

Estado de la publicación: No informado por el autor que envía

¿Qué sentirías si un robot realizara acciones en tu nombre? El caso de Artificial Intelligence Review Assistant (AIRA) y la cobotización del proceso de revisión por pares

Viviana Martinovich, Lucas Drucaroff

<https://doi.org/10.1590/SciELOPreprints.7326>

Enviado en: 2023-11-07

Postado en: 2023-11-07 (versión 1)

(AAAA-MM-DD)

¿Qué sentirías si un robot realizara acciones en tu nombre?: el caso de *Artificial Intelligence Review Assistant* (AIRA) y la cobotización del proceso de revisión por pares

What would you feel if a robot performed actions on your behalf? The case of *Artificial Intelligence Review Assistant* (AIRA) and the cobotization of the peer-review process

Viviana Martinovich¹, Lucas Drucaroff²

¹Profesora adjunta, Instituto de Salud Colectiva, Universidad Nacional de Lanús, Buenos Aires, Argentina. <https://orcid.org/0000-0003-4607-2221> vivianamartinovich@gmail.com

²Doctor en Ciencias Sociales, Éticas y Humanísticas Médicas. Investigador asistente, Consejo Nacional de Investigaciones Científicas y Técnicas, con sede en el Instituto de Salud Colectiva, Universidad Nacional de Lanús, Buenos Aires, Argentina. <https://orcid.org/0000-0002-5079-7607>. lucasdrucaroff@gmail.com

RESUMEN: En el marco del pasaje de procesos de revisión por pares de carácter relacional y artesanal a procesos industriales a gran escala, basados en inteligencia artificial (IA), el objetivo de este artículo es analizar la experiencia de interacción entre personas y robots colaborativos o “cobots” propuesta por estos nuevos entornos y recuperar indicios de posibles sesgos algorítmicos y humanos de esta interacción. Para ello, partimos de la descripción cronológica de un caso de interacción con el *Artificial Intelligence Review Assistant* (AIRA) de la editorial Frontiers que analizamos a partir de la noción habermasiana de *tecnificación del mundo de la vida*, y de conceptos como *sesgo algorítmico*, *sesgo de automatización* y *sesgo de confirmación*. La interacción persona-cobot genera un proceso de cobotización de la revisión por pares, en el que la *tecnificación* de las acciones comunicativas tiende a desarticular los acuerdos y consensos propios del ecosistema relacional de investigación y publicación de la comunidad científica, a la vez que tiene la potencialidad de retroalimentar sesgos algorítmicos y humanos, que ameritan una actitud más reflexiva en torno a la utilización de estas tecnologías.

Palabras claves: Revistas Científicas; Revisión por Pares; Industria Editorial; Inteligencia Artificial.

ABSTRACT: The aim of this article is to analyze the experience of interaction between individuals and collaborative robots or "cobots" introduced in the context of the shift from relational and artisanal peer-review processes to large-scale AI-based industrial processes. Additionally, we seeks to uncover signs of potential algorithmic and human biases within this interaction. To achieve this, we chronologically describe a case of interaction with the Artificial Intelligence Review Assistant (AIRA) developed by the publisher Frontiers. Our analysis is rooted in the Habermasian notion of the *technification of the life-world* and a set of biases (algorithmic bias, automation bias, and confirmation bias). The person-cobot interaction leads to a process of "cobotization" of peer-review, wherein the technification of communicative actions tends to disrupt the agreements and consensus on which the scientific community's relational ecosystem for research and publication are based; additionally, it has the potential to reinforce algorithmic and human biases. All of the aforementioned calls for a more reflective attitude towards the use of these technologies.

Keywords: Scientific Journals; Peer Review; Publishing Industry; Artificial Intelligence.

INTRODUCCIÓN

“Nada de esto es una argumentación contra la tecnología: razonar contra ella sería hacerlo contra nosotros mismos. Es, no obstante, una argumentación en favor de una interacción más reflexiva con la tecnología, junto con una visión radicalmente distinta de lo que es posible pensar y saber sobre el mundo. Los sistemas computacionales, como herramientas, ponen énfasis en uno de los aspectos más poderosos de la humanidad: nuestra capacidad de actuar de manera efectiva en el mundo y amoldarlo a nuestros deseos. Pero sigue siendo nuestra prerrogativa descubrir y articular esos deseos y asegurarnos de que no degradan, anulan, eliminan o borran los deseos de los demás”. (Bridle, 2020)

Estas palabras del escritor británico James Bridle son una invitación a reflexionar críticamente sobre la reciente incorporación de la inteligencia artificial (IA) a los procesos de preselección y revisión por pares, en las revistas científicas. Esta incorporación de la IA se realiza sobre una práctica científica como es la revisión por pares que, con distintos grados de formalización, estuvo presente desde el surgimiento de las ciencias modernas (Kronick, 1990; Moxham & Fyfe, 2018; Zuckerman & Merton, 1971).

Desde hace siglos, previo a su publicación, los textos circulan entre pares, se discuten, se modifican y se enriquecen, sobre la base de las nociones que, en cada momento histórico, las diversas comunidades científicas consensuaron como formas válidas de interpretar científicamente el mundo (Gadamer, 2005, 2006; Heelan, 1998). Este mecanismo de autorregulación (Biagioli, 2002; Daniel, 1993), que las ciencias han ejercido a lo largo de los siglos a escala humana, ha sido extrapolado a escala industrial por la primera generación de grandes editoriales científicas denominadas “*big five*” –entre las que se encuentran Elsevier, Springer Nature, Wiley, Taylor & Francis y American Chemical Society (ACS) (Mehta, 2019)– que se consolidaron a partir de la expansión del complejo industrial científico-editorial en la segunda mitad del siglo XX (Martinovich, 2020).

El pasaje de una revisión relacional y artesanal a un proceso industrializado y a gran escala no se produjo en todos los entornos científico-editoriales, sino solo en aquellos en los que la retroalimentación científico-técnica quedó asociada a la reproducción del capital (Habermas, 2007), lo que llevó a transformar el sistema de comunicación de las ciencias en un mercado mundial de publicaciones que, para 2023, superaría los 28 mil millones de dólares (STM, 2021). Desde estos entornos, ya en la década de 1980, ante la necesidad de acelerar la validación científica de los productos industriales para insertarlos en el mercado, comenzaron a plantear que la revisión por pares como método de consenso

científico no se adaptaba a las nuevas demandas de respuestas inmediatas (Mathis, 1982; Wessel, 1982). Y fue en estos mismos entornos científicos que, en las primeras décadas del siglo XXI, se produjo un aumento de prácticas fraudulentas o errores significativos que se traducen en un incremento del número de “retractaciones” (Peng et al., 2022; Shuai et al., 2017; Van Noorden, 2011), es decir, artículos que se retiran de circulación con posterioridad a su publicación, lo que pone de manifiesto la “debilidad” de los procesos de revisión para detectar conflicto de intereses, manipulación de imágenes, venta de autoría, entre otras prácticas fraudulentas, cuando están en juego grandes intereses económicos (Campanario, 1998a, 1998b; Marcus, 2012).

Del mismo modo que otros procesos industriales a gran escala inciden en los ecosistemas naturales –como la tala indiscriminada de árboles o la contaminación de las aguas por residuos industriales– la producción industrial a gran escala de artículos científicos ha generado profundos cambios en el ecosistema de investigación y publicación de la comunidad científica. Desde el sector científico-académico, desde áreas como biología molecular, biofísica, bioquímica entre otras, hace décadas que vienen señalando que, las exigencias de hiperpublicación, la gran competencia por la obtención de fondos y los criterios de estratificación de esas publicaciones basados en indicadores bibliométricos crearon un mecanismo distorsivo que instrumentaliza el proceso de revisión por pares, a la vez que tiende a desalentar las preguntas o los enfoques más creativos y novedosos que demandan más tiempo de maduración, por lo que los grupos de investigación consolidados, en lugar de explorar nuevos campos, suelen apegarse a sus fórmulas exitosas ya probadas, de manera de asegurar la rápida aprobación de sus publicaciones (Alberts et al., 2014; Horrobin, 1990; Li & Agha, 2015).

A pesar de que los distintos cuestionamientos al proceso de revisión por pares se originan en concepciones y racionalidades científicas muy diversas (Aguado-López & Becerril-García, 2021; Bali, 2015; Ford, 2013; Wolfram et al., 2020), ciertas cualidades intrínsecamente humanas –como la divergencia de criterios a la hora de evaluar un artículo, la subjetividad, y la diversidad de perspectivas– comenzaron a ser consideradas negativas (Cole et al., 1981; Björk & Solomon, 2013). Y este malestar hacia la condición humana, sumado a un proceso generalizado de robotización de la industria y de la vida cotidiana (Hentout et al., 2019), fue el puntapié inicial para que la segunda generación de grandes oligopolios editoriales, que nacieron y alcanzaron mayores niveles de concentración en la era digital (Larivière et al., 2015), apostara al diseño de sistemas

basados en inteligencia artificial (IA) para la selección de manuscritos y la revisión por pares (Ghosal, 2019; Van Noorden, 2021), proceso que en este trabajo denominamos como “cobotización del proceso de revisión por pares”.

¿Qué entendemos por “cobotización del proceso de revisión por pares”?

Según Osika (2022), la “cobotización” de la industria es la implantación masiva de robots colaborativos (cobots) bajo un modelo de producción en el que el uso de soluciones tecnológicas se enriquece con el factor humano, es decir, el proceso productivo se basa en la interacción de personas y máquinas. La “cobotización” basada en inteligencia artificial supone la incorporación de un agente que puede identificar el estado de un entorno y actuar sobre él, para lo cual operan algoritmos de aprendizaje que intentan emular las acciones humanas. Este modo de aprendizaje automático, que surgió del campo de la inteligencia artificial, es el que permite que las máquinas cobren autonomía sin una programación explícita, es decir, harán su propio proceso de aprendizaje imitando no solo el modelo de aprendizaje humano, sino analizando grandes volúmenes de información sobre prácticas humanas específicas (Borboni et al., 2023; Danks & London, 2017; Das et al., 2015; Dastin, 2018; Fu et al., 2020).

En este marco, la “cobotización” de la revisión por pares es un tipo de interacción persona-cobot que, al incorporar la inteligencia artificial a las acciones que llevan a la aceptación o rechazo de una investigación, irrumpe tanto en las nociones que las diversas comunidades acordaron en términos históricos como formas válidas de interpretar científicamente el mundo, como en los entornos relacionales en los que se crean esos acuerdos (Martinovich, 2023). En ese proceso de apropiación de acciones humanas, cultural e históricamente definidas, y su traducción a operaciones algorítmicas para la toma de decisiones, el componente humano de esa “colaboración” queda aislado de su propio entorno relacional, y pasa a interactuar en un nuevo tipo de entorno cuyas acciones se “desacoplan” de los acuerdos del mundo social, proceso que Habermas, ya en la década de 1980, describía como “tecnificación del mundo de la vida” (2002, p. 259).

A su vez, esta “cobotización” de la revisión por pares se da en un esquema de producción editorial científica a gran escala industrial, en el que la interacción humana de la comunidad científica es mediada por un sistema cuyo objetivo es reducir los tiempos de la cadena productiva y disminuir los sesgos humanos. Pero tal como señala Mike Thelwall (2019), los sistemas de IA, al analizar grandes volúmenes de información sobre prácticas

humanas, aprenden los sesgos actuales y los empeoran al tratar de reproducirlos de manera eficiente. Según Runshan Fu *et al.*, si bien inicialmente fueron considerados neutrales y justos, los algoritmos parecen ser cada vez más sesgados, perpetuando y profundizando desigualdades estructurales en la sociedad, lo que genera crecientes preocupaciones sobre el sesgo algorítmico (Fu et al., 2020).

Según Thelwall (2019), evaluar los sesgos de los softwares que no utilizan IA es relativamente sencillo porque el proceso es transparente y puede ser verificado o cuestionado, pero en el caso de la IA es más complejo porque la mayoría de los algoritmos son opacos y matemáticamente abstractos, dado que las decisiones tomadas por la IA dependen del proceso de aprendizaje y de su interacción con el conjunto de datos, de modo que los humanos no acceden a las reglas y, por lo tanto, no pueden desafiarlas.

Ante este escenario, el objetivo de este artículo es analizar la experiencia de interacción persona-cobots en los procesos de revisión por pares gestionados por el *Artificial Intelligence Review Assistant* (AIRA) de la editorial Frontiers, y recuperar ciertos indicios de posibles sesgos algorítmicos y humanos de esta interacción que podrían tender a privilegiar ciertas racionalidades científicas por sobre otras.

ACERCA DE LA INVESTIGACIÓN

Partimos de un estudio de caso, como estrategia de investigación (Eisenhardt, 1989; Yin, 2003), en el que presentamos una descripción cronológica de la experiencia de uno de los autores de este artículo, al aceptar la invitación de la revista *Frontiers in Psychiatry* para actuar bajo la figura de *handling editor*, que traduciremos como *editor encargado*.

Según la propia compañía editorial, la figura del *editor encargado* es quien realiza una evaluación inicial para garantizar que un manuscrito sea “científicamente sólido” y se ajuste al alcance de la revista y la sección a la que se envió. Después de esta primera verificación, el editor encargado debe ser quien decide si envía el manuscrito a revisión externa o recomienda su rechazo en esta instancia inicial. En el caso de decidir iniciar el proceso de revisión externa, el editor encargado es quien debe invitar a las y los revisores y “supervisa directamente la interacción entre revisores y autores durante el proceso de revisión” (Frontiers, 2023d).

La particularidad del caso presentado es que quien articula las comunicaciones entre editores, revisores y autores es el *Artificial Intelligence Review Assistant* (AIRA) y, según

Frontiers, “es la primera vez que la industria utiliza la IA para revisar artículos de investigación y ayudar en el proceso de revisión por pares” (Frontiers, 2020). Si bien, por momentos, la compañía define a AIRA como un “asistente” que brinda soporte a las decisiones humanas, en otros, parecería cobrar autonomía, dado que “AIRA lee cada artículo y actualmente puede realizar hasta 20 recomendaciones en apenas unos segundos, incluyendo la evaluación de la calidad del lenguaje, la integridad de las figuras, la detección de plagio, así como la identificación de posibles conflictos de intereses” (Frontiers, 2020).

En cuanto a la escala de producción, se trata de una compañía editorial que en solo 10 años pasó de recibir 3.557 artículos en 2012 a recibir 240.037 artículos en 2022, de los cuales se aceptó el 51,9% (Frontiers, 2023b), una tasa de aceptación marcadamente más elevada que en la mayoría de las revistas científicas. Dado que, para el proceso de revisión de cada artículo, Frontiers supone la participación de al menos un editor en jefe de la especialidad, una persona que actúe como editor encargado y dos como revisores externos, llevar a cabo los procesos de revisión del total de artículos recibidos demandaría un gran volumen de participación humana.

La descripción cronológica del caso se divide en tres etapas: la fase 0, que narra el momento previo a aceptar ser editor encargado del manuscrito; la fase 1, centrada en las acciones inmediatamente posteriores a aceptar la invitación a ser editor encargado del manuscrito; y la fase 2, que detalla la participación del editor encargado del manuscrito en el proceso de revisión.

A partir de la noción de “tecnificación del mundo de la vida” (Habermas, 2002), y de conceptos como sesgo algorítmico (Danks & London, 2017), sesgo de automatización y sesgo de confirmación (Bridle, 2020), analizamos la experiencia desde dos ejes: 1) *De la revisión relacional humana a la cobotización*, en el que abordamos el tipo de interacción persona-cobot que propone la incorporación de la inteligencia artificial en los procesos de revisión; y 2) *Del concepto humano al algoritmo*, en el que recuperamos ciertos indicios de posibles sesgos algorítmicos y humanos en la experiencia analizada que podrían tender a privilegiar ciertas racionalidades científicas por sobre otras.

Para facilitar la lectura, se transcriben y se traducen al castellano algunos fragmentos de las comunicaciones generadas en el proceso de revisión. Con relación a las normas éticas, en todo momento se preservó la confidencialidad de las (escasas) identidades humanas

involucradas y no se expone ningún dato que pueda incumplir con el acuerdo de colaboración suscrito al inicio del proceso de revisión.

RELATO DEL CASO

Fase 0: Momento previo a aceptar ser editor encargado de un manuscrito

A las 5:06 hs del jueves 24 de marzo de 2022 recibí un correo electrónico titulado “Invitación para editar un manuscrito”, cuyo remitente era *Frontiers in Psychiatry*. El contenido del correo electrónico comenzaba diciendo (traducción literal): “Querido Dr. Drucaroff, considerando su experticia, yo le estoy escribiendo para preguntarle si usted sería capaz de ser editor de un manuscrito”, la utilización de un pronombre en primera persona del singular mostraba una discordancia con la firma que figuraba al final de la comunicación: “*Frontiers in Psychiatry Editorial Office*” (Oficina Editorial de *Frontiers in Psychiatry*). Esta discordancia no desentonaba de la variada gama de invitaciones que recibe cotidianamente una persona que se dedica a la investigación: invitaciones para participar de congresos en el exterior, para escribir y publicar artículos dentro de su “campo de trabajo” para revistas internacionales y otras invitaciones similares. En todos estos casos se alude a “la inmensa relevancia del aporte de su producción a la ciencia” (lo cual en mi caso es absolutamente inverosímil, según mi propio juicio). Se trata de frases que endulzan el ego con el fin de facilitar el cobro de altas sumas de dinero para hacer efectiva la invitación. Si bien cuento con cierta experiencia previa en publicación de artículos en revistas internacionales con referato, esta era la primera vez que recibía una invitación para ser editor de un artículo. También era la primera vez en la que una invitación de estas características no llevaba la firma de una persona, por lo que decidí no responder el correo electrónico.

Los antecedentes como revisor por pares y editor de revistas con referato suelen ser valorados positivamente a la hora de evaluar la carrera de investigación en casi todas las instituciones científicas de prestigio. Además, cada vez que participaba en un proceso de revisión por pares, en lo personal tenía la sensación de estar aportando al desarrollo de la ciencia, sensación compartida también por otros colegas a quienes consulté. Por tales motivos, 48 horas después, al recibir un segundo correo electrónico que me “recordaba amablemente la invitación”, decidí considerar la posibilidad. En primer lugar, me interioricé sobre la revista *Frontiers in Psychiatry*, la cual no conocía previamente. Constaté que la revista se encontraba indizada en PubMed y que ocupaba un lugar en el

primer cuartil de revistas del área de Psiquiatría y Salud Mental de Scimago Journal Rank, desde 2013 en adelante. En segunda instancia, consulté con gente que tiene mayor conocimiento del ámbito editorial científico, para evaluar si la invitación podría ser genuina. Tuve una conversación con quien a la postre es autora del presente artículo, quien me informó que este tipo de invitaciones eran perfectamente plausibles y me sugirió aceptar para que nos interioricemos en los procesos editoriales de esa revista. Basado en esto, decidí aceptar la invitación para ser editor del manuscrito propuesto.

Fase 1: Aceptar la invitación a ser editor encargado del manuscrito

Luego de aceptar la invitación, recibí un correo electrónico de confirmación firmado por el *gerente de operaciones de revisión (review operations manager)* que, si bien explicitaba su nombre y apellido, en la actualidad ese nombre no es encontrable en Google asociado a la posición, y no firmó ninguna otra comunicación a lo largo del proceso de revisión. Ese correo electrónico fue recibido inmediatamente al pulsar en el botón de aceptación, el 29 de marzo de 2022 a las 14:21 hs, lo cual hace que no sea plausible que haya sido enviado por un ser humano.

Lo que me resultó sumamente llamativo fue el contenido de un segundo correo electrónico recibido también en el instante posterior a la aceptación. El título del correo era “*El revisor no ha respondido a su invitación para revisión*”. El título se refería a “su” invitación, adjudicándome la acción de haber enviado invitaciones para revisar el artículo. Esto me suscitó varios interrogantes en forma simultánea: ¿qué revisor?, ¿qué invitación?, ¿mi identidad está siendo suplantada? El contenido del correo electrónico proveía la respuesta al primer interrogante: “...*La invitación a los siguientes revisores aún está pendiente...*” y, a continuación, figuraban los nombres de tres personas. Al respecto de esas personas decía:

“Si sabe que tienen dificultades para aceptar la invitación, hágamelo saber para que podamos brindarle asistencia. De lo contrario, le recomiendo que se lo recuerde personalmente de nuevo o que revoque esta invitación de revisión e invite a nuevos revisores. Puede realizar estas acciones en el foro de revisión: [...]. Por favor, no dude en ponerse en contacto conmigo si tiene alguna pregunta. Atentamente, Frontiers in Psychiatry Editorial Office”.

Nuevamente el correo estaba redactado en primera persona del singular (salvo por una conjugación en primera persona del plural), pero carecía de firma o del nombre de una persona a quien referenciarse.

Al ingresar al enlace suministrado para dar inicio a mi tarea como editor encargado del manuscrito, noté que dicho proceso ya estaba largamente iniciado. Se explicitaba un reporte de actividades que contaba con un total de 26 acciones que ya se habían realizado previamente. De ellas, 16 consistían en invitaciones para ser editor encargado del artículo en cuestión, enviadas en tres tandas: cinco el día 7 de marzo, otras cinco el día 17 de marzo, y por último seis más el día 24 de marzo (entre las cuales se encontraba la invitación que recibí en dicha fecha). Tres cuestiones resultaban destacables: por un lado, el envío simultáneo de más de una invitación para hacerse cargo del mismo rol; segundo, el hecho de que el rol en cuestión era justamente el que yo acababa de aceptar; y, por último, coincidentemente con los correos electrónicos que había recibido, las acciones estaban genéricamente adjudicadas a “la oficina editorial” sin consignar el nombre de una persona que las hubiera llevado a cabo, por lo que podía considerarse que el proceso estaba automatizado.

La ya mencionada “oficina editorial” había invitado a cinco personas para realizar la revisión por pares del artículo, nuevamente sin especificar ninguna persona que hubiera efectuado dichas invitaciones. La elección de revisores suele ser una tarea humana, potestad de quien tiene a cargo el seguimiento del proceso de revisión del manuscrito que, en el esquema propuesto por la editorial Frontiers, sería la figura de editor encargado del manuscrito, tal como está explicitado en la guía para editores: “*si un manuscrito se envía para revisión por pares, el editor encargado es responsable de invitar y supervisar a los revisores expertos*” (Frontiers, 2023c). El hecho de que se hubieran enviado estas invitaciones en forma inconsulta, e incluso antes de haber asignado un editor encargado, indicaba una vez más que la participación humana parecía ser una excusa para darle el sustento formal requerido a un sistema automático de revisión.

Esta presunción ganó fuerza al ver que, entre las acciones realizadas antes de mi incorporación, había cinco recordatorios de plazos de revisión vencidos o por vencer, también enviados por “la oficina editorial”, dos de los cuales estaban dirigidos a la *editora en jefe de la especialidad (Specialty Chief Editor)*, quien no había realizado ninguna acción antes de dichos recordatorios, tampoco después de ellos, ni en ningún momento del proceso de revisión.

Fase 2: Participación en el proceso de revisión como *editor encargado del manuscrito*

Al ingresar al sistema, observé que había varias invitaciones a revisores que se encontraban pendientes de respuesta, las cuales habían sido enviadas previamente y en forma automática. Sin embargo, decidí comenzar mi tarea e invitar a revisar el manuscrito a dos colegas por su pertinencia temática. Al recibir su rápida negativa decidí ingresar nuevamente a la plataforma de Frontiers y descubrí cómo elige el sistema a los potenciales revisores para enviar las invitaciones automáticas: hay un listado de personas sugeridas, virtualmente interminable (nunca logré llegar al final) cuya pertinencia temática es estimada, sin participación humana, aparentemente según coincidencia entre los términos utilizados en el manuscrito y los utilizados por los potenciales revisores en sus propias publicaciones. Al observar los perfiles de las personas sugeridas encontré varias inconsistencias, con supuestas coincidencias temáticas que parecían más fruto del azar o de la arbitrariedad que de trayectorias científicas pertinentes para revisar el artículo. De todos modos, elegí minuciosamente algunos de los nombres sugeridos por el sistema, corroboré la pertinencia según publicaciones previas y antecedentes académicos, y envié invitaciones para la revisión del manuscrito en cuestión.

Continué familiarizándome con el sistema. Me sorprendió ver que la editorial cuenta con un asistente de revisión de IA, cuyo nombre es *Artificial Intelligence Review Assistant* (AIRA). En muchos casos, AIRA relataba en primera persona lo que ya había hecho o lo que haría más adelante en el proceso, utilizando expresiones como:

- “no detecté similitudes con ningún otro manuscrito enviado a Frontiers”;
- “busqué en la base de datos iThenticate y este manuscrito tiene un nivel aceptable de superposición textual con los artículos publicados”;
- “no detecté imágenes humanas en los archivos de figuras”;
- “no detecté similitudes entre las figuras del artículo”;
- “no detecté ningún elemento controvertido en este manuscrito”;
- “no detecté potenciales conflictos de intereses comerciales”.

Además, entre otras acciones más, AIRA refiere que evalúa el “*comportamiento de los revisores*”, es decir, muestra cuántos revisores se han retirado del proceso luego de aceptar participar, y cuántos han recomendado rechazar la publicación del artículo. En particular este ítem muestra una advertencia con la leyenda “*No confiable*”, cuyo significado concreto no está explicado.

En la mayoría de las acciones no estaba explicitado, ni era evidente, qué procedimiento aplicaba AIRA para llegar al resultado deseado. Sin embargo, pensé que en caso de realizar con efectividad las acciones esgrimidas, la asistencia que brindaba AIRA era invaluable.

Al tiempo que la fascinación por esta plataforma ganaba terreno, noté que uno de los potenciales revisores sugeridos era uno de los coautores del trabajo a revisar, y aun así el sistema lo sugería entre las primeras opciones y permitía enviarle una invitación para revisar su propio artículo. Obviamente decidí revocar la invitación. Sorprendido por la magnitud del error y por la aparente facilidad que tendría evitarlo, comencé a preguntarme por la infalibilidad del sistema de inteligencia artificial: ¿cometerá errores así de groseros en otros aspectos que son imposibles de corroborar?

Dos días después, recibí un correo electrónico que se titulaba “*Haciendo que comience el proceso de revisión del manuscrito*”, siendo que el proceso había comenzado semanas antes, pero sin mi participación. En su contenido se podía leer:

“Para iniciar el proceso de revisión, complete una verificación preliminar para garantizar que el manuscrito sea adecuado para la revisión. [...] Complete su evaluación inicial dentro de los próximos 2 días para evitar que se envíe un manuscrito inadecuado para su revisión”

Sin embargo, el manuscrito había sido enviado a revisión automáticamente sin haber completado este paso. Decidí completar el paso y esperar las respuestas pendientes de las personas invitadas para hacer la revisión.

Seis días después, recibí un correo titulado “*Nombramiento de revisor demorado*”, que me instaba a conseguir revisores lo más rápido posible y me informaba que el sistema había enviado nuevas invitaciones en forma automática a varios *editores de revisión*, que integraban el cuerpo editorial de la revista. De los 24 *editores de revisión*, todos miembros del cuerpo editorial de la revista, identificados por el sistema e invitados automáticamente, ninguno respondió nunca a la invitación, ni siquiera dando una respuesta negativa a participar.

En ese momento, descubrí que, en el reporte de acciones, el sistema se refería a mí como “*editor asociado*”. Busqué entre los correos enviados por Frontiers alguno en el que se me invitara a ser editor asociado o que explicitara que se me había asignado ese rol, pero no existía tal notificación. Tampoco encontré mi nombre en el extensísimo listado de editores de la revista *Frontiers in Psychiatry*, que al día 19 de septiembre de

2023 incluía 5.464 editores, de los cuales 1.071 eran editores asociados. Nuevamente, me sentí invadido por la sensación de ser una de las tantas piezas intercambiables de un engranaje que tenía como único objetivo acelerar el proceso de publicación; sensación muy diferente de aquella de estar contribuyendo al desarrollo de la ciencia, que había sido una de mis motivaciones iniciales.

A raíz de esto, decidí declinar mi participación como *editor encargado del manuscrito*. Sin embargo, dentro de este sistema altamente automatizado me fue imposible encontrar un botón mediante el cual comunicar mi decisión, y tampoco pude hallar un dato de contacto de una persona responsable a quien informarle. Finalmente, frustrado y sin opción para excluirme formalmente de la tarea, opté por tomar la misma actitud que la *editora en jefe de la especialidad* y los 24 miembros del cuerpo editorial de la revista mencionados anteriormente: dejar de interactuar con el sistema. Eso fue lo que hice durante las semanas posteriores con todos los mensajes automáticos que recibí. Cotidianamente recibía correos electrónicos que me mostraban cómo las personas invitadas de forma automática por el sistema declinaban realizar la revisión. Incluso no realicé acción alguna al ser notificado de que los revisores #1 y #2 aceptaron realizar la revisión, ni al recibir el aviso de que se encontraba listo el reporte de revisión del revisor #1, ni de la aceptación a participar del revisor “#3”. El sistema había enviado automáticamente decenas de invitaciones a revisores, sin anular las invitaciones excedentes, una vez que se completó el número requerido. Consideré que se trataba de un error, y con el objetivo de evitar que más personas realizaran un trabajo innecesario, decidí cancelar manualmente las invitaciones pendientes de respuesta. Así fue como, por empatía con otros colegas, estaba nuevamente activo en el proceso que había decidido abandonar cinco semanas atrás.

Unos días después recibí la notificación de que el revisor #3 había completado su reporte de revisión, junto con el siguiente aviso:

“...El tipo de artículo manuscrito requiere un mínimo de dos revisores que presenten un informe de revisión independiente. Dado que se alcanzó este número y algunos revisores solicitaron ver la respuesta de los autores, la discusión de revisión interactiva debería activarse ahora. Tenga en cuenta que, si la discusión interactiva no se activa manualmente dentro de los tres días, el manuscrito pasará a la etapa interactiva automáticamente para evitar retrasos en el proceso de revisión por pares”

Ante la inevitabilidad del paso a la siguiente etapa, y con la intención de no demorar más el proceso y no perjudicar a las personas involucradas, decidí activar la “*discusión de revisión interactiva*”.

Cuatro días después recibí un correo electrónico absolutamente novedoso dentro de lo acostumbrado en este proceso: estaba escrito por seres humanos. El remitente era el autor principal del artículo, quien notó que aún estaba pendiente el reporte del revisor #2 y me preguntaba si tenía información al respecto. En esta instancia, noté que el sistema, al no cancelar automáticamente las invitaciones pendientes una vez que se alcanzaba el número requerido de revisores, permitía que participara un mayor número de revisores, aumentando la probabilidad de que al menos dos revisores entreguen sus reportes y acepten el artículo. En términos de eficiencia se aumentaba la probabilidad de aprobar un artículo y disminuía el tiempo hasta lograrlo, sin importar cuántos revisores dejaban vacantes sus reportes o recomendaran rechazar el artículo. En el caso de que dos revisores lo aceptaran, ya se cumplían las condiciones para su publicación.

Sin posibilidades de salirme de la lógica del sistema, le respondí a los autores que se procedería con los reportes de revisión disponibles hasta el momento. En sus reportes, los revisores solicitaban cambios menores, que los autores respondieron con celeridad.

El 18 de julio de 2022, 148 días después del envío inicial por parte de los autores, el artículo era aprobado para publicación. La media de tiempo que tarda un artículo biomédico desde su envío hasta su aceptación oscila entre 50 y 276 días (Andersen et al., 2021), por lo que, sumado a las vicisitudes de la experiencia vivida y al hecho de que los autores pagaron aproximadamente \$3.000 euros, al finalizar el proceso me pregunté: ¿realmente se eficientizó el proceso en algún aspecto o simplemente se diluyó al máximo la intervención humana y sus costos?, ¿esta experiencia habrá sido un hecho subjetivo aislado u otros colegas tendrán vivencias similares?

ANÁLISIS Y DISCUSIÓN

De la revisión relacional humana a la cobotización

El relato del caso nos muestra un panorama muy alejado del vínculo que prima en las prácticas de los entornos científico-editoriales no industrializados. En los procesos de revisión por pares centrados en prácticas relacionales persona-persona, en la elección de un editor o una editora para realizar el seguimiento de un manuscrito, además de la

trayectoria, tienen relevancia los vínculos de confianza y respeto que se establecen al interior de la comunidad científica. En ese marco, ser elegido como editor o editora resulta ser efectivamente un reconocimiento a la labor realizada. Por el contrario, la automatización de invitaciones masivas y simultáneas produce una sensación de despersonalización y de distanciamiento de aquellos valores consensuados en las prácticas relacionales.

Cuando en el relato se menciona *“cada vez que participaba en un proceso de revisión por pares, tenía la sensación de estar aportando al desarrollo de la ciencia, sensación compartida también por otros colegas a quienes consulté”*, se invocan esas acciones que respondían a acuerdos y a valores compartidos en el plano relacional, al interior de la propia comunidad científica (Martinovich, 2023). Habermas diría que este tipo de acciones comunicativas no pueden desconectarse del “contexto de saber cultural compartido, normas vigentes y motivaciones imputables, que constituye el mundo de la vida, porque tienen que servir de los recursos que caracterizan la formación lingüística del consenso” (Habermas, 2002, p. 259). Sin embargo, en los subsistemas en los que opera una *tecnificación* de las acciones comunicativas, no solo se simplifica la comunicación, sino que se la sustituye, por lo que los procesos sociales de formación de consenso quedan desvalorizados y sometidos al tipo de interacciones que imponen los entornos tecnificados.

En estos subsistemas “el mundo de la vida ya no es necesario para la coordinación de las acciones”, diría Habermas (2002, p. 259). Y esta desconexión genera un efecto paradójico en el componente humano de la interacción: “como un alivio de la necesidad de comunicación y una reducción de los riesgos que la comunicación comporta” (Habermas, 2002, p. 259), desarticulando no solo el compromiso humano con la acción, sino los consensos y los valores creados por las propias comunidades.

Este efecto descrito por Habermas está presente en diversas partes del relato. Está presente en las innumerables invitaciones automáticas que no obtuvieron respuesta por parte de los integrantes humanos del sistema, *“ni siquiera dando una respuesta negativa a participar”*. Está presente también en el desacoplamiento de los procesos sociales de formación de consenso que bloquean la posibilidad de sentirse parte de algo:

“...me sentí invadido por la sensación de ser una de las tantas piezas intercambiables de un engranaje que tenía como único objetivo acelerar el proceso de publicación; sensación muy diferente de aquella de estar

contribuyendo al desarrollo de la ciencia, que había sido una de mis motivaciones iniciales.

A raíz de esto, decidí declinar mi participación como editor encargado del manuscrito”.

Esta desarticulación de las prácticas creadas históricamente por la cultura no parece ser un efecto colateral indeseado para aquellas industrias cuya finalidad es la reproducción del capital. La cantidad de artículos que publica Frontiers aumenta exponencialmente año a año. En 2022, aceptó 124.508 artículos y cobró, en promedio, U\$2.270 por los cargos de procesamiento de cada artículo o *article processing charges* (APC), lo que lleva a un volumen de ingresos que superarían los U\$282 millones de dólares. Para llegar a estas cifras en tan solo 10 años, la editorial creó desde sus inicios un amplio “*staff*” dedicado a custodiar los intereses de la compañía, que fue denunciado en 2015 por la junta de editores de ciencias médicas de Frontiers, quienes publicaron el *Manifiesto of editorial independence of editors of Frontiers medical journals*, en el que denunciaban que la editorial se proponía “reinventar” las modalidades de publicación médica:

“El manifiesto sintetiza la continua interferencia de la Editorial con nuestra independencia editorial, las transgresiones documentadas, los inaceptables procedimientos de revisión por pares de manuscritos de artículos médicos [...]. Se requiere que la Editorial responda al manifiesto, implemente cambios para que se cumplan los estándares internacionales de publicación médica y se establezca y garantice la independencia editorial total”. (Editors of Frontiers medical journals, 2015)

Efectivamente la editorial respondió al manifiesto y lo hizo despidiendo a todas las personas firmantes. La pregunta que cabe es: ¿esa vacancia de editores estables es suplida con un elenco inestable de *editores encargados* “coordinados” por el propio sistema?

Según Bauer *et al.* (2018), uno de los cuatro tipos de interacción propuesta por los nuevos modelos de industrialización es la completa sustitución de los seres humanos, lo que aún no ha sido explicitado en el caso de Frontiers. Sin embargo, tal como se expresa en el relato, el hecho de que AIRA hubiera enviado “*estas invitaciones en forma inconsulta, e incluso antes de haber asignado un editor encargado, indicaba una vez más que la participación humana parecía ser una excusa para darle el sustento formal requerido a un sistema automático de revisión*”.

Al describir sus prácticas, la editorial destaca que “la revisión por pares está a cargo de investigadores activos, cuidadosamente designados por nuestros consejos editoriales según estrictos criterios de excelencia” (Frontiers, 2023c), es decir, hace hincapié en el

factor humano. Sin embargo, esa mención contrasta con la automatización del proceso, bajo la firma genérica de “*oficina editorial*”:

“...coincidentalmente con los correos electrónicos que había recibido, las acciones estaban genéricamente adjudicadas a “la oficina editorial” sin consignar el nombre de una persona que las hubiera llevado a cabo, por lo que podía considerarse que el proceso estaba automatizado. La ya mencionada “oficina editorial” había invitado a cinco personas para realizar la revisión por pares del artículo, nuevamente sin especificar ninguna persona que hubiera efectuado dichas invitaciones”.

Del concepto humano al algoritmo

En el relato del caso encontramos menciones a acciones realizadas por AIRA que implican la traducción de conceptos humanos, cultural e históricamente definidos, a operaciones algorítmicas para la toma de decisiones. En este sentido, Danks y London señalan que “si bien los sistemas autónomos podrían considerarse neutrales o imparciales, podrían emplear algoritmos sesgados (en cierto sentido) que causan un daño significativo que pasa desapercibido y no se corrige, tal vez hasta que sea demasiado tarde” (2017, p. 1). De hecho, entre los múltiples sesgos reportados por la bibliografía, se encuentran el *sesgo algorítmico* de la IA, que solo es posible de detectar cuando los humanos analizan los resultados de su aplicación, y los *sesgos de automatización* y *de confirmación* de los humanos en la interacción con la IA.

En cuanto al *sesgo algorítmico*, tal como señala Mike Thelwall (2019), en los softwares que no utilizan IA, evaluar los sesgos es relativamente sencillo porque el proceso es transparente y puede ser verificado o cuestionado. Pero en el caso de la IA es más complejo porque la mayoría de los algoritmos son opacos y matemáticamente abstractos, dado que las decisiones tomadas por la IA dependen del proceso de aprendizaje y de su interacción con el conjunto de datos, de modo que los humanos no acceden a las reglas y, por lo tanto, no pueden desafiarlas. De hecho, en el caso de AIRA, no hay una descripción exhaustiva de las concepciones humanas que subyacen a los conceptos presentes en la preselección de los manuscritos. Por ejemplo, cuando AIRA solicitó al editor encargado que “*complete su evaluación inicial dentro de los próximos dos días para evitar que se envíe un manuscrito inadecuado para su revisión*”, tal como se describe en el relato, al momento de enviar ese mensaje, AIRA ya había enviado el artículo a revisión externa. Y cabe la pregunta ¿qué noción de “adecuado” o “inadecuado” configuró AIRA? De forma

similar, cuando AIRA afirmó “*no detecté ningún elemento controvertido en este manuscrito*” ¿qué nociones de “controversia” configuró? ¿rechaza automáticamente AIRA a aquellos artículos que considera “controvertidos”?

Como señalamos en la introducción, en esa traducción de conceptos humanos a procesos automáticos operan algoritmos de aprendizaje que intentan emular las acciones humanas, lo cual permite que las máquinas hagan su propio proceso de aprendizaje analizando grandes volúmenes de información sobre prácticas humanas específicas. Entre los múltiples ejemplos que ilustran los sesgos algorítmicos ya reportados en la bibliografía, se encuentra la herramienta de contratación experimental de Amazon basada en IA, cuyos modelos fueron entrenados para examinar a las personas que se postulaban a los cargos de desarrollo de software, mediante la observación de patrones en los currículos enviados a la empresa durante un período de 10 años. Un año después de su puesta en funcionamiento, la empresa se dio cuenta de que su nuevo sistema no calificaba a los candidatos de manera neutral en cuanto a género, sino que reproducía de manera más eficiente los sesgos precedentes (Dastin, 2018).

En el caso de Frontiers, el editor encargado es convocado, en teoría, para realizar la “*evaluación inicial*”. Sin embargo, cuando el sistema envía invitaciones a decenas de editores encargados y no obtiene respuestas, AIRA comienza a realizar las tareas propias del editor encargado. En este caso, si los algoritmos de aprendizaje de AIRA intentaran emular las acciones humanas vinculadas al proceso de revisión de artículos, analizando patrones en grandes volúmenes de información sobre qué artículos son aceptados y cuáles rechazados por la comunidad científica que interactúa con el sistema, ponderarían la racionalidad que prime en los artículos que hayan sido aceptados, transformando esa racionalidad en un estándar, con la potencialidad de que descarte aquellos trabajos que intenten no seguir el patrón y ser más innovadores o creativos.

En este sentido, Danks y London señalan que “la posibilidad de un sesgo algorítmico es particularmente preocupante para los sistemas autónomos o semiautónomos, ya que no necesitan involucrar a un ser humano ‘en el circuito’ (ya sea activo o pasivo) que pueda detectar y compensar los sesgos en el algoritmo o modelo” (2017, p. 1). De hecho, si bien en el marco de la experiencia con AIRA, en el relato surgen varias inconsistencias en la elección automatizada de revisores, basada en supuestas coincidencias temáticas “*que parecían más fruto del azar o de la arbitrariedad que de trayectorias científicas pertinentes para revisar el artículo*”, el sistema no preveía ningún tipo de interacción que

permitiera a los humanos señalar esas inconsistencias. Según Danks y London, el problema de esta autonomía es que:

“...a medida que los sistemas se vuelven más complicados y su funcionamiento más inescrutable para los usuarios, puede resultar cada vez más difícil comprender cómo los sistemas autónomos toman sus decisiones. Cuando el sesgo está determinado por el proceso de toma de decisiones y no únicamente por los resultados, esta inescrutabilidad puede cuestionar la noción misma de monitoreo humano del sesgo”. (Danks & London, 2017, p. 1)

A esta imposibilidad de comprender cómo los sistemas autónomos toman sus decisiones, se suma el *sesgo de automatización* asociado al *sesgo de confirmación*, que nos hablan de la obnubilación que nos provoca la automatización de procesos: “damos más valor a la información automatizada que a nuestras propias experiencias, incluso cuando aquella entra en conflicto con otras observaciones” (Bridle, 2020, p. 33).

En el relato, ese valor otorgado a la automatización se percibe al inicio, cuando aún “...*la fascinación por esta plataforma ganaba terreno*”. Según Bridle, el sesgo de automatización no se origina necesariamente en la tecnología, sino que cuando las personas se enfrentan a un problema y tienen poco tiempo para resolverlo, “ante la posibilidad de delegar la toma de decisiones, el cerebro toma la senda del menor esfuerzo cognitivo, el atajo más corto, que los asistentes automatizados le ofrecen de forma casi instantánea” (Bridle, 2020, p. 35).

Sin embargo, según Habermas (2007), no se le atribuiría un valor diferencial a la automatización en sí, sino a la racionalidad técnica que estructuró las ciencias modernas y que motorizó la industrialización y la automatización. Esta ponderación de la racionalidad técnica por sobre otras racionalidades se trasladaría al amplio espectro de acciones y productos propuestos desde estos mismos entornos. Este mecanismo, sumado a la gran capacidad operativa que han desarrollado estos sistemas, tiene la potencialidad de perfeccionar los sesgos precedentes y ponderar aún más ciertas racionalidades científicas por sobre otras.

En el caso de AIRA, el sistema ofrece continuamente soluciones que simplificarían la participación humana. Pero ¿qué sucede cuando el sistema explicita un nivel de automatización tal que la acción humana pasa a ser irrelevante? “*Tenga en cuenta que, si la discusión interactiva no se activa manualmente dentro de los tres días, el manuscrito pasará a la etapa interactiva automáticamente para evitar retrasos en el proceso de*

revisión por pares”, le comunicó AIRA al editor encargado, quien fue abandonando paulatinamente su interés, sin responder a las múltiples comunicaciones enviadas por el sistema.

Pero no solo el editor encargado perdió interés, sino que “de los 24 editores de revisión, todos miembros del cuerpo editorial de la revista, identificados por el sistema e invitados automáticamente, ninguno respondió nunca a la invitación, ni siquiera dando una respuesta negativa a participar” y, más aún, tampoco participó del proceso la editora en jefe de la revista. Esta nula participación por parte de editores asociados de la revista y de múltiples revisores contactados por el sistema, contrasta con los más de 2.700 artículos publicados en el último año por *Frontiers in Psychiatry*, lo que nos lleva a preguntarnos si realmente son humanos quienes están evaluando los artículos. La paradoja es que, aunque la respuesta explicitara la casi nula participación de la comunidad científica en esta nueva configuración de lo que se considere publicable y, por ende, científicamente válido, tal como señala Bridle, los procesos automatizados generan otro fenómeno, que está asociado al *sesgo de confirmación*, el cual

“...reajusta nuestra percepción del mundo para que se adapte mejor a la información automatizada, reafirmando aún más la validez de las soluciones computacionales, hasta el extremo de que podamos descartar por completo las observaciones incompatibles con el punto de vista de la máquina”. (Bridle, 2020, p. 35)

Habermas diría que ese reajuste de “nuestra percepción del mundo” del que nos habla Bridle es producto de la institucionalización del dominio técnico sobre las acciones sociales, que logra un nivel tal de aceptación y de legitimidad que, según Marcuse, no es irracional, sino que se expresa más bien como “una sumisión al aparato técnico que aumenta las comodidades de la vida y aumenta la productividad” (Marcuse, 1984, p. 186). Si bien ese aparato técnico está impulsado por los intereses de las fuerzas productivas, su poder transformador y su capacidad operativa es de tal magnitud que esa trama de intereses queda oculta, permanece subsumida, sin que se reflexione sobre esos intereses como tales y sin que sean confrontados con los intereses de otros grupos sociales.

“...como consecuencia de ello, nuevas oleadas de saber técnico irrumpen cada día en la vida social cogiéndola desprevenida, y los nuevos potenciales de un poder ampliado de disposición técnica hacen cada día más manifiesta la desproporción existente entre los resultados de una racionalidad técnica al máximo de su tensión productiva y unos fines perseguidos sin reflexión”. (Habermas, 2007, p. 147)

Y este mecanismo tiene un efecto de retroalimentación creciente. De hecho, si a la comunidad científica se le propusiera un sistema automatizado de revisión por pares, libre de humanos, probablemente una parte sustancial de la propia comunidad lo aceptaría, sin preguntarse cuáles serían las consecuencias para la propia práctica científica y la concepción de ciencia que los sistemas de IA crearían.

CONCLUSIONES

El análisis de este caso nos permitió acercarnos a qué tipo de interacción persona-cobot propone el *Artificial Intelligence Review Assistant* (AIRA) de la editorial Frontiers, al incorporar la inteligencia artificial a la gestión de la revisión por pares de un artículo. Ese pasaje de prácticas de carácter relacional a prácticas productivas industriales a gran escala se enmarca en procesos sociales más amplios, como la industrialización del artículo científico y la robotización de las industrias y de la vida cotidiana, los cuales proponen entornos colaborativos entre humanos y cobots que dan paso a la cobotización de diversas prácticas.

Por un lado, la tecnificación genera procesos productivos altamente industrializados que demandan la inversión de grandes capitales que lleva, a su vez, a una concentración aún mayor y a una escala de producción sin precedentes, que estimula aún más su retroalimentación. Este círculo productivo se torna el motor de un proceso de tecnificación de las acciones sociales asociadas, en este caso, a la revisión por pares, que tiende no solo a quebrar los acuerdos culturales que orientan las acciones, sino a anular la dialéctica interrogativa y reflexiva, lo cual genera un efecto recursivo en espiral que potencia su dominio, su legitimación y su reproducción.

En el caso de la cobotización de la revisión por pares, se crean entornos que tecnifican las acciones comunicativas, y esta tecnificación de las interacciones lleva a una disminución de la participación humana calificada, a una desarticulación de los valores y acuerdos creados por las propias comunidades científicas, que expulsa a los humanos del sistema. Esta cobotización de la revisión por pares, por el momento, pareciera requerir de la participación de un mínimo de personas para que el sistema pueda actuar en su nombre, de manera de desligar la responsabilidad, incluso legal (Frontiers, 2023a), de sus decisiones en los pocos humanos que interactúan, aunque podríamos estar presenciando la antesala de una completa automatización.

Entendemos que es un proceso que despliega una capacidad operativa tal que tiene la potencialidad de legitimar ciertas racionalidades científicas por sobre otras, de desvalorizar o excluir aquellas propuestas más innovadoras o creativas, condicionando el tipo de ciencias que se valide a partir de estos sistemas. En términos históricos, la racionalidad técnica irreflexiva no fue el único camino que tomaron las ciencias. No todos los entornos industrializados tendieron a instrumentalizar las ciencias y tecnificar sus prácticas, del mismo modo que no todos los entornos no industrializados tendieron a realizar ciencias más interrogativas, reflexivas o críticas (Martinovich, 2023). Por lo tanto, no es la propuesta argumentar en contra de la tecnología, dado que aún no es una entidad en sí misma, sino proponer una interacción más reflexiva que nos permita traer a un primer plano la trama social de intereses y las estrategias puestas en juego por la racionalidad técnica, para abordar las dinámicas que generan en el conjunto social, que impactan en las formas de validación y puesta en circulación del conocimiento generado por las diversas comunidades científicas.

CONFLICTO DE INTERESES

Los autores declaran no tener vínculos o compromisos que condicionen lo expresado en el texto y que puedan ser entendidos como conflicto de intereses.

CONTRIBUCIÓN AUTORAL

Ambos autores contribuyeron en la conceptualización, el análisis formal, el desarrollo de la investigación, la metodología, la supervisión y validación de los resultados, la redacción del borrador original y la revisión de la versión final.

REFERENCIAS

- Aguado-López, E., & Becerril-García, A. (2021). El tiempo de la revisión por pares: ¿obstáculo a la comunicación científica? *Interciencia*, 46(2), 56–64.
- Alberts, B., Kirschner, M. W., Tilghman, S., & Varmus, H. (2014). Rescuing US biomedical research from its systemic flaws. *Proceedings of the National Academy of Sciences*, 111(16), 5773–5777. <https://doi.org/10.1073/pnas.1404402111>
- Andersen, M. Z., Fonnes, S., & Rosenberg, J. (2021). Time from submission to publication varied widely for biomedical journals: A systematic review. *Current Medical Research and Opinion*, 37(6), 985–993. <https://doi.org/10.1080/03007995.2021.1905622>
- Bali, M. (2015). A new scholar's perspective on open peer review. *Teaching in Higher Education*, 20(8), 857–863. <https://doi.org/10.1080/13562517.2015.1085857>

- Bauer, W., Schlund, S., Hornung, T., & Schuler, S. (2018). Digitalization of industrial value chains? A review and evaluation of existing use cases of industry 4.0 in Germany. *LogForum*, 14(3), 331–340.
- Biagioli, M. (2002). From book censorship to academic peer review. *Emergences: Journal for the Study of Media & Composite Cultures*, 12(1), 11–45. <https://doi.org/10.1080/1045722022000003435>
- Björk, B.-C., & Solomon, D. (2013). The publishing delay in scholarly peer-reviewed journals. *Journal of Informetrics*, 7(4), 914–923. <https://doi.org/10.1016/j.joi.2013.09.001>
- Borboni, A., Reddy, K. V. V., Elamvazuthi, I., AL-Quraishi, M. S., Natarajan, E., & Azhar Ali, S. S. (2023). The Expanding Role of Artificial Intelligence in Collaborative Robots for Industrial Applications: A Systematic Review of Recent Works. *Machines*, 11(1), Article 1. <https://doi.org/10.3390/machines11010111>
- Bridle, J. (2020). *La nueva edad oscura: La tecnología y el fin del futuro*. Debate. <https://incom.uab.cat/portalcom/books/libros/la-nueva-edad-oscura-la-tecnologia-y-el-fin-del-futuro-james-bridle-2020-2/>
- Campanario, J. M. (1998a). Peer Review for Journals as it Stands Today—Part 1. *Science Communication*, 19(3), 181–211. <https://doi.org/10.1177/1075547098019003002>
- Campanario, J. M. (1998b). Peer Review for Journals as it Stands Today—Part 2. *Science Communication*, 19(4), 277–306. <https://doi.org/10.1177/1075547098019004002>
- Cole, S., Cole, J. R., & Simon, G. A. (1981). Chance and Consensus in Peer Review. *Science*, 214(4523), 881–886. <https://doi.org/10.1126/science.7302566>
- Daniel, H.-D. (1993). *Guardians of Science: Fairness and Reliability of Peer Review*. VCH.
- Danks, D., & London, A. J. (2017). Algorithmic bias in autonomous systems. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, 4691–4697. <https://doi.org/10.24963/ijcai.2017/654>
- Das, S., Dey, A., Pal, A., & Roy, N. (2015). Applications of Artificial Intelligence in Machine Learning: Review and Prospect. *International Journal of Computer Applications*, 115(9), 31–41.
- Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>
- Editors of Frontiers medical journals. (2015). *Manifesto of editorial independence of editors of Frontiers medical journals*. <https://doi.org/10.13140/RG.2.1.1744.0164>

- Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *The Academy of Management Review*, 14(4), 532–550. <https://doi.org/10.2307/258557>
- Ford, E. (2013). Defining and Characterizing Open Peer Review: A Review of the Literature. *Journal of Scholarly Publishing*, 44(4), 311–326. <https://doi.org/10.3138/jsp.44-4-001>
- Frontiers. (2023a). *Response to the Office of Science and Technology Policy (OSTP), Request for Information: National Priorities for Artificial Intelligence, Docket ID: OSTP-TECH-2023-0007*. Frontiers. https://blog.frontiersin.org/wp-content/uploads/2023/07/Frontiers_response_OSTP_RFI_AI_OSTP-TECH-2023.pdf
- Frontiers. (2020). Artificial Intelligence to help meet global demand for high-quality, objective peer-review in publishing—Science & research news | Frontiers. *Frontiers Science News*. <https://blog.frontiersin.org/2020/07/01/artificial-intelligence-to-help-meet-global-demand-for-high-quality-objective-peer-review-in-publishing/>
- Frontiers. (2023b). *Fee policy*. <https://www.frontiersin.org/about/fee-policy>
- Frontiers. (2023c). *Peer review*. <https://www.frontiersin.org/about/peer-review>
- Frontiers. (2023d). *What is the role of the Frontiers Handling Editor?* Help Center. <https://helpcenter.frontiersin.org/s/article/What-is-the-role-of-the-Frontiers-Handling-Editor>
- Fu, R., Huang, Y., & Singh, P. V. (2020). *AI and Algorithmic Bias: Source, Detection, Mitigation and Implications* (SSRN Scholarly Paper 3681517). <https://doi.org/10.2139/ssrn.3681517>
- Gadamer, H.-G. (2005). *Verdad y método* (11^a ed.). Ediciones Sígueme.
- Gadamer, H.-G. (2006). *Verdad y método II*. Ediciones Sígueme.
- Ghosal, T. (2019). Exploring the Implications of Artificial Intelligence in Various Aspects of Scholarly Peer Review. *Bulletin of IEEE Technical Committee on Digital Libraries*, 15(1).
- Habermas, J. (2002). *Teoría de la acción comunicativa II: Crítica de la razón funcionalista*. Taurus.
- Habermas, J. (2007). *Ciencia y técnica como “ideología”* (5^a ed.). Tecnos.
- Heelan, P. A. (1998). The scope of hermeneutics in natural science. *Studies in History and Philosophy of Science Part A*, 29(2), 273–298. [https://doi.org/10.1016/S0039-3681\(98\)00002-8](https://doi.org/10.1016/S0039-3681(98)00002-8)
- Hentout, A., Aouache, M., Maoudj, A., & Akli, I. (2019). Human–robot interaction in industrial collaborative robotics: A literature review of the decade 2008–2017. *Advanced Robotics*, 33(15–16), 764–799. <https://doi.org/10.1080/01691864.2019.1636714>

- Horrobin, D. F. (1990). The Philosophical Basis of Peer Review and the Suppression of Innovation. *JAMA*, 263(10), 1438–1441.
<https://doi.org/10.1001/jama.1990.03440100162024>
- Kronick, D. A. (1990). Peer Review in 18th-Century Scientific Journalism. *JAMA*, 263(10), 1321–1322. <https://doi.org/10.1001/jama.1990.03440100021002>
- Larivière, V., Haustein, S., & Mongeon, P. (2015). The oligopoly of academic publishers in the digital era. *PLOS ONE*, 10(6), e0127502.
<https://doi.org/10.1371/journal.pone.0127502>
- Li, D., & Agha, L. (2015). Big names or big ideas: Do peer-review panels select the best science proposals? *Science*, 348(6233), 434–438.
<https://doi.org/10.1126/science.aaa0185>
- Marcus, A. (2012). Does anesthesiology have a problem? Final version of report suggests Fujii will take retraction record, with 172. *Retraction Watch*.
<https://tinyurl.com/4v23yhs4>
- Marcuse, H. (1984). *El hombre unidimensional*. Ariel.
- Martinovich, V. (2020). Indicadores de citación y relevancia científica: Genealogía de una representación. *Dados*, 63(2), e20190094.
<https://doi.org/10.1590/001152582020218>
- Martinovich, V. (2023). *Las revistas científicas como objetos narrativos de las ciencias* [Preprint]. <https://doi.org/10.1590/SciELOPreprints.7148>
- Mathis, J. F. (1982). An industrial scientist's perspective on scientific peer review. *Fundamental and Applied Toxicology*, 2(6), 280–282.
[https://doi.org/10.1016/S0272-0590\(82\)80005-5](https://doi.org/10.1016/S0272-0590(82)80005-5)
- Mehta, A. (2019). 75% of European spending on scientific journals goes to ‘big five’ publishers. *Chemistry World*. <https://www.chemistryworld.com/news/75-of-european-spending-on-scientific-journals-goes-to-big-five-publishers/4010616.article>
- Moxham, N., & Fyfe, A. (2018). The Royal Society and the prehistory of peer review, 1665–1965. *The Historical Journal*, 61(4), 863–889.
<https://doi.org/10.1017/S0018246X17000334>
- Osika, G. (2022). Humanistic and Social Dimensions of Cobotization in the Context of Implementation Industry 5.0. *Scientific Papers of Silesian University of Technology. Organization & Management*, 165, 259–274.
<https://doi.org/10.29119/1641-3466.2022.165.19>
- Peng, H., Romero, D. M., & Horvát, E.-Á. (2022). Dynamics of cross-platform attention to retracted papers. *Proceedings of the National Academy of Sciences*, 119(25), e2119086119. <https://doi.org/10.1073/pnas.2119086119>
- Shuai, X., Rollins, J., Moulinier, I., Custis, T., Edmunds, M., & Schilder, F. (2017). A Multidimensional Investigation of the Effects of Publication Retraction on

- Scholarly Impact. *Journal of the Association for Information Science and Technology*, 68(9), 2225–2236. <https://doi.org/10.1002/asi.23826>
- STM. (2021). *STM Global Brief 2021 – Economics & Market Size* (STM Report Supplement). https://www.stm-assoc.org/wp-content/uploads/2021_10_19_STM_Global_Brief_2021_Economics_and_Market_Size-1.pdf
- Thelwall, M. (2019). *Artificial Intelligence, Automation and Peer Review*. <https://www.semanticscholar.org/paper/Artificial-Intelligence%2C-Automation-and-Peer-Review-Thelwall/c3af18797fa9dadd45c53729c7a64fbdb8ed4936>
- Van Noorden, R. (2011). Science publishing: The trouble with retractions. *Nature*, 478(7367), Article 7367. <https://doi.org/10.1038/478026a>
- Van Noorden, R. (2021). Journals adopt AI to spot duplicated images in manuscripts. *Nature*, 601(7891), 14–15. <https://doi.org/10.1038/d41586-021-03807-6>
- Wessel, M. R. (1982). The “state of the science” conference: A new approach to scientific decision making. *Fundamental and Applied Toxicology*, 2(6), 283–288. [https://doi.org/10.1016/S0272-0590\(82\)80006-7](https://doi.org/10.1016/S0272-0590(82)80006-7)
- Wolfram, D., Wang, P., Hembree, A., & Park, H. (2020). Open peer review: Promoting transparency in open science. *Scientometrics*, 125(2), 1033–1051. <https://doi.org/10.1007/s11192-020-03488-4>
- Yin, R. K. (2003). *Case Study Research: Design and Methods* (3^a ed.). SAGE Publication.
- Zuckerman, H., & Merton, R. K. (1971). Patterns of evaluation in science: Institutionalisation, structure and functions of the referee system. *Minerva*, 9(1), 66–100.

Este preprint fue presentado bajo las siguientes condiciones:

- Los autores declaran que son conscientes de que son los únicos responsables del contenido del preprint y que el depósito en SciELO Preprints no significa ningún compromiso por parte de SciELO, excepto su preservación y difusión.
- Los autores declaran que se obtuvieron los términos necesarios del consentimiento libre e informado de los participantes o pacientes en la investigación y se describen en el manuscrito, cuando corresponde.
- Los autores declaran que la preparación del manuscrito siguió las normas éticas de comunicación científica.
- Los autores declaran que los datos, las aplicaciones y otros contenidos subyacentes al manuscrito están referenciados.
- El manuscrito depositado está en formato PDF.
- Los autores declaran que la investigación que dio origen al manuscrito siguió buenas prácticas éticas y que las aprobaciones necesarias de los comités de ética de investigación, cuando corresponda, se describen en el manuscrito.
- Los autores declaran que una vez que un manuscrito es postado en el servidor SciELO Preprints, sólo puede ser retirado mediante solicitud a la Secretaría Editorial deSciELO Preprints, que publicará un aviso de retracción en su lugar.
- Los autores aceptan que el manuscrito aprobado esté disponible bajo licencia [Creative Commons CC-BY](#).
- El autor que presenta el manuscrito declara que las contribuciones de todos los autores y la declaración de conflicto de intereses se incluyen explícitamente y en secciones específicas del manuscrito.
- Los autores declaran que el manuscrito no fue depositado y/o previamente puesto a disposición en otro servidor de preprints o publicado en una revista.
- Si el manuscrito está siendo evaluado o siendo preparando para su publicación pero aún no ha sido publicado por una revista, los autores declaran que han recibido autorización de la revista para hacer este depósito.
- El autor que envía el manuscrito declara que todos los autores del mismo están de acuerdo con el envío a SciELO Preprints.