

¿Qué peligros presenta la IA?

ALBERTO PRIETO ESPINOSA

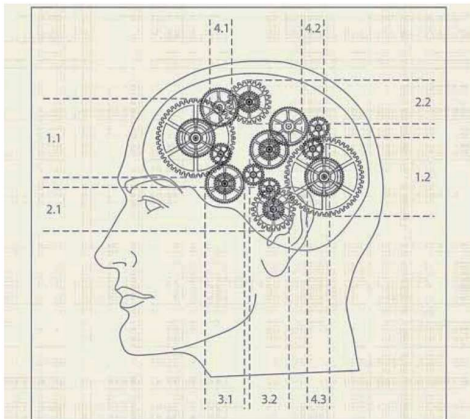
En estos días en redes sociales circulan imágenes y videos de Donald Trump y Kamala Harris creadas con aplicaciones de Inteligencia Artificial (IA) que parecen totalmente reales si no fuera por el tono jocoso y de chanza de lo que transmiten. Este fenómeno genera temor sobre los límites de la tecnología, pero como señaló el filósofo R. W. Emerson, el conocimiento es un antídoto contra el miedo. La IA, como toda tecnología, presenta ventajas e inconvenientes que es importante comprender, y recapitando sobre ellas he redactado las siguientes reflexiones para facilitar su conocimiento.

Entre las ventajas de la IA se encuentran las siguientes:

- La IA analiza grandes volúmenes de datos de manera mucho más rápida y eficiente que los humanos, lo que permite decisiones más objetivas, justas y transparentes por estar basadas en evidencias, evitando conflictos de interés, corrupciones y sesgos.
- La IA, y en general la Informática, facilita la interacción entre personas de diferentes idiomas y contextos a través de redes sociales y aplicaciones de reconocimiento de voz y traducción automática.
- La IA está revolucionando áreas clave como la salud, la energía, el transporte y la educación, mejorando, por ejemplo, la gestión energética, los vehículos autónomos y la educación personalizada.
- Las IA generativas pueden crear arte, música, textos y diseños con mínima intervención humana.
- La IA permite personalizar productos y servicios, mejorando la interacción empresarial con los consumidores y satisfaciendo necesidades específicas, como en medicina personalizada o cuidado de personas dependientes.
- Gracias a su capacidad para generar diseños y modelos de manera rápida, la IA acelera el desarrollo de nuevos productos y reduce el tiempo de lanzamiento al mercado, incrementando la competitividad global.

Sin embargo, la IA también presenta riesgos reales, no especulativos, como los siguientes:

- La automatización de trabajos rutinarios o peligrosos por la IA requiere políticas de adaptación para las profesiones en transformación o en desaparición.
- Con mucha frecuencia el concepto de IA se utiliza fraudulentamente para promover servicios o productos que no contienen realmente IA. Además, grandes empresas exageran sus logros en IA para atraer inversiones y mejorar su valor en los mercados.
- La carrera por la superioridad en IA puede alterar el equilibrio de poder mundial, estableciendo nuevos estándares en diplomacia y defensa. Además, grandes plataformas tecnológicas y redes sociales, dirigidas por personas sin



control democrático, pueden acceder y explotar datos de miles de millones de personas; permitiéndoles inferir información detallada sobre nuestra vida y potencialmente manipular nuestro comportamiento de manera subliminal, llegándonos a conocer mejor que nosotros mismos.

- La IA podría transformar la gobernanza, desde la vigilancia inteligente hasta sistemas de votación seguros, planteando un dilema entre seguridad y derechos humanos, con riesgos de violaciones a la privacidad y control masivo.
- La IA facilita la creación de noticias, imágenes y videos falsos, con un gran potencial de desinformación que afecta a la opinión pública, la política y las elecciones. La suplantación y falsificación de personas debería considerarse como un delito tanto o más grave que, por ejemplo, falsificar dinero.
- Los sistemas de IA, como otros sistemas informáticos, deben construirse siguiendo las metodologías de la Ingeniería del Software para asegurar que sean eficientes, seguros y éticos. Esto es especialmente crucial en aplicaciones de IA sensibles o críticas, ya que pueden producir errores no intencionados o diseños no rigurosos o incluso mal intencionados, que no extremen las medidas de seguridad.
- Los sistemas con aprendizaje automático responden de acuerdo con la información con que aprendieron, por lo que reproducen los sesgos presentes en los datos con los que son entrenados, lo que puede dar lugar, por ejemplo, a textos sexistas o racistas, o en algoritmos sesgados hacia ciertas tendencias. En cualquier sistema con aprendizaje automático es de suma importancia la selección suficientemente completa de los patrones de entrenamiento ya que la cali-

dad de los resultados depende de ellos.

- En el desarrollo de una aplicación de IA participan muchas personas, y es difícil identificar responsabilidades por las decisiones automáticas que toman los sistemas. Las máquinas por sí solas no hacen nada, todo lo hacen de acuerdo con programas hechos por humanos; siempre tenemos que prever los resultados inmediatos y últimos de las posibles decisiones que programamos, lo cual no siempre es fácil.

Existen otros riesgos, pero son a largo plazo, entre los que se encuentran los derivados de conseguir la llamada 'singularidad tecnológica' dentro de la cual especulativamente se lograría una superinteligencia que podría monitorizar y controlarlo todo.

Los diseñadores de IA deben proteger sus productos de los ciberataques y usos maliciosos, y garantizar en lo posible que no tengan consecuencias negativas para las personas, ni escapen al control humano. Los avances tecnológicos deben impulsar el progreso, la igualdad y la prosperidad de toda la sociedad, no solo de unos pocos. Nuestro objetivo debe ser que los sistemas de IA operen con transparencia, estén sujetos a supervisión humana y sean susceptibles de evaluación y certificación por autoridades externas competentes.

El conocimiento nos permite los grandes avances científicos, pero también evaluar y prever futuras amenazas.

* Alberto Prieto es miembro de la Academia de Ciencias Matemáticas, Físico-Química y Naturales de Granada Profesor Emérito. Departamento de Ingeniería de Computadores, Automática y Robótica. Universidad de Granada