



*Muchos de nosotros pensamos que la IA acaba de llegar a nuestras vidas, pero ¿realmente es así? ¿Desde hace cuánto tiempo lleva existiendo?*

Los avances (casi) diarios en el terreno de la Inteligencia Artificial están creando muchas dudas y miedos a su alrededor. Algunas de ellas hacen referencia a si estos sistemas son tan inteligentes como nos dicen, a cómo funcionan o qué riesgos supone su uso. Una serie de preguntas que hemos tratado de responder junto a Melanie Mitchell, una de las mayores expertas en la materia y autora del libro 'Inteligencia artificial. Guía para seres pensantes' (Capitán Swing).

La idea de construir una «máquina pensante» existe desde hace siglos, pero el campo de la IA comenzó en 1956 en un taller del Dartmouth College, en el que pioneros como John McCarthy, Marvin Minsky, Herbert Simon y Allen Newell idearon su nombre y sus objetivos generales.

*«La red se entrena tratando de predecir la siguiente palabra de un bloque de texto entre miles de millones o billones de bloques de texto»*

**Lo que sí que ha hecho es cambiar a lo largo del tiempo. ¿A qué nos referíamos en el pasado con IA y a qué hoy en día?**

Al principio, la mayoría de los investigadores de IA pensaban que estos sistemas podían programarse utilizando la lógica y las reglas concebidas por el ser humano. Pero hoy en día casi todos los sistemas de IA se crean utilizando redes neuronales entrenadas con enormes cantidades de datos. Una red neuronal es un programa informático inspirado en los cerebros biológicos, en el que neuronas simuladas se conectan entre sí mediante conexiones ponderadas. Esta se entrena cambiando los pesos de las conexiones entre neuronas, dependiendo de si la respuesta de la red a una entrada es «correcta» o «incorrecta». Los grandes modelos lingüísticos, como ChatGPT, son redes neuronales enormes (es decir, miles de millones o billones de conexiones ponderadas) que se entrenan principalmente con texto extraído de Internet. La red se entrena tratando de predecir la siguiente palabra de un bloque de texto entre miles de millones o billones de bloques de texto. El resultado se denomina «gran modelo lingüístico», ya que a partir de todo el entrenamiento, el sistema aprende la estructura del lenguaje generado por el ser humano.

**Escuchamos, vemos, leemos casi todos los días noticias sobre los avances en la IA. ¿En qué situación se encuentra ahora mismo? ¿Existe realmente una IA equiparable a la inteligencia del ser humano?**

Los sistemas de IA como ChatGPT pueden realizar muchas tareas que consideramos «inteligentes», como escribir ensayos, traducir entre idiomas, entablar conversaciones y resolver problemas matemáticos. Sin embargo, también tienen limitaciones no humanas, como no saber distinguir entre información verdadera y falsa o cometer a menudo errores difíciles de explicar. En resumen, estos sistemas parecen inteligentes en muchos aspectos, pero siguen sin ser fiables. Por último, hay que apuntar que la «inteligencia humana» no está bien definida. ¿Caminar por una acera llena de gente sin chocar con alguien es una forma de «inteligencia»? ¿Lo es ser capaz de arreglar un problema de fontanería? Ahora mismo, los robots de IA no pueden competir con los humanos en la mayoría de las tareas que requieren interacción física con el mundo real.

**¿Puede crear arte o incluso razonar como un humano?**

La IA puede crear «arte», pero es discutible que sea realmente creativa como lo son los humanos. En cuanto a la segunda cuestión, a menudo parece que los sistemas de IA pueden razonar como nosotros, pero luego cometen errores muy poco humanos que demuestran que sus procesos probablemente no sean similares a los nuestros.

**¿Cuánta de toda esta burbuja alrededor de la IA tiene que ver con el marketing y el dinero que mueve?**

No puedo cuantificarlo, pero creo que hay mucho bombo publicitario en torno a la IA creado por el marketing y la necesidad de las empresas de ganar dinero. Esto siempre ha sido así: el bombo publicitario de la IA ha existido desde los inicios de este campo en la década de los cincuenta.

*«Hay mucho bombo publicitario en torno a la IA»*

*Muchas personas y empresas han empezado a dejar algunas de sus decisiones en manos de la IA. ¿A qué riesgos se enfrentan? ¿Es realmente neutral como nos repiten muchas veces?*

Los sistemas de IA pueden cometer errores, es decir, tener «alucinaciones» en las que generan información falsa, y también pueden tener sesgos sutiles que han aprendido de sus datos de entrenamiento, lo que puede causar perjuicios en términos de raza, género y otros factores. Por ello, las personas y organizaciones que utilizan la IA necesitan algún tipo de sistema (que incluya auditores humanos) para asegurarse que los errores y los sesgos sutiles se detectan antes de que estas decisiones acaben perjudicando. Se trata de un problema difícil de resolver pero urgente.

**Aparte de esos puntos negativos, tiene otros muy positivos. ¿Cuáles son?**

Los sistemas actuales de IA pueden ser enormemente útiles cuando se usan para aumentar, que no sustituir, a la inteligencia humana. Estos sistemas pueden tener -y ya están teniendo-una enorme repercusión a la hora de ayudar a médicos, científicos, programadores informáticos y otros profesionales a realizar tareas concretas, así como a los usuarios en general en tareas como mejorar la escritura y ayudar a buscar en Internet. Siempre que las personas que utilicen estos sistemas sean conscientes de los tipos de errores que pueden cometer y de hasta qué punto se puede confiar en ellos, estos aumentarán nuestra propia inteligencia.

*«Se ha hecho demasiado hincapié en fomentar la innovación y demasiado poco en la seguridad y la equidad»*

**Uno de los retos actuales a los que nos enfrentamos en el terreno de la IA es la regulación. ¿Hacia dónde se tiene que dirigir?**

El principal compromiso de los reguladores es, por un lado, garantizar que los sistemas de IA sean seguros y justos y, por otro, que la normativa no sea tan estricta o engorrosa que dificulte la innovación de las empresas de IA o la creación de nuevas empresas. Ahora mismo quizá se ha hecho demasiado hincapié en fomentar la innovación y

demasiado poco en la seguridad y la equidad. También existe el deseo de garantizar que las decisiones de los sistemas de IA sean «explicables». De hecho, esa comprensión forma parte de la normativa de la UE sobre IA. Sin embargo, la forma en que se desarrollan los actuales sistemas de IA hace muy difícil que su comportamiento sea explicable, y nadie sabe aún cómo solucionarlo.

**¿Hasta dónde crees que va a evolucionar la IA? ¿Podrá llegar a multiplicar la inteligencia humana como dicen?**

Creo que los sistemas de IA «aumentarán» la inteligencia humana de muchas maneras -en ciencia, medicina, salud pública y muchas otras áreas-, al igual que lo han hecho otras tecnologías en el pasado (por ejemplo, los ordenadores, Internet, etcétera).

**Carlos Madrid en [ethic.es](https://ethic.es)**