

Editorial: Inteligencia artificial: cuatro planteos filosóficos Artificial Intelligence: Four Philosophical Approaches

TAMARA CHIBEY-RIVAS¹

Universidad Nacional de Córdoba, Córdoba, Argentina
tchibey@gmail.com

ROBERTO RUBIO²

Universidad Alberto Hurtado, Santiago, Chile
rorubio@uahurtado.cl

El presente número recibe su impulso inicial del coloquio “Inteligencia Artificial: desafíos presentes y futuros”. Dicho coloquio, realizado el 7 de diciembre de 2023, fue organizado por Tamara Chibey, Roberto Rubio y Martín Gendler y auspiciado por el Centro de Estudios Mediales de la Universidad Alberto Hurtado. El evento reunió a especialistas nacionales y extranjeros, quienes abordaron filosóficamente las implicancias inmediatas y futuras de los avances recientes en el campo de la inteligencia artificial generativa.

Entre las principales preguntas que se plantearon y desarrollaron en el coloquio, destacan las siguientes: 1) Al interactuar con máquinas operadas por sistemas de inteligencia artificial, ¿corresponde hacerlo desde una actitud intencional, orientada hacia una relación de yo-tú, o más bien hay que orientar nuestra comunicación con ellas a partir de la idea de que son máquinas diseñadas para fines específicos?; 2) ¿De qué manera se pueden limitar los sesgos de género en sistemas de IA que ofrecen asistencia en la toma de decisiones? ¿Cómo asegurar

CÓMO CITAR ESTE ARTÍCULO:

En APA: Chibey-Rivas, T. & Rubio, R. (2024). Inteligencia artificial: cuatro planteos filosóficos. *Resonancias*, (17), 81-84. DOI: 10.5354/0719-790X.2024.75334

En MLA: Chibey-Rivas, T., and R. Rubio. “Inteligencia artificial: cuatro planteos filosóficos.” *Resonancias*, n.º 17, julio de 2024, pp. 81-84. DOI: 10.5354/0719-790X.2024.75334

¹ Cursa un Doctorado en Filosofía en la Facultad de Filosofía y Humanidades de la Universidad Nacional de Córdoba en Argentina y es becaria de Conicet en el Instituto de Humanidades de la misma universidad. También es Investigadora Asociada del Centro de Estudios Mediales de la Universidad Alberto Hurtado en Chile. ORCID 0000-0003-3670-4598.

² Director del Centro de Estudios Mediales. ORCID 000-0002-2482-9062.

allí criterios éticos de transparencia, justicia, privacidad y responsabilidad desde una perspectiva de género?; 3) ¿Qué consecuencias políticas tiene el actual salto de escala en el procesamiento de datos, un salto que va del uso de datos directos e indirectos en smartphones y plataformas digitales a la creación y utilización de bancos de datos para el entrenamiento de grandes modelos de lenguaje y grandes modelos de agencia?; 4) ¿Qué efectos socio-políticos conlleva caracterizar los sistemas de IA mediante rasgos humanos como la cognición, por ejemplo? ¿Cómo pensar los ensamblajes computacionales operantes en los procesos de datificación social y propagación de plataformas y algoritmos?

El carácter filosófico del encuentro quedó plasmado en los siguientes rasgos de las ponencias y de las discusiones que les siguieron: la aproximación sociotécnica al estudio de los sistemas de IA, un profundo interés y preocupación por los alcances éticos y políticos de la implementación masiva de dichos sistemas, y un cuestionamiento de fondo dirigido a la imbricación entre técnica y cultura como clave para el análisis de la así llamada inteligencia artificial.

A partir de las ponencias y debates mencionados se redactaron los cuatro artículos que componen el presente número. El primer artículo, titulado “La IA entre nosotros: ¿qué debemos hacer?”, aborda las formas en las que los seres humanos interactúan con los sistemas con inteligencia artificial actualmente existentes. La autora Diana Pérez analiza las formas de interacción que están a la base de la consideración de tales sistemas como “inteligentes”, tomando en cuenta el rol programático del “test de Turing”. Con vistas a ello, destaca la siguiente tensión: “la IA es un objeto de diseño humano que nos obliga a no considerarlo un artefacto, una máquina, sino a considerarlo uno de nosotros”. Habría que desarrollar, entonces, frente a la IA, una actitud de diseño y a la vez una actitud intencional. Sin embargo, sostiene Pérez: “no podemos hacer las dos cosas simultáneamente”.

La autora divide en dos grandes grupos los avances actuales de la IA: por una parte, la creación de IAs que simulan las conductas de los seres humanos en la interacción con ellos en el entorno cotidiano social; por otra parte, máquinas más inteligentes que las mentes humanas que ayuden de manera eficiente y rápida en tareas y entornos específicos. Con vistas a estos avances, Pérez advierte los siguientes desafíos: en primer lugar, incentivar la regulación de los desarrollos tecnológicos relacionados con la inteligencia artificial; y en segundo lugar, limitar de hecho ciertos desarrollos de esa tecnología, bajo el entendido de que pueden generar un efecto adverso masivo.

En el artículo “Una mirada crítica a la ética de la IA: de preocupaciones emergentes a principios orientadores”, Gabriela Arriagada ofrece una reconstrucción crítica de la interdisciplina denominada “Ética de la inteligencia artificial”. La reconstrucción comienza por las preocupaciones que dieron impulso a los

debates iniciales, luego considera algunas críticas de peso relevantes para el campo de la ética de la IA y, por último, ofrece tres principios orientadores para su desarrollo actual y futuro.

Con respecto a las preocupaciones que dieron inicio a la ética de la IA en la última década, la autora cita el influyente artículo de Mittelstadt *et al.* (2016). En dicho artículo se identifican seis preocupaciones éticas: tres respecto a la evidencia que ofrecen los sistemas de IA (evidencia inconclusa, inescrutable y equivocada) y las otras tres referentes a los resultados injustos, los efectos transformativos y la trazabilidad, respectivamente.

Entre las principales críticas hacia el campo de la ética de la IA, Arriagada destaca los planteamientos acerca de la falacia aditiva y de la inutilidad de los principios. La falacia aditiva consiste en asumir erróneamente que la ética de la IA se alcanza añadiendo sin más enfoques tradicionales de la ética al estudio sobre la inteligencia artificial. Por su parte, la crítica acerca de la inutilidad de los principios fue planteada por Mittelstadt, y consiste en advertir sobre la estrategia de crear principios ambiguos y generales que no se hacen cargo de las correspondientes discusiones normativas y políticas. Teniendo en cuenta esas críticas, la autora ofrece los siguientes principios para orientar el desarrollo una ética aplicada a la inteligencia artificial: una mirada sociotécnica que reconoce la interdependencia entre IA y contextos sociales, estructuras regulatorias y educativas para una integración efectiva de la ética y la necesidad de estándares profesionales que den marco a la práctica ética.

El tercer artículo pertenece a Martín Gendler y se titula “Datificación Social e Inteligencia Artificial: ¿hacia un nuevo ‘salto de escala’?”. El autor analiza los desarrollos de la inteligencia artificial y su relación con los procesos de datificación social. Para ello, ofrece una breve genealogía de tales procesos en conjunto con un análisis de las distintas fases del desarrollo de IA. Entre los principales momentos señalados por dicha genealogía, cabe destacar los siguientes: a) el sistema de recolección, almacenamiento, procesamiento, perfilamiento y aplicación (R.A.P.P.A) como la lógica principal de la operatoria de las principales empresas de Internet y de los objetos técnicos digitales (OTD); b) la industria 4.0 y la correspondiente configuración de objetos técnicos digitales, a los que cabría llamar OTD 4.0. Gendler advierte una relación de doble reciprocidad entre el proceso de datificación y los desarrollos de IA, “dado que el primero opera como condición de base para el entrenamiento y desarrollo de los segundos, mientras que estos, fomentan, optimizan y perfeccionan los diversos mecanismos de recolección, almacenamiento y desarrollo de datos, luego utilizados para los procesos de perfilado y aplicación”.

La propuesta genealógica se despliega en una doble perspectiva: por una parte, Gendler analiza los avances realizados en los sistemas de IA como un elemento clave para potenciar la datificación; por otra parte, nos da a pensar acerca de la

posibilidad de un nuevo “salto de escala” en los actuales procesos de personalización y modulación algorítmica.

El último artículo pertenece a Javier Blanco y lleva por título “Cognición computacional”. Blanco plantea que la noción de cognición computacional, tomada de William Rapaport, es más precisa que la de inteligencia artificial, especialmente para caracterizar los sistemas de *machine learning*.

Blanco advierte acerca de la confusión que suele dominar actualmente las discusiones y análisis acerca de las tecnologías de IA, y propone distinguir entre diversos niveles y tipos de mediaciones computacionales. De ese modo, argumenta, se evitaría “una frecuente falacia de composición que consiste en tomar propiedades específicas de algún subconjunto determinado, en particular a los sistemas producidos por *machine learning*, como aplicables a cualquier sistema computacional”.

Con respecto a la tecnología del *machine learning*, Blanco advierte sus limitaciones, en particular respecto al margen de error en los resultados que se ofrecen, dado el carácter estadístico de los métodos de procesamiento. Sin embargo, no descarta un eventual salto conceptual que traiga consigo otro tipo de programa de aprendizaje maquínico que se automejore. Blanco entiende al *machine learning* como un método de metaprogramación situado en lo que, con Dijkstra, denomina el interjuego entre programación y computación. Este interjuego posibilita un paisaje de oportunidades evolutivas que contribuye a desarrollar y renovar los acoplamientos cognitivos. La propuesta de Blanco apunta a favorecer la coevolución entre seres humanos e IA o, más en general, entre humanos y programas.

A fin de profundizar en la noción cognición computacional, el autor problematiza la distinción entre cognición computacional y cognición humana planteada por Brian Cantwell Smith. Según esto, tal distinción radica en la capacidad de “juicio”, la cual da cuenta de un compromiso existencial humano. Frente a ello, Blanco se pregunta “cómo se establece hoy un compromiso existencial con un mundo poblado de clasificadores estadísticos, cómo estos clasificadores integrados ya a la cotidianidad, a la toma de decisiones, a la cultura, afectan y transforman la capacidad ¿humana? de juicio y sus alcances”.

Para finalizar, deseamos agradecer a las autoras y los autores de los artículos, como así también al Centro de Estudios Mediales, institución que hizo posible el coloquio que dio impulsos para la redacción de este Dossier: *Last but not least*; vaya nuestro sincero agradecimiento a *Resonancias. Revista de Filosofía* por invitarnos a organizar y editar el presente Dossier.

