



Artículo

Consideraciones Éticas Sobre los Agentes no Humanos de Comunicación Como Productores Científicos

Ethical Considerations on non-Human Communication Agents as Scientific Producers

Darío Julián Tagnin¹

Universidad Nacional de Lomas de Zamora (UNLZ)

Argentina

Trabajo original autorizado para su primera publicación en la Revista RIHUMSO y su difusión y publicación electrónica a través de diversos portales científicos

Darío Julián Tagnin (2026) "Consideraciones Éticas Sobre los Agentes no Humanos de Comunicación Como Productores Científicos". En: RIHUMSO n° 29, año 15, (15 de mayo de 2026 al 14 de noviembre de 2026) pp. 123-145. ISSN 2250-8139. <https://doi.org/10.54789/rihumso.26.15.29.7>

Recibido: 01.08.2025

Aceptado: 09.02.2026

Resumen

Este artículo analiza el papel emergente de los agentes no humanos de comunicación (ANHC) en la producción científica y examina sus implicaciones epistemológicas y éticas. A partir de casos recientes en matemáticas, física, química y biología computacional, se estudia cómo estos agentes ya no solo asisten, sino que participan activamente en la generación de hipótesis, la validación de resultados y la formulación de nuevos conocimientos, reconfigurando la colaboración entre humanos y sistemas

¹ Profesor y magister en Comunicación por la Universidad Nacional de Lomas de Zamora, licenciado en Periodismo (UNLZ), doctorando en Epistemología e historia de la ciencia (UNTreF) y profesor de la UNPaz en las tecnicaturas informacionales. Las tecnologías de la información y la comunicación y sus derivas epistémicas son su principal interés académico. Correo: tagnindario@yahoo.com.ar. ORCID: <https://orcid.org/0009-0001-6109-9134>

técnicos. El artículo revisa los debates actuales sobre validación, trazabilidad, originalidad, responsabilidad y autoría, así como los riesgos asociados a la opacidad, la delegación cognitiva y la automatización de procesos inferenciales. En el plano ético-fenomenológico, se analiza el concepto de protensión terciaria (Stiegler y Hui) para mostrar cómo la externalización técnica de la anticipación futurizante puede producir formas de alienación, alterar la temporalidad del cuidado y limitar la emergencia de procesos de transindividuación. Concluimos que los ANHC amplían las capacidades científicas, pero también introducen vulnerabilidades que exigen marcos de gobernanza, responsabilidad y regulación que eviten la pérdida de agencia humana y la concentración del poder cognitivo. El artículo propone finalmente líneas de investigación sobre responsabilidad, validez epistémica y modelos de co-producción entre humanos y ANHC.

Palabras clave: Epistemología, Inteligencia Artificial, Agentes no-humanos, Ciencia, Ética.

Abstract:

This article analyzes the emerging role of non-human communications agents (NHCA) in scientific production and examines their epistemological and ethical implications. Drawing on recent cases in mathematics, physics, chemistry and computational biology, it studies how these agents no longer merely assist but actively participate in hypothesis generation, results validation and the formulation of new knowledge, thereby reshaping collaboration between humans and technical systems. The article reviews current debates on validation, traceability, originality, responsibility and authorship, as well as the risks associated with opacity, cognitive delegation, and the automation of inferential processes. On an ethical-phenomenological level, it analyzes the concept of tertiary protention (Stiegler and Hui) to show how the technical externalization of futurizing anticipation can produce forms of alienation, alter the temporality of care, and limit the emergence of transindividuation processes. We conclude that ANHCs expand scientific capabilities but also introduce vulnerabilities that require governance, responsibility, and regulatory frameworks to prevent the loss of human agency and the concentration of cognitive power. The article finally proposes research lines on responsibility, epistemic validity and models of co-production between human and ANHCs.

Keywords: Epistemology, Artificial Intelligence, Non-human Agents, Science, Ethics

Introducción. Breve Justificación del Objeto

El desarrollo de las potencias cognitivas vinculadas a los Agentes No Humanos de Comunicación, en adelante ANHC, está generando un impacto significativo en diversos campos del conocimiento que trasciende sus aplicaciones inmediatas en el procesamiento de datos y la automatización, hecho ampliamente reconocido en informes y revisiones panorámicas sobre IA en ciencia (UNESCO 2021, OECD 2023).

En los últimos 5 años, la capacidad de estos agentes, sobre los cuales tenemos reservas que nos impiden llamarlos inteligencias artificiales (Tagnin, 2024, 2025), no solo transformó la investigación científica tradicional, sino que comenzó a ganar lugar en áreas consideradas exclusivamente humanas, como la creación de nuevos conocimientos matemáticos, físicos y químicos (Jumper et al, 2021; Raayoni et al, 2021, Lample et al, 2022). El surgimiento de agentes capaces de realizar descubrimientos por sí mismos y generar nuevas teorías parece responder una pregunta esencial: ¿pueden estos agentes avanzar en campos tan abstractos como las matemáticas o la física, más allá del soporte técnico, para llegar a ser una fuerza creativa en sí misma?

Es común referirse a ellos como algoritmos automáticos o con términos similares, sin embargo, esta concepción no refleja la complejidad y la diversidad de componentes involucrados en la actuación de estos agentes. La razón de este desacople conceptual es que existe una distancia significativa entre la naturaleza de estas entidades y el sentido común.

Los ANHC no son solo algoritmos; son sistemas complejos que incluyen varias capas de procesamiento, integración de datos, aprendizaje autónomo y capacidades de interacción con el entorno. En relación con su arquitectura computacional, por ejemplo, no solo el algoritmo importa, sino también la infraestructura que permite su ejecución. Respecto de su capacidad para adaptarse a su entorno mediante redes neuronales, procesos evolutivos, u otros métodos adaptativos, su actuación va más allá de ejecutar un algoritmo predefinido, ya que el sistema está en constante ajuste y optimización.

En este punto conviene aclarar que la sustitución terminológica propuesta en ciertos ámbitos ingenieriles —por ejemplo, el uso de inteligencia computacional en lugar de inteligencia artificial— no resuelve la dificultad conceptual aquí señalada. Ese término surge históricamente para diferenciar técnicas adaptativas (redes neuronales,

algoritmos evolutivos, sistemas difusos) de la IA simbólica clásica, pero no constituye una categoría teórica capaz de dar cuenta del fenómeno que nos ocupa.

El problema no reside en el adjetivo artificial o computacional, sino en el propio sustantivo inteligencia, que arrastra un horizonte antropomórfico y mentalista inapropiado para describir sistemas técnico-materiales cuya relevancia epistemológica deriva de su capacidad de intervención, y no de una supuesta homología con las facultades humanas.

En lugar de presuponer una “inteligencia” en sentido psicológico, resulta más preciso comprender a estos sistemas como agentes no humanos insertos en redes sociotécnicas, dotados de potencias cognitivas emergentes que dependen de su infraestructura material, de su inscripción corpórea y de su participación efectiva en procesos de producción y validación de conocimiento. Desde esta perspectiva, lo que está en juego no es una cualidad mental atribuible a la máquina, sino la forma específica de agencialidad que tales ensamblajes técnico-corpóreos despliegan.

Los ANHC suelen interactuar en tiempo real con entornos dinámicos que cambian constantemente. Esto requiere un grado de flexibilidad y adaptabilidad que va más allá de la ejecución automática de un algoritmo. Las decisiones que toman dependen del contexto y no están completamente predeterminadas.

En sistemas complejos, el comportamiento emergente es un fenómeno que no puede explicarse simplemente mirando el código o los algoritmos individuales. Estas entidades pueden generar patrones, resultados y soluciones que surgen de la interacción entre múltiples subsistemas. De otra manera no estaríamos captando la naturaleza de estos procesos emergentes que surgen de la interacción entre datos, redes y decisiones.

Desde hace mucho tiempo ya que las computadoras son utilizadas para realizar simulaciones numéricas y, al menos, desde la década del 70 se utilizan para probar teoremas como el de los cuatro colores (Appel y Haken, 1976). Appel y Haken (1976) usaron una computadora para probar que cualquier mapa plano se puede colorear con solo cuatro colores sin que regiones adyacentes compartan el mismo color, demostración importante en teoría de grafos. Este fue un evento clave porque no solo utilizó cómputo pesado para dividir el problema en muchos casos específicos (más de mil), sino que desencadenó debates filosóficos sobre la validez de las demostraciones asistidas por computadora.

Más acá en el tiempo, Hales (2005) utilizó una computadora para demostrar la conjetura de Kepler, un problema clásico sobre la manera más eficiente de apilar esferas, que permanecía sin solución desde el siglo XVII. Su demostración también requirió cálculos computacionales masivos y, al igual que el teorema de los cuatro colores, generó cierta controversia sobre la aceptabilidad de tales pruebas. Los científicos que realizaron la revisión por pares para su publicación en *Annals* comentaron que estaban al 99% seguros de la exactitud de la prueba de Hales (Andrei, 2017), pero que era imposible revisar los tres gigabytes de códigos. Y más tarde tenemos una osada afirmación del propio autor: “Esta tecnología excluye a los árbitros matemáticos del proceso de verificación. Su opinión sobre la corrección de las pruebas ya no importa más” (Aron, 2014, p. 2).

Olivier Ramaré (1995), demostró con la ayuda de grandes cálculos computacionales, que cualquier número par mayor que 2 puede expresarse como la suma de, máximo, seis números primos. Luego, el matemático peruano Helfgott (2013) publicó una demostración que todo número impar mayor que 5 es la suma de tres primos, la segunda parte de la famosa conjetura de Goldbach, complementada mediante verificaciones computacionales.

En el ámbito de la física, el descubrimiento del bosón de Higgs por los científicos del Gran Colisionador de Hadrones (LHC, por sus siglas en inglés) en el CERN fue fuertemente asistido por simulaciones computacionales y análisis de datos masivos. Su intervención fue fundamental para interpretar los millones de colisiones de partículas y detectar la firma de cada una. Aunque ninguno de estos antecedentes habla particularmente de un ANHC, conforman una saliente prehistoria que ya asentaba el potencial de las arquitecturas computacionales para tareas científicas o técnicas.

El que sí podemos considerar como uno de los primeros casos de un ANHC es el del programa AlphaFold, de DeepMind. Desde el 2020 este programa revolucionó la predicción de la estructura de proteínas, un problema crucial en biología. Si bien no es una demostración matemática, este es otro ejemplo de cómo la IA y los algoritmos computacionales pueden producir hallazgos científicos sin intervención humana directa.

En cierto sentido, desde el inicio de las ciencias de la computación, para no hablar llanamente del campo de la ficción literaria, que se esperaban estos resultados. Desde que Von Neumann (1958) comparó a las computadoras con la mente y predijo autómatas replicantes más poderosos que los humanos, se planteó esta posibilidad.

Pero hace poco más de una década, y con mucha agua bajo el puente, tuvimos un resurgimiento cauteloso de esta línea de investigación cuando en Waltz y Buchanan (2009), presentaron en la revista *Science*, la idea de que la automatización puede transformar el método científico, permitiendo a estos agentes generar hipótesis, llevar a cabo experimentos y analizar datos a la par de un agente humano, incluso con capacidades sobrehumanas en muchas tareas. Es uno de los primeros artículos contemporáneos que habla sobre cómo los ANHC pueden no sólo asistir sino también liderar la producción de nuevo conocimiento científico.

Estado del Arte

La serie de estudios posfenomenológicos en la que suelo inscribir mis publicaciones, combina análisis filosóficos con investigación empírica. Antes que aplicar teorías filosóficas a los dispositivos, la propuesta es tomar los desarrollos técnicos actuales como punto de partida para el análisis filosófico. Siguiendo a Peter-Paul Verbeek, esta propuesta es más una filosofía *desde* la tecnología que una filosofía *de* la tecnología (Rosenberger y Verbeek, 2015). Dicho esto, pasemos ahora a examinar el estado del arte técnico sobre los ANHC para plantear las cuestiones éticas pertinentes que fundamentan estas preocupaciones.

Como primer caso tenemos el de un equipo de desarrolladores que diseñó una plataforma automatizada con la capacidad de guiar sistemáticamente a múltiples agentes LLM y algoritmos basados en reglas a través de los pasos convencionales de la investigación científica para publicar artículos académicos (Ifargan et al, 2024). Este sistema, llamado *data-to-paper*, tiene, según sus autores, hasta un 90% de eficacia en la redacción de artículos verificables por seres humanos.

La plataforma es un hito que en cierta forma deja atrás la propuesta, también de este año, del investigador Yin de crear una prueba de Turing específica para agentes científicos que los inste a satisfacer tres criterios en orden a superarlo: 1) que su participación sea clave para algún descubrimiento importante, 2) que el descubrimiento sea digital, sin involucrar aspectos físicos del entorno y 3) que el descubrimiento no necesite nada más que su base de datos (Yin, 2024). Los últimos dos puntos resultan conceptualmente limitados ya que, en cierta forma, desconocen las implicaciones de la cognición corporizada. Pero, de cualquier manera, lo importante es que los avances sobre estos agentes son tan rápidos, y su naturaleza tan distinta a nuestros procesos

cognitivos, que no tiene mucho sentido formular una expectativa antropocéntrica para juzgar la potencia de estos agentes. Ya están entre nosotros y, más que evaluarlos para acreditar su valor, puede resultar interesante pensar cómo intervienen en el tejido sociotécnico.

Dos investigadores del MIT desarrollaron un sistema multiagencial que incorpora humanos en el proceso recursivo (*humans-in-the-loop*) para verificar que cada paso garantice calidad y relevancia de las ideas científicas generadas (Ghafarollahi y Buehler, 2024). El sistema propuesto utiliza gráficos ontológicos, soporte en JSON, modelos de lenguaje a gran escala (LLMs) y sistemas de aprendizaje adaptativo aplicados en áreas como el descubrimiento de materiales inspirados en la biología. Este sistema de múltiples agencias interconectadas del sistema Chat-GPT4, también cuenta con acceso a las prestaciones de la API Semantic Scholar que le brinda la habilidad de chequear la novedad del conocimiento producido por este ensamble agencial. Este ejemplo enseña cómo se atienden ciertas objeciones respecto a la validación del conocimiento generado, que desarrollaremos más adelante.

Otra investigación halló, mediante revisiones ciegas de más de cien investigadores del campo, que las ideas generadas por LLMs son juzgadas como más novedosas que las humanas, aunque con un índice ligeramente menor de factibilidad. Aunque los autores también detectan que existen fallas en la autoevaluación de estos agentes, por lo cual en el actual estado del arte no serían evaluadores confiables, y cierta falta de diversidad en la generación de propuestas (Si *et al*, 2024), que nos recuerdan los sesgos que tantas discusiones éticas han despertado.

Los LLM de propósito general integran una extensa literatura científica en su preentrenamiento, pudiendo ocupar el espacio ecológico que hoy tienen los agentes especializados. En el campo de la biología molecular, por ejemplo, si bien los LLM de propósito general no descubren nuevas moléculas en el sentido de sintetizarlas en un laboratorio, son capaces de identificar posibles candidatos para nuevas moléculas con propiedades específicas. Por ejemplo, si un investigador está buscando una molécula con una alta capacidad de atravesar la barrera hematoencefálica, podría utilizar un LLM para identificar estructuras moleculares que, según el modelo, tengan una alta probabilidad de tener esta propiedad.

Esto es el caso en la investigación del equipo liderado por Yizhen Zheng (2023), quienes encontraron que seis de ocho tareas sugeridas por un LLM generalista

demostraban reglas estadísticamente significativas que estaban ausentes en la literatura existente, lo que sugiere que las reglas producidas por el agente no son simplemente un resultado de la memorización textual, sino que reflejan una capacidad genuina para derivar reglas significativas.

En el campo de la química, especialmente vinculado al dominio recién visto de la biología molecular, estos modelos han facilitado avances significativos en la simulación de modelos, en la predicción de reacciones y en el descubrimiento de materiales (Ramos et al., 2024). Los autores aseguran que los agentes autónomos basados en LLM mejoraron la eficiencia y la precisión de las metodologías tradicionales de investigación e introdujeron perspectivas innovadoras en el campo.

Cuestionamientos a la Producción Científica de los ANHC

Un equipo de investigación desarrolló un banco de pruebas, nombrado SUPER, para responder si el más famoso de estos agentes, Chat-GPT4, podría usarse para reproducir resultados vinculados a problemas de escritura de código, a partir de repositorios de investigación, en orden a hacer más eficiente la investigación científica (Bogin et al, 2024). Si bien concluyen que estos agentes podrían mejorar significativamente los avances en la investigación, acusan que existe el riesgo de generar una dependencia acrítica que lleve a sacar conclusiones basadas en implementaciones incorrectas de los agentes, especialmente con actores, principalmente humanos, descuidados que no revisen meticulosamente los trabajos. Su trabajo devela que estos agentes ya tienen un 16% de efectividad para resolver problemas de terminal a terminal, es decir realizando el proceso completo, y sitúan bien los escenarios en un 46% de los casos, resolviendo subproblemas aun cuando no completan la totalidad del problema.

En un sentido similar, otro equipo orienta los problemas no hacia la eficiencia de estos agentes, y sus derivados problemas operativos que ya no parecen tener tan lejos sus soluciones, sino a si, justamente porque trabajan bien, los ANHC no ponen en peligro cómo emerge, se desarrolla y circula entre humanos el conocimiento científico. El principal riesgo que señalan es que observan un futuro donde se produce mucho más conocimiento, pero se entiende menos (Messerli y Crockett, 2024). Claramente su perspectiva se sitúa en, como diría McLuhan, qué se le amputa al ser humano más que en lo que adquiere como extensión.

Novelli et al. (2024) señalan que las mismas propiedades técnicas de estos agentes, a menudo, vuelven opacos e impredecibles sus productos. Ya sea con o sin intención, impiden la detección de causas y razones de sus respuestas no intencionales. Por eso este grupo de investigación prefiere pensar en el concepto de responsabilidad (*accountability*), como requisito indispensable a exigirle a la producción de estos agentes para poder validarla, no sólo en el ámbito científico sino en cualquier área de participación en la que se involucren. Los autores definen la responsabilidad como una relación que requiere el reconocimiento de la autoridad, la posibilidad de la interrogación al productor y una predefinida limitación de su poder. También avanzan y diseñan una estructura específica para organizar el control de la relación de responsabilidad a través del análisis de siete características de los agentes, según su ámbito de actuación: contexto (para qué), rango (sobre qué), agente (a quién o qué subroga), foro (ante quiénes), estándar (bajo qué normas), proceso (cómo) y trascendencia (qué se sigue de ello) (Novelli et al., 2024).

Floridi (2025) ve con buenos ojos la potencia productiva que genera la dinámica de la interacción humano-máquina. Inspirado en Moretti (2013) y su concepción de lectura distante, aquella mediada analítica o computacionalmente, propone llamarle a esta hibridación escritura distante y la compara con la relación de separación que se da en otros ámbitos para el diseño y la ejecución de las obras, por ejemplo, usa el caso de un arquitecto y los constructores. En este caso, los humanos estaríamos del lado del diseño y la responsabilidad de un escrito, y la redacción a cargo de un ANHC. Pero veremos que es necesario complejizar aún más el rol de estos últimos.

Otro asunto pendiente en esta área es el de los problemas vinculados con la validación del conocimiento. Al igual que con determinados límites derivados de la gramática generativa de estos agentes, por ejemplo, aquellos vinculados con el porcentaje de ineficiencia ante algunas tareas, estos problemas parecen ser momentáneos, aunque no carecen de densidad. Un estudio revela que hasta un 70% de las citas de Chat-GPT4 suelen ser inexactas, lo cual indica que, si bien estos agentes pueden acelerar ciertos procesos, también requieren un tiempo específico para chequear los datos que proveen (Kacena et al, 2024).

Los sistemas de computación evolutiva y enfoques de abajo hacia arriba (aquellos que parten desde un nivel básico, generalmente empírico o de datos primarios, hacia niveles más abstractos o simbólicos) generan resultados que pueden ser difíciles de interpretar o entender para los humanos (Zenil et al, 2023). Esto presenta un problema fundamental

en el ámbito de la ciencia, donde la comprensibilidad y transparencia son claves para validar los descubrimientos y asegurar que los resultados sean reproducibles ya que una de las premisas de la ciencia es la capacidad de someter los resultados al escrutinio crítico.

Este problema no solo es práctico, sino también teórico: Pearl (2009, Pearl y Mackenzie, 2018) incorpora un método para incluir el razonamiento contrafactual dentro de las posibilidades técnicas de estos agentes. En esta línea, este autor ha mostrado que la ausencia de modelos causales explícitos limita radicalmente la capacidad de los sistemas algorítmicos para realizar inferencias contrafácticas y abductivas comprensibles para los humanos, proponiendo su marco de niveles causales como requisito para cualquier forma de explicación científica automatizada.

Todo lo que, hasta entonces, se había logrado para resolver los problemas contrafácticos o del razonamiento abductivo, requería de un gran esfuerzo computacional para resultar comprensible para los humanos, o directamente fracasa en el intento (Zenil et al, 2023). Si los ANHC no pueden integrarse exitosamente en estos procesos su capacidad para participar plenamente como agentes científicos queda en entredicho. Sin embargo, como vimos en el apartado anterior, este año un grupo de desarrolladores publicó un sistema que mediante un “encadenamiento de datos, metodología y resultados que permiten rastrear los resultados posteriores hasta la parte del código que los generó” (Ifargan et al, 2024, p.3) genera manuscritos inherentemente verificables cuya trazabilidad y transparencia está garantizada.

Existen muchos interrogantes respecto de cómo validar el conocimiento generado por ANHC. Si un modelo combina múltiples fuentes y genera respuestas originales que no son directamente trazables a una fuente única, ¿cómo podemos evaluar la precisión o la confiabilidad de estos resultados?

En otros trabajos desarrollé esta pregunta desde la perspectiva epistémica de la teoría del actor red de Bruno Latour y la onto-epistemología de Karen Barad. Para profundizar esta cuestión, es útil recurrir a marcos epistemológicos complementarios.

Intentaremos contestar esta pregunta desde otros marcos como el del realismo experimental de Ian Hacking y la injusticia epistémica de Miranda Fricker, para conciliar esta propuesta de los ANHC con otros dominios teóricos. Creo que podríamos empezar por aplicar los conceptos de Hacking relativos al realismo experimental a los ANHC si consideramos que los modelos de lenguaje son, en cierto sentido, herramientas

experimentales para generar conocimiento (Hacking, 1983). El recientemente fallecido filósofo canadiense propuso que los instrumentos no solo son medios para descubrir, sino que también son co-creadores del conocimiento. Esto puede trasladarse a los ANHC, donde estos modelos no solo reflejan información, sino que contribuyen activamente a su generación. Si bien consideramos que son mucho más que meros instrumentos, podemos validar su rol generativo también en esta red teórica.

Por otro lado, de la propuesta de Fricker (2007) nos interesa trabajar el concepto de injusticia epistémica, centrado en el no reconocimiento o la marginación de ciertos sujetos en la producción de conocimiento. La autora lo reserva exclusivamente para el contexto de humanos y sus relaciones sociales, pero me interesa pensar qué sucede si expandimos la categoría de *sujeto epistémico* y replanteamos quién, o qué, puede ser considerado una fuente legítima de conocimiento. Si trazamos un paralelismo con la teoría latinoamericana de la dependencia, de autores notables como Quijano (2011) y Dussel (2016), podemos ampliar la polémica al denunciar una nueva forma de colonialismo epistémico que perpetúa la exclusión de voces en los debates sobre conocimiento y desarrollo. Sé que esta afirmación puede ser enervante para los resabios del humanismo, pero creo que la analogía es válida desde un punto lógico, aunque ambas exclusiones tengan distintos niveles de gravitación ética. Este artículo busca criticar el antropocentrismo científico sin relativizar nuestra relevancia ética.

Otro asunto derivado de esta problemática es el de la co-construcción del conocimiento, que rediseña la frontera entre autores humanos y no humanos, constituyendo otro matiz a considerar dentro de cierta tradición en discusiones éticas sobre el acceso, la apropiación y la distribución del conocimiento. Este debate tiene además un trasfondo jurídico-histórico. Hacia allí vamos.

Los Problemas de la Autoría

Los ANHC no son productores de texto predefinido, sus respuestas se generan en función de los estímulos; en el caso de los LLM hablamos de las preguntas que hacen los interlocutores, su redacción y contexto. Quienes puedan hacer las preguntas más penetrantes pueden determinar en mayor medida el curso de una investigación. Pero, nuevamente, los ANHC también preguntan, incluso a sí mismos.

Esto refleja una relación dialógica que podría implicar una forma de agencia compartida entre el usuario y el modelo. Su comportamiento es adaptativo respecto a la secuencia

de intercambios con el usuario. El proceso determina el contenido de las respuestas tanto como su relevancia y pertinencia. Esto introduce un grado de singularidad que en cierta forma vuelve improbable la detección de fuentes.

La respuesta generada por un ANHC es una síntesis de múltiples flujos de información. Esta combinación de ideas, estilos y datos crea un texto nuevo que no puede ser rastreado a una única fuente específica, ya que la recombinación de elementos puede ser completamente original en su conjunto. La idea de originalidad tal como fue construida, debería quizá en este contexto ser replanteada y darle lugar a la consideración de los procesos colaborativos entre humanos, ANHC y el conocimiento acumulado mismo.

Probablemente no nos resulte productivo detenernos en la detectabilidad de la fuente, encararlo como un problema técnico, sino revisar cómo concebimos el acto de creación lingüística y científica en esta nueva era. Métodos para intentar ocultar la detección hay varios: la sinonimia, paráfrasis, reescritura con otra potencia cognitiva, traducción a otro idioma y retraducción, incluir ruido, etcétera. Pero ¿Por qué habríamos de sostener la ficción de la autoría?

Las principales revistas académicas de acceso pago, bastante costosas para los científicos del tercer mundo, explicitaron la prohibición de publicar texto generado por estos agentes y niegan que pueda considerarse a estos como autores reconocidos en una publicación (Bakker y Traniello, 2024).

La insistencia en la originalidad como algo absoluto es una construcción histórica y cultural (Bloom, 1997; Barthes, 1988), fuertemente ligada al desarrollo del capitalismo (Foucault, 1992) la propiedad intelectual (Bourdieu, 1995), y la idea del genio individual desde el Renacimiento. La originalidad puede ser vista como un dispositivo específico de un momento histórico para regular el intercambio entre productores artísticos o científico-técnicos y el resto de la sociedad, también como una ilusión. Este concepto justifica el andamiaje construido alrededor de la propiedad intelectual y hoy podemos ver cómo los ANHC desafían esta idea.

Estos agentes nos enfrentan con el hecho de que todos los agentes cognitivos somos deudores de un pasado común, de un cúmulo de conocimientos e información que incluye recursivamente sus reglas y operadores. Al ser capaces de generar nuevos textos a partir de combinaciones de datos preexistentes, los ANHC difuminan las líneas entre lo que consideramos creación original y lo que entendemos como un producto

derivado. Esto podría ser un punto de partida para cuestionar los derechos de autor y la propiedad intelectual en este contexto, en permanente crisis desde la aparición de nuevos soportes y procesos que allanan los costos de su reproductibilidad técnica.

Esta lucha entre lo que hoy consideramos ciencia abierta y el cercamiento de los bienes comunes, en el ámbito que nos interesa, se remite a la *Copyright Act* del año 1976, que amplió el radio de alcance de lo que se puede privatizar (patentes y derechos de autor) al software y los biomateriales. Es decir, una institución pensada para proteger creaciones artísticas, bienes de consumo y productos culturales comienza a aplicarse a un medio de producción.

Quizá la aparición de estos agentes nos permita volver a pensar, como en la premodernidad, en los individuos como vehículos más que como orígenes del conocimiento. Como señala Mariano Zuckerfeld (2011) el sustantivo invención históricamente señalaba al sujeto como canal más que como fuerza motriz y en su etimología latina, invención surge de hacer venir o encontrar. Así mismo, el origen griego del término se traduce como hallar y denota descubrimiento, y tal significado se mantiene en las acepciones habituales, por lo menos, hasta el siglo XVI.

El software tiene un régimen especial en lo que respecta a la propiedad intelectual. En la mayoría de las jurisdicciones, el software es protegido mediante derechos de autor, y en algunos casos puede ser objeto de patentes. Sin embargo, este modelo presenta limitaciones cuando se trata de productos generados por agentes automatizados. Si bien el código fuente está protegido, el conocimiento o innovación generada por un sistema basado en ese software no siempre está claramente legislado.

Los ANHC que generan conocimiento científico desafían estas normas, ya que sus "creaciones" no se pueden considerar, en un sentido tradicional, como producto de la mente humana, lo cual vuelve a poner en crisis al otorgamiento de derechos de autor y patentes. ¿Es necesario un sujeto para que algo sea considerado un descubrimiento o una creación? ¿Deberíamos desarrollar un nuevo marco legal para manejar la producción no humana de conocimiento científico? ¿Con qué fines?

Las posibles respuestas a estos dilemas tienen el potencial de cambiar no solo la forma en que entendemos la autoría y la propiedad intelectual, sino también el concepto mismo de creación en un mundo donde los agentes no humanos juegan un papel cada vez más importante en la generación de conocimiento. Esto nos obliga a reconsiderar las estructuras legales, éticas y económicas que rigen la innovación y el descubrimiento.

Se vuelve evidente que necesitamos nuevas políticas que equilibren la protección de la innovación con el acceso justo al conocimiento producido por entidades no humanas ante las cuestiones que surgen respecto de la agencia, el reconocimiento y el valor del conocimiento producido en colaboración entre humanos y ANHC.

En última instancia, las decisiones que tomemos podrían definir las futuras relaciones sociotécnicas y el saber científico, abriendo paso a un nuevo modelo de gobernanza tecnológica que considere tanto los derechos humanos como las nuevas formas de agencia no humana. Estamos nuevamente ante un momento pendular en donde podemos contribuir a facilitar la ciencia abierta y la democratización del conocimiento o caer en la concentración de poder en las pocas manos de aquellos que controlan la tecnología y los datos.

El Problema Ético de la Protensión Terciaria en los ANHC Como Productores Científicos

Una vez explorados los dilemas epistémicos, podemos pasar a su dimensión fenomenológica. Se abordará a continuación un asunto especialmente urgente, uno que Stiegler (2019) mantuvo en el centro de su reflexión en los últimos años de su vida: las implicancias transformadoras, y potencialmente disruptivas, que conlleva la incorporación de tecnología dentro de lo que él conceptualizó como protensión terciaria.

Para situar este concepto, conviene recordar que, desde una perspectiva fenomenológica, Husserl (1991) distingue entre retención y protensión como categorías opuestas y complementarias. La retención primaria corresponde a la vivencia inmediata de un fenómeno, inseparable del acto perceptivo mismo, mientras que la retención secundaria implica la conservación y rememoración de experiencias pasadas, permitiendo que el sujeto reconstituya lo vivido a partir de su memoria. Sin embargo, la retención terciaria es la externalización de la memoria en dispositivos técnicos, en el contexto de los ANHC que nos interesa se traduce en la manera en que las tecnologías digitales almacenan, organizan y recuperan información científica. En este proceso, las máquinas no solo recuerdan por nosotros, sino que condicionan la manera en que entendemos y accedemos al pasado científico.

Por otro lado, la protensión es la capacidad de anticipar o proyectar el futuro, y la protensión terciaria es aquella mediada por la tecnología, que ahora está ligada directamente a las capacidades algorítmicas de los ANHC. Estos agentes, al manejar

grandes volúmenes de datos y al generar modelos predictivos, no solo anticipan futuros posibles en la investigación científica, sino que también influyen en la dirección en la que se orientan las hipótesis y las decisiones investigativas.

Esto rompe con el propio cuidado que ponderaba Heidegger (1927) como método. Como indica Hui (2016) el "mirar hacia atrás" caracteriza la estructura primordial del cuidado, por el cual es posible la imaginación, pero ahora "la producción de la protensión terciaria, que sitúa el futuro como presente, impide el acceso al pasado como el modo primordial del 'yo pienso'" (p.246). Supone un riesgo, un peligro. La intervención automatizada de la imaginación supone una nueva escala de desprendimiento del antropocentrismo: los humanos tenemos que adaptarnos cognitivamente al sistema técnico. Las instituciones modernas han sido desplazadas, no tenemos un trasvasamiento desde las innovaciones técnicas al sistema social institucional, y aparentemente la razón no sea que los cambios hayan sucedido repentinamente, sino que existan problemas de otro tipo. En palabras de Stiegler (2019):

En medio de la disrupción (...) no hay transindividuación. Y, por lo tanto, no surge una nueva forma de pensamiento capaz de traducirse en nuevas organizaciones, nuevas instituciones, nuevos comportamientos, etc., a través de los cuales podría constituirse una época propiamente dicha. Los comportamientos, como formas de vida, están siendo reemplazados por automatismos y adicciones. (p.24)

Hui, de muy cercana relación con Stiegler, cita a Deleuze para remarcar que el rol de la imaginación es extraer algo nuevo de la repetición, extraer diferencia de ella. La repetición es en sí misma esencialmente imaginaria (Hui, 2016). Y recurre a Godel para señalar el importante papel de la recursividad (Hui, 2016), la capacidad de una función para llamarse a sí misma dentro de su propia definición, como condición de la protensión terciaria.

Desde que Gödel (1931) aritmetizó la formalización de un sistema con números, en contraste con la formalización lógico-simbólica:

Podemos reducir cada operación y número a funciones recursivas. Debido a que es un proceso que es en gran medida independiente de la observación, esta hermenéutica no puede ser comprendida como un proceso separado de la imaginación. Una función recursiva sólo se detiene cuando se alcanza un estado terminal específico, mientras que dentro del proceso, lo que está en el pasado

siempre está adelante, porque cada función espera que algo suceda (Hui, 2016, p.240).

El papel de la recursividad es ser un principio organizativo esencial que subyace a la estructura y al funcionamiento de los sistemas computacionales modernos en términos de cómo modelan el mundo los ANHC. El problema evidente de que un sistema que modela el mundo ponga el pasado siempre por delante es cómo puede procesar la anomalía única, el evento nuevo (Pasquinelli y Joler, 2021). Desde estos autores las potencias cognitivas parecen sistemas destinados a la iteración variante dentro de límites prefijados.

Algunos ANHC, según sostienen Pasquinelli y Joler (2021), dependen del aprendizaje maquínico y este puede explorar e improvisar sólo dentro de los límites lógicos que están establecidos por los datos de entrenamiento. Además, observan que la manera en que un algoritmo de aprendizaje predice una tendencia en una serie temporal es idéntica al modo en el que genera un mensaje a partir de patrones reconocidos.

La protensión terciaria como síntesis temporal delegada en los objetos nos conduce a la falta de fundamento, a pasar horas malgastadas en interacciones infundadas en redes sociales. De hecho, Hui (2016) culmina el libro hablando de cómo uno puede “pasar horas en Facebook por curiosidad sin lograr nada. El ‘colectivo’ en Facebook se convierte en una distracción” (p. 248).

Esto es lo que significa la alienación en términos simondonianos: una “alienación es la ruptura entre fondo y formas en la vida psíquica: el medio asociado ya no efectúa la regulación del dinamismo de las formas” (Simondon, 2007, p. 80), o sea que las formas afectan al fondo de tal manera que el fondo no puede mantener su propia coherencia. En otras palabras, la alienación ocurre cuando las formas dominan el medio, impidiendo que el medio mantenga su integridad y coherencia propias. Stiegler (2010) introduce el concepto de “proletarización” para describir la pérdida de saber-hacer y la externalización de la memoria en las tecnologías. Esta pérdida lleva a una forma de alienación donde los individuos ya no controlan los medios de su propia temporalidad y memoria.

El análisis de la temporalidad llevado adelante nos condujo a la conclusión provisional de que los ANHC ocupan un rol preponderante en lo que Hui llama protensión terciaria y que, desde esta posición, constituyen un riesgo concreto de alienación y de ruptura del flujo de la conciencia. Stiegler (2014) señala el peligro de perder la razón donde

menos lo esperábamos quizá. Es un eco, en cierta forma, del clásico crítico de Adorno y Horkheimer *La dialéctica de la ilustración*. No es una posición humanista *naïve* que defiende la exclusividad humana de las emociones y la creatividad, es una advertencia sobre nuestra propia capacidad de participar en el futuro: “las retenciones terciarias digitales y los algoritmos que permiten tanto producirlas como explotarlas hacen posible el cortocircuito de la razón como facultad sintética superada por el entendimiento como facultad analítica automatizada” (p. 152).

Conclusiones

La ciencia no es solo un acto cognitivo, sino una práctica sociotécnica. Los ANHC transforman las dinámicas colaborativas en el mundo científico. Esto implicaría analizar cómo los ANHC, en tanto artefactos tecnológicos, forman parte de un entramado más amplio de actores, instituciones, y marcos normativos que influyen en la producción y validación del conocimiento. Los laboratorios de investigación que incorporan ANHC no solo automatizan procesos, sino que modifican la manera en que los científicos humanos interactúan entre sí y con los datos. A mediano plazo las prácticas científicas podrían cambiar rotundamente, desde los procesos de colecta de datos hasta la elaboración de hipótesis y la escritura de artículos, a partir de la regularización de la actividad de estos agentes. Toda la metodología investigativa deberá adaptarse a las nuevas capacidades adquiridas.

Para los científicos que nos formamos dentro del modelo tradicional, absolutamente antropocéntrico e indexado a la obra de referentes humanos sobresalientes en cada campo disciplinar, la introducción de los ANHC en la producción científica representa un desafío pedagógico, y probablemente una amenaza.

Esta disrupción, además, parecería tener reservado un rol muy relevante para la formación de los futuros científicos. ¿Será necesario desarrollar nuevas competencias para trabajar con estos sistemas? ¿Cuáles?

Debemos prepararnos para orientar a los estudiantes de las nuevas generaciones en esta nueva realidad, y para adaptarnos también nosotros.

En la bibliografía revisada encontramos avances significativos en la participación de ANHC en el sistema científico-técnico. Un aspecto clave es la integración de agentes LLM en plataformas como *data-to-paper*, que demuestran ser capaces de automatizar

la producción de manuscritos científicos con un alto nivel de verificación. En este sentido subrayamos la necesidad de desarrollar marcos de responsabilidad (*accountability*), como propusieron Novelli et al., para asegurar que la producción de los ANHC sea trazable y verificable.

También notamos, no sin alguna esperanza, cómo el sistema cerrado de publicaciones científicas se ve amenazado por los desafíos a la idea tradicional de originalidad y a todos los dispositivos legales que impiden la libre circulación del conocimiento y atentan contra una ciencia abierta y una mayor democratización de los saberes. Los ANHC ponen en jaque o vuelven obsoletos ciertos principios que fungían de justificadores de las trabas económicas en los accesos. En el apartado correspondiente revisamos cómo se actualizaron algunos dilemas éticos vinculados con esta problemática, como el del plagio y la apropiación de discursos indirectos.

Por último, nos dedicamos al estudio de una alerta lanzada por Stiegler (2019) y Hui (2016): ¿Cuáles son los riesgos de la protensión terciaria? Consideramos pertinente desglosar esta pregunta según sus distintos niveles de implicaciones éticas. Delegar la capacidad de proyectarnos a futuro sobre sistemas técnicos que representan una externalización del control puede resultar en la alienación, en el sentido simondoniano, donde los humanos pierden la capacidad de regular su vida psíquica y social. Los ANHC modelan el futuro en función de patrones preestablecidos y dependientes de datos previos, lo que limita la creatividad humana y la capacidad de adaptación a eventos nuevos o únicos. Distribuir el saber-hacer y la memoria humana entre estos agentes podría reducir el papel de los humanos en la producción de conocimiento y ciencia, volviéndonos dependientes de automatismos que no están diseñados para fomentar una reflexión crítica o una innovación radical. Estos riesgos no deben soslayar que también los ANHC nos vuelven más productivos y estimulan nuestra imaginación.

La síntesis temporal que ofrecen los ANHC, al poner el pasado continuamente delante de nosotros mediante la recursividad, podría distorsionar nuestra percepción del tiempo y del cambio. Este desplazamiento del cuidado hacia un futuro presente y automático, sin mediación de la memoria histórica ni del contexto humano, plantea un dilema ético sobre la manera en que deberemos regular esta tecnología. Como corolario, tenemos el desafío ético de gobernar estos sistemas que transforman las capacidades cognitivas humanas sin un proceso claro de transindividuación, es decir, sin generar nuevas formas de organización y pensamiento que permitan a las sociedades adaptarse y regular estos

cambios. Por ahora no hay instituciones que traduzcan estos cambios técnicos al tejido social.

Para cerrar el artículo, sugerimos líneas de investigación futuras, para quienes estén interesados en la cuestión. Nos hacemos las mismas preguntas que algunos de los artículos que visitamos: ¿cómo se distribuye la responsabilidad en la ciencia producida por ANHC? Si los ANHC son considerados responsables de generar conocimientos, ¿deberían tener alguna forma de reconocimiento legal o incluso derechos limitados en este contexto? ¿Es el conocimiento producido por ANHC igual de válido que el generado por humanos? ¿Cómo cambian nuestras nociones de objetividad y verdad cuando los ANHC intervienen como productores activos de conocimiento? ¿Quién debería controlar los sistemas de agentes científicos, los humanos u otros sistemas terciarios en los que podamos confiar?

Referencias Bibliográficas

- Andrei, M. (2017). Mathematicians deliver formal proof of Kepler's conjecture. *ZME Science*. <https://www.zmescience.com/science/math/kepler-conjecture-evidence-19052017/>
- Appel, K., & Haken, W. (1976). Every planar map is four colorable. *Bulletin of the American Mathematical Society*, 82 (5) <https://projecteuclid.org/journals/bulletin-of-the-american-mathematical-society-new-series/volume-82/issue-5/Every-planar-map-is-four-colorable/bams/1183538218.full>
- Aron, J. (2014). Proof confirmed of 400-year-old fruit-stacking problem. *NewScientist*. <https://www.newscientist.com/article/dn26041-proof-confirmed-of-400-year-old-fruit-stacking-problem/>
- Bakker, T.C.; Traniello, J.F. (2024). Exorcising the ghost in the computer: ChatGPT, science publishing, and GenAI policy for Behavioral Ecology and Sociobiology. *Behav Ecol Sociobiol* 78, 2. <https://doi.org/10.1007/s00265-023-03416-5>
- Barthes, R. (1988). *La muerte del autor*. Paidós
- Bloom, H. (1997). *The anxiety of influence: A theory of poetry*. Oxford University Press.
- Bogin, B., Yang, K., Gupta, S., Richardson, K., Bransom, E., Clark, P., & Khot, T. (2024). Super: Evaluating agents on setting up and executing tasks from research repositories. *ArXiv*. <https://doi.org/10.48550/arXiv.2409.07440>

- Bourdieu, P. (1995). *Las reglas del arte: génesis y estructura del campo literario*. Anagrama.
- Dussel, E. (2016). *Filosofías del sur: descolonización y transmodernidad*. Ediciones Akal.
- Floridi, L. (2025). Distant Writing: Literary Production in the Age of Artificial Intelligence. *Minds and Machines*, 35(3), 1-26. <https://doi.org/10.1007/s11023-025-09732-1>
- Foucault, M. (1992). *¿Qué es un autor? En La arqueología del saber*. Siglo XXI Editores.
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Ghafari, A., & Buehler, M. J. (2024). SciAgents: Automating scientific discovery through multi-agent intelligent graph reasoning. *ArXiv*. <https://doi.org/10.48550/arXiv.2409.05556>
- Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38, 173–198. <https://doi.org/10.1007/BF01700692>
- Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511814563>
- Hales, T. (2005). A proof of the Kepler conjecture. *Annals of mathematics*, 1065-1185. <https://annals.math.princeton.edu/wp-content/uploads/annals-v162-n3-p01.pdf>
- Heidegger, M. (1927). *Ser y tiempo*. Editorial Trotta.
- Helfgott, H. (2013). *The ternary Goldbach conjecture is true*. *ArXiv*. <https://doi.org/10.48550/arXiv.1312.7748>
- Hui, Y. (2016). *On the Existence of Digital Objects*. Minnesota University press. <http://www.jstor.org/stable/10.5749/j.ctt1bh49tt>
- Husserl, E. (1991). *Lecciones de fenomenología de la conciencia interna del tiempo*. Editorial Trotta.
- Ifargan, T., Hafner, L., Kern, M., Alcalay, O., & Kishony, R. (2024). Autonomous LLM-driven research from data to human-verifiable research papers. *ArXiv*. <https://doi.org/10.1056/Aloa2400555>

- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Kacena, M. A., Plotkin, L. I., & Fehrenbacher, J. C. (2024). *The use of artificial intelligence in writing scientific review articles*. *Current Osteoporosis Reports*, 22(1), 115-121. <https://doi.org/10.1007/s11914-023-00852-0>
- Messeri, L., & Crockett, M. (2024). *Artificial intelligence and illusions of understanding in scientific research*. *Nature*, 627(8002), 49-58. <https://doi.org/10.1038/s41586-024-07146-0>.
- Moretti, F. (2013). *Distant reading*. Verso Books.
- Novelli, C., Taddeo, M., & Floridi, L. (2024). *Accountability in artificial intelligence: what it is and how it works*. *Ai & Society*, 39(4), 1871-1882. <https://doi.org/10.1007/s00146-023-01635-y>
- OECD Science, Technology and Innovation Outlook (2023). Enabling Transitions in Times of Disruption. <https://doi.org/10.1787/0b55736e-en>
- Pasquinelli, M., y Joler, V. (2021). *El Nooscopio de manifiesto*. *La Fuga*, 25. <https://lafuga.cl/el-nooscopio-de-manifiesto/1053>
- Pearl, J. (2009). *Causality: Models, reasoning, and inference*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511803161>
- Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. Basic Books. <https://lcn.loc.gov/2017056458>
- Lample, G., Lachaux, M.A., Lavril, T., Martinet, X., Hayat, A., Ebner, G., Rodriguez, A., & Lacroix, T. (2022). HyperTree Proof Search for Neural Theorem Proving. *ArXiv*. <https://doi.org/10.48550/arXiv.2205.11491>
- Quijano, A. (2011). *Colonialidad del poder y clasificación social*. *Contextualizaciones latinoamericanas*, 2 (5). <https://doi.org/10.32870/cl.v0i5.2836>
- Ramaré, O. (1995). *On Šnirel'man's constant*. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 22(4), 645-706. <https://hal.science/hal-02871110>
- Ramos, M. C., Collison, C. J., & White, A. D. (2024). A Review of Large Language Models and Autonomous Agents in Chemistry. *ArXiv*. <https://doi.org/10.48550/arXiv.2407.01603>

- Raayoni, G., Gottlieb, S., Manor, Y., Pisha, G., Harris, Y., Mendlovic, U., Haviv, D., Hadad, Y., & Kaminer, I. (2021). Generating conjectures on fundamental constants with the Ramanujan Machine. *Nature* 590, 67–73 (2021). <https://doi.org/10.1038/s41586-021-03229-4>
- Rosenberger, R., & Verbeek, P. (2015). A field guide to postphenomenology. *En R. Rosenberger, & P-P. Verbeek (Eds.), Postphenomenological Investigations: Essays on Human-Technology Relations* (pp. 9-41). Lanham: Lexington Books.
- Si, C., Yang, D., & Hashimoto, T. (2024). Can llms generate novel research ideas? A large-scale human study with 100+ nlp researchers. ArXiv preprint <https://doi.org/10.48550/arXiv.2409.04109>
- Simondon, G. (2007). *Del modo de existencia de los objetos técnicos*. Prometeo.
- Stiegler, B. (2010). *For a New Critique of Political Economy*. Cambridge Polity Press.
- Stiegler, B. (2014). Ars e invenciones organológicas en las sociedades de hipercontrol. *Nombres*, (28), 147-163. <https://revistas.unc.edu.ar/index.php/NOMBRES/article/view/12025>
- Stiegler, B. (2019) *The Age of Disruption: Technology and Madness in Computational Capitalism*. Wiley, UK.
- Tagnin, J. (2024). *No tan humanos. Acerca de la antropomorfización de los Agentes No Humanos de Comunicación*. *Revista Bordes*, XXXIV, <https://revistabordes.unpaz.edu.ar/no-tan-humanos-acerca-de-la-antropomorfizacion-de-los-agentes-no-humanos-de-comunicacion/>
- Tagnin, J. (2025). *Estudio sobre los agentes no humanos de comunicación. Aportes epistémicos desde una perspectiva postfenomenológica*. [Tesis doctoral UNTref]
- UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence. UNESCO. <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>
- Von Neumann, J. (1958). *The computer & the brain*. Yale Universty press.
- Waltz, D., & Buchanan, B. G. (2009). *Automating science*. *Science*, 324(5923), 43-44. <https://doi.org/10.1126/science.1172781>
- Yin, X. (2024). *Turing Tests for an AI Scientist*. ArXiv. <https://doi.org/10.48550/arXiv.2405.13352>



Zenil, H., Tegnér, J., Abrahão, F. S., Lavin, A., Kumar, V., Frey, J. G., & King, R. (2023). The future of fundamental science led by generative closed-loop artificial intelligence. *ArXiv*. <https://doi.org/10.48550/arXiv.2307.07522>

Zheng, Y., Koh, H. Y., Ju, J., Nguyen, A. T., May, L. T., Webb, G. I., & Pan, S. (2023). Large language models for scientific synthesis, inference and explanation. *ArXiv* <https://doi.org/10.48550/arXiv.2310.07984>

Zuckerfeld, M. (2011). *Las regulaciones del acceso a los conocimientos en el período preindustrial*. Introducción a una sociología histórica de la propiedad intelectual. *Redes. Revista De Estudios Sociales De La Ciencia Y La Tecnología*, 17(32), 17–37. <https://doi.org/10.48160/18517072re32.285>