimiento del Sistema Nacional de Publicaciones Científicas y Humanísticas (SNPCyH) de la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI). A nivel internacional, aparece catalogada dentro de Scopus y su SCImago Journal & Country Rank, en el nivel Q2 del área de Artes y Humanidades, y en el nivel Q3 del área de Ciencias Sociales. Figura, asimismo, en múltiples índices y bases de datos, entre los cuales destacan: Biblat, Dialnet, EBSCO y sus Academic Search Ultimate Magazines and Journals, Fuente Académica Premier Database Coverage List y Sociology Source Ultimate Database Coverage List; Hispanic American Periodicals Index (HAPI), International Bibliography of the Social Sciences (IBSS), Latindex (Catálogo 2.0), LatinRev, Matriz de Información para el Análisis de Revistas (MIAR), Political Database of the Americas (PDBA), Redalyc, Scielo-México, USC Libraries. University of Southern California; y Web of Science (WoS) y sus Social Sciences Citation Index (SSCI) y Journal Citation Reports (JCR), entre otros. En su versión electrónica tiene presencia a través del portal institucional de la Universidad Autónoma de la Ciudad de México (UACM) y el Open Journal Systems (OJS).

1. Secciones de la revista

Toda colaboración debe indicar con claridad para cuál de las tres secciones se propone (si va destinada a un *dossier* debe señalar con precisión a cuál); y atenderá, sin excepción, todos los Lineamientos editoriales (ver, a continuación, el numeral 5):

1.1. *Dossier*. Trabajos de investigación que se ajustan a la convocatoria de la sección monográfica coordinada por destacados especialistas en ciencias sociales y humanidades. Este apartado cambia de número en número, de acuerdo con el tema específico objeto de reflexión y, a juicio de la o las personas que coordinan ese número, podrá incluir hasta 9 artículos, una entrevista, una traducción y una bibliografía especializada en el tema abordado.

1.2. *Artículos*. Trabajos de naturaleza teórica o aplicada que contribuyan a difundir los temas, problemas y discusiones en cualquier área de las humanidades y las ciencias sociales. Esta sección se encuentra permanentemente abierta pero, en caso de que el artículo se considere publicable se enlistará entre los que han sido considerados de la misma forma con anterioridad, por lo que se indicará a su debido tiempo el número donde aparecerá el texto dentro de la revista. Podrá incluir hasta 10 artículos por número.

1.3. *Reseñas*. Comentarios críticos de alguna novedad bibliográfica de carácter académico, significativa para el mundo de las humanidades y las ciencias sociales, sea nacional o extranjera. Al tratarse de un texto crítico se espera que argumente tanto las cualidades como las posibles desventajas del trabajo que comenta, a la luz de otros estudios sobre la especialidad. Versará sobre libros publicados, como máximo, el año inmediato anterior, o sobre nuevas ediciones de clásicos que resulta oportuno revisar a la luz de problemáticas vigentes.

En los dos primeros casos pueden ostentar la autoría hasta tres personas. En el tercer caso sólo una.

2. Arbitraje

2.1. El Comité Editorial remitirá a la autora o autor el acuse de recibo de su trabajo en un plazo no mayor a un mes. En el plazo de un año, que podría ampliarse en circunstancias extraordinarias, le remitirá la resolución final sobre el mismo.

2.2. Fase 1 del arbitraje

Primero, toda contribución será sometida a revisión con hasta tres programas de detección de similitudes e IA, siempre bajo supervisión humana, con la finalidad de evitar plagios y autoplagios. Si los hubiere, el artículo no podrá continuar en el proceso de arbitraje. Cabe precisar que, una vez obtenidos los informes ofrecidos por los programas antedichos, el material revisado se eliminará de ahí con la finalidad de proteger los derechos de autor.

2.2.1. Política de uso de IA

En aras de mantener los principios de integridad, transparencia y responsabilidad autoral, *Andamios. Revista de Investigación Social* no admite artículos ni reseñas cuya autoría o coautoría se acredite a alguna herramienta de Large Language Model (LLM) de inteligencia artificial generativa (IA), ya sean propiamente modelos, aplicaciones de IA, asistentes o programas que incorporan IA para búsqueda y revisión bibliográfica; para lectura, escritura, traducción o corrección de documentos; para análisis cualitativo o cuantitativo; para generación de imágenes y esquemas, o propias de las humanidades digitales (por ejemplo: ASReview, Canva, ChartGen, ChatGPT, Copilot, Claude, Consensus, Deep Seek, DeepL, Elicit, Excelmatic, Gemini, Google Translate, Grammarly, GraphAI, Grok, Jasper, Jenni AI, Julius AI, JSTOR AI Research Tool, LanguageTool, Litmaps, MAXQDA AI

Assist, PaLM-E, Paperpal, Powerdrill AI, Rayyan, Research Rabbit, Scholarcy, Scite Assistant, SciSpace, Tableau, Transkribus, Trinka AI, Writesonic, entre otras).

Ahora bien, si alguna autora o autor emplea una o más de esas herramientas —o similares— debe dejar constancia clara y explícita de ello *en la primera nota a pie de página* dentro de su artículo o reseña, indicando qué usó, por qué y en qué parte del proceso de investigación o escritura aplicó contenido generado por la IA. De igual manera, en caso de no haber usado alguna de ésas u otras herramientas, debe declararlo *en la primera nota a pie de página*; recomendamos esta leyenda: "Declaro que no he usado ninguna herramienta de inteligencia artificial para la realización de este artículo". Si durante la verificación mencionada en el ítem precedente (2.2) se identifica uso de IA no declarado o no justificado en el espacio destinado para ello, el artículo será sancionado con su retiro del proceso de revisión y la autora o autor con la prohibición de enviar futuras contribuciones a *Andamios. Revista de Investigación Social*.

2.3. Fase 2 del arbitraje

Habiendo solventado la etapa previa, el artículo pasará por una fase de revisión "doble ciego" donde el Comité Editorial determinará si corresponde con las áreas temáticas atendidas por la revista y cumple con los requisitos académicos indispensables en un artículo científico. A saber:

Desarrollar temas originales, aportar nuevos enfoques o proporcionar información reciente sobre temas ya estudiados.

Mostrar consistencia, rigor y coherencia metodológica.

Desarrollar una temática relevante dentro de las disciplinas sociales o humanísticas.

Respetar los valores de claridad, concisión, corrección gramatical y orden lógico en la composición general del texto.

Mostrar un uso crítico de fuentes actualizadas y adecuadas según el área de conocimiento.

2.4. Fase 3 del arbitraje

De cumplir con lo antedicho, las contribuciones serán sometidas al dictamen de dos especialistas en la materia, en la modalidad "doble ciego". (Cuando la autora o autor forme parte del personal de la UACM su trabajo necesariamente será dictaminado por árbitros externos a esta institución. En caso contrario, el arbitraje podrá ser interno o externo a la UACM.) Si existe contradicción entre ambos dictámenes, se procederá a una tercera evaluación, que se considerará definitiva. El proceso de dictaminación será secreto y en ningún caso se dará información nominal sobre éste. El formato de dictamen empleado por *Andamios. Revista de Investigación Social* puede consultarse en la página *web* de la revista.

2.5. Una vez emitidas las evaluaciones, se enviará a los autores el acta de dictamen y tendrán un plazo no mayor a tres semanas para entregar la versión final

del artículo con las correcciones pertinentes. El Equipo de Redacción de la revista cotejará la versión final con los dictámenes para corroborar la actualización y la corrección del documento, e informará a los autores en qué número de la revista será publicado su trabajo.

2.6. Sin excepción, las colaboraciones aceptadas se someterán a corrección de estilo, siempre basada en la norma lingüística de la lengua española.

2.7. La publicación estará sujeta a la disponibilidad de espacio en cada número. En los casos donde el número de colaboraciones dictaminadas favorablemente para la sección *Dossier* exceda el espacio destinado a la misma, el Comité Editorial se reserva el derecho de determinar cuáles de ellas se ajustan más a los lineamientos establecidos en la convocatoria.

2.8. En todas las fases descritas la decisión será inapelable.

3. Políticas generales

3.1. Los trabajos enviados a *Andamios. Revista de Investigación Social* deberán ser investigaciones de alta calidad académica en materia de ciencias sociales y humanidades.

3.2. Para ser consideradas y sometidas a dictaminación, las contribuciones deberán ser originales y cumplir con todas las normas editoriales.

3.3. El envío o entrega de un trabajo a esta revista compromete a su autora o autor a no someterlo simultáneamente a la consideración de otras publicaciones. De hacerlo, esa falta de probidad académica impedirá que se admitan ése o cualquier otro artículo de la misma persona, incluso en los casos de autoría compartida.

3.4. El envío de cualquier colaboración a la revista implica la aceptación de lo establecido en estas normas editoriales y la autorización al Comité Editorial de *Andamios. Revista de Investigación Social* para que difunda su trabajo en la revista, en la página electrónica de la misma, en reimpresiones, en colecciones y en cualquier base de datos o sistema de indización impreso o electrónico que se decida para lograr una mayor y mejor difusión del mismo. En la versión electrónica, las contribuciones serán publicadas en formato PDF, ePub, XML SPS, XML Marcalyc, algunos MP3 y HTML compatible con marcaje especializado en la interpretación de *Lectores de pantalla con voz* (Screen Readers) convencionales como: NonVisual Desktop Access (NVDA), Jaws for windows, a fin de que la revista pueda ser consultada en versiones de audio.

3.5. Esta revista está bajo una licencia Creative Commons México 2.5. Está permitida la reproducción y difusión de sus contenidos para fines educativos o de investigación, sin ánimo de lucro, siempre y cuando éstos no se mutilen y se cite la procedencia (*Andamios. Revista de Investigación Social*) y a la autora o autor.

3.6. Los derechos patrimoniales de los artículos publicados en *Andamios. Revista de Investigación Social* son cedidos por cada autora o autor a la UACM. Sin embargo, tal como lo establece la ley, cada autora o autor conserva sus derechos morales.

3.7. Las y los autores podrán usar el material de su artículo en otros trabajos o libros, con la condición de citar a *Andamios. Revista de Investigación Social* como la fuente original de los textos.

3.8. Los artículos contenidos en esta publicación son responsabilidad de sus autoras o autores y no comprometen la posición oficial de *Andamios. Revista de Investigación Social*, ni la de la UACM.

4. Datos de la autora o autor

Siempre en un archivo aparte del texto y *nunca dentro de éste*, las contribuciones anexarán los datos completos de la persona que ostenta la autoría (su institución de adscripción, dirección postal, dirección electrónica, teléfono y ORCID), y una breve reseña curricular (estudios profesionales, máximo grado académico, nombramiento en su institución de adscripción, líneas de investigación y principales publicaciones —en el caso de los libros se mencionará el título, el lugar de edición, la editorial y el año—). Cuando se trate de coautorías, se incluirán los datos de todas las personas que colaboran en el artículo, pero sólo una fungirá como "autor de correspondencia". Tal como se indicó con anterioridad, los artículos no podrán ir firmados por más de tres personas, en tanto que las reseñas sólo podrán ostentar un autor o autora.

5. Lineamientos editoriales

5.1 Sólo pasarán a dictamen los trabajos que cumplan cabalmente las normas de recepción de originales, políticas generales y lineamientos editoriales, especificados en *Andamios. Revista de Investigación Social*.

5.2 Las contribuciones, rigurosamente inéditas y en formato *Word* anonimizado, se presentarán con letra Arial de 12 puntos, con interlineado doble y páginas numeradas, con márgenes superior e inferior de 2.5 cm, e izquierdo y derecho de 3 cm.

5.3 La escritura (gramática y sintaxis) deberá atender siempre la norma lingüística de la lengua española. No hay excepciones en este punto.

5.4 Las contribuciones para las secciones *Dossier* y *Artículos* deberán tener como mínimo 5,000 palabras y como máximo 10,000, incluyendo notas, gráficas, tablas, citas, bibliografía, así como el título, los resúmenes en español e inglés y las palabras clave; las correspondientes a *Reseñas* no podrán superar las 3,000 palabras.

5.5 Se recomienda que el título de los trabajos no exceda de 70 caracteres, incluyendo los espacios.

5.6 Las contribuciones, salvo las reseñas, deberán anexar: a) un resumen en español y otro en inglés, que no superen, cada uno, las 120 palabras; deben destacar las principales aportaciones y conclusiones del artículo, así como su originalidad, valor y relevancia disciplinar; y b) cinco palabras clave, en español e inglés, que identifiquen el contenido del texto. Los metadatos con los cuales se etiquetan los archivos de las distintas versiones electrónicas son generados a partir de estas palabras clave y títulos tanto en español como en inglés.

5.7 Cada párrafo, a partir del 2º, comenzará con sangría.

5.8 El aparato crítico se presentará de acuerdo con el formato utilizado por la American Psychological Association (APA). Entre otros sitios, puede consultarse en la dirección electrónica siguiente: https://www.bidi.unam.mx/pdfs/3_Normas-APA-7-ed-2019-11-6.pdf

ENVÍO DE COLABORACIONES

Toda contribución deberá enviarse exclusivamente
como archivo adjunto vía electrónica al correo:
andamios@uacm.edu.mx

Andamios. Revista de investigación Social
del Colegio de Humanidades y Ciencias Sociales
de la Universidad Autónoma de la Ciudad de México
fue impresa en octubre de 2026 en los talleres de
Arte en Diseño y Diagramación Digital, S.A. de C.V.
Calle Bolivar, núm. 455-L, Col. Obrera, Alcaldía Cuauhtémoc,
C.P. 06800, Ciudad de México, México. Con un tiraje de 300 ejemplares.

Andamios, Revista de Investigación Social
del Colegio de Humanidades y Ciencias Sociales Universidad
Autónoma de la Ciudad de México.