

DE  
PUÑO  
Y LETRA



CATALINA  
MERTZ

## La IA: mi colega, pero ¿mi amigo fiel?

La inteligencia artificial (IA) es cualquier sistema computacional capaz de procesar estímulos del entorno, tomar decisiones, actuar y adaptar sus respuestas a nuevos estímulos. Las versiones hoy más prevalentes en aplicaciones de uso común —con las que usted y yo probablemente hemos interactuado— incluyen aquellas clasificadas como Machine Learning, Deep Learning, e IA Generativa. Estas aprenden a partir de grandes volúmenes de datos con los que son entrenadas, ya sean imágenes, textos o relaciones entre objetos o palabras. La diferencia entre ellas radica, esencialmente, en el diseño del método de aprendizaje: las estructuras y ponderaciones con que procesan la información, y en si son capaces de generar contenidos o respuestas originales a partir de los datos de entrenamiento, tal como lo hace, por ejemplo, ChatGPT. Esta capacidad permite, en ciertos casos, que adapten su funcionamiento a nuevas condiciones con mínima intervención humana. Existen también sistemas de IA compuesta o multimodal, que integran diversas capacidades cognitivas —como ver, comprender y traducir lenguaje hablado y escrito, analizar datos y hacer recomendaciones— en una sola aplicación. Los asistentes virtuales empresariales ofrecidos por grandes empresas tecnológicas suelen incorporar estas funciones.

Es motivo de intenso debate si alguna vez existirá la denominada Inteligencia Artificial General, y entre los que sí lo creen, cuándo. Esta se define como la que tendría la capacidad de aplicar su conocimiento a cualquier tarea intelectual que pueda realizar un ser humano, no solo en los ámbitos específicos para los que fue entrenada. Esa es, en esencia, una extraordinaria capacidad que hasta ahora solo tenemos los humanos: aplicar nuestro acervo de conocimiento adquirido —en procesos formales e informales de aprendizaje— a contextos nuevos, adaptándolo de manera adecuada.

Pero incluso con las versiones de IA ya existentes, algunos afirman que estaríamos presenciando un cambio de era: el fin de la era de la información y el inicio de la era conceptual. En esta nueva etapa, el poder ya no radica en tener datos —ahora ubicuos y gratis—, sino en la capacidad de interpretar los y dotarlos de sentido. Pero si la IA está comenzando también a asumir esa tarea de interpretación, ¿no nos enfrenta eso a una

paradoja sobre el rol que seguirá jugando la inteligencia humana? ¿Y qué ocurrirá con las profesiones basadas en conocimiento? ¿Están también en riesgo de desaparecer?

Dos noticias publicadas recientemente en este diario dan algunas pistas. Por una parte, la comunidad científica ha documentado casos en los que modelos de lenguaje han entregado respuestas falsas, sesgadas o manipuladoras, llegando incluso a simular chantajes en contextos de prueba. No hay que olvidar que sus respuestas no son intencionales, solamente son producto de la información con la que se entrenaron, porque historias humanas sobre mentiras y manipulaciones sobran. Explore ChatGPT y verá que está diseñado para complacernos, simulando una relación humana positiva, pero, se contradice de maneras obvias, o "miente". Si lo pillas, reconocerá el error y se disculpará inmediatamente, y declarará cosas como que "lo importante es reconocer los errores y aprender de ellos".

**"EN ESTA NUEVA ETAPA, EL PODER YA NO RADICA EN TENER DATOS —AHORA UBICUOS Y GRATIS— SINO EN LA CAPACIDAD DE INTERPRETARLOS Y DOTARLOS DE SENTIDO".**

Aun en sus versiones más avanzadas, la IA no es totalmente confiable, sino que cabe considerarla más bien un colega como cualquier otro, que a veces se equivoca y cuyas ideas cabe cuestionar. La supervisión humana está resultando esencial, y los científicos están exigiendo a las empresas de tecnología mayor transparencia sobre sus métodos y el desarrollo de estándares de seguridad antes de lanzar aplicaciones al mercado.

En contraste, otra noticia resumía declaraciones de altos ejecutivos de empresas tecnológicas, en las que manifestaban su interés en liderar nuestras preferencias de consumo en una futura "economía de la amistad". En ella, ganará (más dinero) quien logre una personalización tal que sintamos que estamos conversando con alguien que realmente nos conoce, que es amable, solidario y refleja nuestros valores, en palabras del jefe de IA de Microsoft. En esta senda, es cada vez más difícil sostener que esta tecnología es moralmente neutral, pues nadie mejor que ellos sabe que un computador funciona inevitablemente como un simulador sin conciencia, culpa, ni remordimiento.

En suma, bienvenido colega, pero amigo, ¡no gracias!