

# IA: ¿arrancar para adelante?

## ANÁLISIS

Felipe Irarrázabal



**COMO ESTE TEMA DE LA INTELIGENCIA ARTIFICIAL (IA) ES —AL MENOS PARA MÍ— PRÁCTICAMENTE INASIBLE, PARTO CON DOS RELATOS MAGNÍFICOS** y dos anécdotas menores y termino comentando un libro.

El primer relato se refiere a Aprendiz de Brujo, poema de Goethe recogido por Disney en 1940, sobre un aprendiz que decide ensayar sus poderes con una escoba para que acarree agua (en YouTube). Luego se duerme y, al despertar, descubre que toda la habitación está inundada. Ante la imposibilidad de revertir el embrujo, decide trozar la escoba con un hacha. Al rato, de las astillas surgen nuevas escobas y ahora es un regimiento trasladando agua y aumentando la inundación hasta que vuelve el brujo para revertir el entuerto, ante la confesión del aprendiz: “He invocado a los espíritus y no me desprendo de ellos”.

La película es 2001: Odisea en el Espacio, dirigida por Kubrick y estrenada en 1968 (en Max), y trata de un viaje espacial a Júpiter, en una nave manejada por un computador infalible. El computador habla como un humano y es representado por un ojo rojo omnipotente. El computador inventa un desperfecto y luego de que los humanos discuten su desconexión, asesina a uno de ellos. Cuando el tripulante sobreviviente se dispone a desconectar el computador, este le dice que se equivocó, que lo perdone y que “tiene sentimientos”. Luego de la desconexión, aparece una grabación que le comunica al astronauta que solo el computador sabía de la misión.

Respecto a las anécdotas, la primera es sobre el Centro de Competencia de la Universidad Adolfo Ibáñez que dirijo. Su página web cuenta con un chatbot cuya base de conocimiento incluye los conceptos vertidos en un glosario de derecho de la competencia elaborado por nosotros. Nuestro desafío es que el chatbot se atenga al glosario y no invente. Para testear esto, le pregunto cuál es la conexión entre libre competencia y la felicidad (algo que no está en el glosario), y su respuesta es tajante: no tengo información para responder esa pregunta. Meses después le hago la misma pregunta y su respuesta me sorprende: la economía de mercado implica precios bajos, aumento de productos e innovación y eso debiera hacer aumentar la felicidad. Un tiempo después insisto en el mismo tema y la

respuesta es aún más articulada y completa. Le pregunto, luego, por la relación entre la economía centralmente planificada y la infelicidad, y el chatbot vuelve a limitarse y decirme que no puede responder.

La otra anécdota se refiere a que a minutos de realizar una prueba de alternativas a través de Google Forms a un curso de pregrado, le pido a ChatGPT que me las responda. El formulario tiene triquiñuelas. IA las responde en segundos. Todas buenas.

Me concentro ahora en el último libro de Harari (Ne-

**La IA puede minar a las democracias porque centraliza mucho poder en un puñado de empresas digitales, que enfrentan bajos mecanismos de autocorrección (...) nos satura con simples y atractivos relatos, aunque muchas veces falsos”.**

xus), que ya aparece liderando el ranking de libros de “El Mercurio”. Harari, luego de reconocer que su libro podría haber sido escrito por IA, aclara que “la IA no es una herramienta, es un agente” porque “es la primera tecnología de la historia —en contraste con la escritura, la imprenta y la radio/televisión— que puede tomar decisiones y generar nuevas ideas por sí misma”. Insiste en lo mismo: “IA puede aprender por sí mismo cosas que ningún ingeniero humano ha programado, y puede decidir cosas que ningún ejecutivo humano ha previsto”.

Harari nos advierte que el actual manejo lingüístico de algunos modelos de IA es superior al humano, aunque aún no generan conciencia. Sin embargo, estos modelos pueden entablar relaciones íntimas con personas y después usar el poder de esa intimidad para influir en noso-

tros, sin que eso implique haber desarrollado sentimientos propios. Esa capacidad, en parte, se basa en el conocimiento que las empresas tecnológicas tienen de nosotros, producto de las interacciones con el celular.

La IA puede minar a las democracias porque centraliza mucho poder en un puñado de empresas digitales, que enfrentan bajos mecanismos de autocorrección (en contraste con el control que tradicionalmente han ejercido las universidades, los medios de comunicación y el poder judicial). Además, nos satura de información, alterando nuestros ciclos de descansos, con simples y atractivos relatos, aunque muchas veces falsos. La verdad, en cambio, es compleja y difícil de develar.

Harari advierte que podríamos enfrentarnos, si los avances de IA siguen a galope firme, al fin de la historia de los humanos y su reemplazo por una superinteligencia algorítmica, con o sin conciencia, pero siempre altamente manipuladora. Esa nueva forma de organización podría aumentar los riesgos de una guerra a escala planetaria, además de impactar los empleos de manera vertiginosa.

Europa ha estado reaccionando, con leyes digitales, de datos y de IA. Estados Unidos, en cambio, se ha concentrado en casos de libre competencia, pero también está discutiendo proyectos de ley para aumentar los niveles de vigilancia. China, por su parte, busca desarrollar su propia tecnología, y prohíbe el uso en su país de Google, YouTube, Facebook y Wikipedia.

Difícil saber qué hacer. Me temo que darle la espalda a IA, y pensarnos como si siguiéramos navegando en el siglo XX, no parece ser una estrategia ni razonable ni posible. Los beneficios de IA son profundos e insondables. Quizás haya que “arrancar para adelante”, incorporando IA lo más posible en nuestros procesos, pero siempre testeando su uso por humanos. Más que entrar a redactar proyectos de ley “copy/paste”, el Estado y las universidades deberían destinar ingentes recursos a estudiar el fenómeno, en especial sus riesgos asociados. No vaya a ser que terminemos como el aprendiz de brujo o el astronauta solitario.

P.D.: Le pido a IA que me comente esta columna y recibo elogios. Le pido críticas, y me dispara cuatro —bien interesantes—, que las dejo para después.