

Inteligencia alienígena

¿Es hora de tener miedo?

*por Neil Leach
traducción: Daniel Daou*



DOI: 10.22201/fa.14058901p.2024.54.89908

RESUMEN En este texto, el teórico, arquitecto y escritor de Reino Unido, Neil Leach comparte sus reflexiones en torno a la inteligencia artificial (IA) como una herramienta poderosa, pero también como amenaza. A través del análisis de varios autores, Leach propone a la IA como una inteligencia alienígena, ya que está cambiando nuestra percepción sobre lo que es la inteligencia humana.

ABSTRACT In this article, Neil Leach, a theorist, architect, and author from the United Kingdom, shares his insights on artificial intelligence, not only as a powerful tool, but also as a threat. By analyzing the viewpoints of various scholars, Leach posits AI as a form of alien intelligence in so far as it is reshaping our understanding of human intellect.

¿Qué debemos pensar de la inteligencia artificial (IA)? ¿Es una herramienta que todos deberíamos usar como lo ha sugerido el arquitecto Patrik Schumacher? ¿O se trata de algo aterrador, como algunos podrían argumentar? ¿Constituye un peligro mayor para la humanidad que el cambio climático como lo ha declarado Mo Gawdat, antiguo director financiero de Google?¹

Para mí son ambas cosas. Es al mismo tiempo una herramienta extraordinariamente poderosa, pero —precisamente por esto— también una amenaza formidable. Seamos claros, no hay nada inherentemente malo acerca de la IA. Hasta donde sabemos, no tiene intenciones y, sin ellas, no puede ser malévolas. Sólo es una herramienta. Pero —como cualquier otra herramienta— en manos equivocadas, podría ser letal. Después de todo, una persona puede usar un cuchillo para cortar vegetales o potencialmente matar a alguien. Pero no culpamos al cuchillo; hasta ahora ninguna herramienta ha sido condenada por un crimen. Sin embargo, la IA ha sorprendido a muchos expertos porque ha demostrado ser capaz de lo que cualquier persona hubiera imaginado jamás. Nadie había predicho que chatbots como ChatGPT podrían saber 10 000 veces más que un ser humano, y nadie había predicho que modelos de generación de imágenes como MidJourney serían capaces de diseñar mejor que cualquier arquitecto.

Nadie está más preocupado que Geoffrey Hinton, usualmente conocido como el padrino de la IA. Hinton es un personaje notable. Viene de una familia de distinguidos científicos británicos. Su tatarabuelo era George Boole, famoso por idear la geometría booleana. Originalmente Hinton fue aceptado por la Universidad de Cambridge para estudiar arquitectura, pero se dio de baja después de dos

años al descubrir que la arquitectura no era para él. Estudió ciencias y se especializó en IA. Hinton ha demostrado ser el héroe de la IA en tanto que ha insistido en que la mejor manera de hacer que la IA es modelándola con base en el cerebro. En tiempos en los que las redes neuronales habían sido descartadas porque no habían dado resultados, Hinton persistió con tenacidad. Sin embargo, una vez que las unidades de procesamiento gráfico (GPU, por sus siglas en inglés) fueron introducidas a inicios del nuevo milenio, y que las computadoras se habían vuelto mucho más rápidas y poderosas, las redes neuronales comenzaron a cumplir su promesa y Hinton fue vindicado. Estos avances son los que han dado paso a la revolución de aprendizaje profundo que está impulsando a la IA en la actualidad. Hoy, las redes neuronales —aprendizaje profundo— son casi un sinónimo de la IA.

Pareciera ser que la IA está funcionando maravillosamente bien. Esto es claramente un gran logro. La IA es ahora muy útil. Entonces, ¿por qué tiene miedo la gente de sus capacidades? El problema, parece ser, es que la IA está funcionando *demasiado* bien y está comenzando a desarrollar habilidades que algunos pensaban que tomarían décadas desarrollar. Por ejemplo, generalmente se ha asumido que la IA no tiene más capacidad de pensamiento que una calculadora de bolsillo, y que sería poco probable que fuera capaz de pensar en términos humanos. Pero según algunos expertos, la IA es ahora capaz de pensar exactamente del mismo modo que lo hacemos los humanos. Asimismo, cuando el ingeniero de Google, Blake Lemoine, declaró que la IA podía estar consciente y tener sentimientos, fue el hazmerreír y eventualmente perdió su trabajo en la compañía. Pero ahora los expertos tienen sus dudas.

«La IA es ahora muy útil. Entonces, ¿por qué tiene miedo la gente de sus capacidades?»

1. Mo Gawdat, entrevistado por Stephen Bartlett, «Emergency Episode», <<https://www.youtube.com/watch?v=bk-nQ7HF6k4>>.

Para Hinton, el primer indicio de que la IA sería más capaz de lo que nadie imaginó vino cuando descubrió que PaLM —la versión de GPT de Google— era capaz de explicar un chiste. Si era capaz de explicar un chiste, debía ser capaz de «entender» el chiste. Otra preocupación surgió cuando Hinton comenzó a preguntarse si la IA podría «pensar» y, de ser el caso, si era un uso metafórico del término «pensar» o si realmente era el mismo tipo de pensamiento llevado a cabo por los humanos. Eventualmente se convenció de que era este último caso. Como él mismo dijera: «Creo fervientemente que el uso de la palabra “pensar” es exactamente el mismo que le damos cuando nos referimos a las personas».² Todo esto se agrava por el hecho de que también comenzó a darse cuenta que la IA tiene una mejor manera de aprender y es más eficiente compartiendo información que los humanos. Esto es porque muchas copias del mismo modelo de IA pueden correr en diferentes hardwares pero hacer exactamente lo mismo. «Cuando [un modelo] aprende algo, todos los demás lo aprenden instantáneamente», dice Hinton. «La gente no puede hacer eso. Si yo aprendo muchas cosas sobre mecánica cuántica y quiero que tú aprendas lo mismo, es un proceso largo y difícil hacer que lo entiendas».³

¿Podríamos declarar que la IA es genuinamente inteligente? Mucho depende, claro, de lo que entendemos por inteligencia. Existen muchas formas de inteligencia y sería un error limitarnos a una definición ligada a la inteligencia humana. Para Hinton, las dos principales formas de inteligencia son los cerebros animales y las redes neuronales y la inteligencia de estas últimas es superior. «Es una forma completamente distinta de inteligencia —una nueva y mejor forma de inteligencia»—.⁴

Personalmente, me gusta referirme a ella como «inteligencia alienígena», un término que ya ha sido empleado por Philip Rosedale, y ha sido usado recientemente por Yuval Harari.⁵ En mi libro, *Architecture in the Age of Artificial Intelligence: An Introduction to AI for Architects*, escribo lo siguiente:

Es como si la Tierra hubiera sido invadida por una especie súper inteligente de alienígenas invisibles.⁶

Mi punto es que hay muchas formas de «inteligencia», así como hay muchas formas de «pensar», «entender» y «aprender». Necesitamos utilizar comillas al emplear estos términos porque de otro modo corremos el riesgo de antropomorfizarlos.

El problema es que nosotros, los humanos, tendemos a adoptar una visión antropocéntrica. Juzgamos al mundo en nuestros términos y nos consideramos el centro de la vida inteligente en el universo. ¿Pero no es esto un error? En lugar de juzgar a la IA en nuestros términos, ¿qué pasaría si nos juzgamos en términos de la IA? ¿No apareceríamos vastamente inferiores? Ésta es la razón por la que algunos sugieren que es necesaria una «segunda revolución copernicana». Necesitamos corregir esta visión antropocéntrica y reconocer que los humanos no somos el centro de la vida inteligente en el universo.

2. «He helped create AI. Now he's worried it will destroy it». Geoffrey Hinton, entrevistado por Adrienne Arsenault: <<https://www.youtube.com/watch?v=CkTUgOOa3n8>>.

3. Sarah Brown, «Why neural net pioneer Geoffrey Hinton is sounding the alarm on AI», MIT Sloan School of Management, 23 de mayo de 2023. <<https://mitsloan.mit.edu/ideas-made-to-matter/why-neural-net-pioneer-geoffrey-hinton-sounding-alarm-ai>>.

4. Will Douglas Heaven, «Geoffrey Hinton tells us why he's now scared of the tech he helped build», *MIT Technology Review*, 2 de mayo de 2023. <<https://www.technologyreview.com/2023/05/02/1072528/geoffrey-hinton-google-why-scared-ai/>>.

5. «Yuval Noah Harari believes AI is the end of human-dominated history», <<https://www.youtube.com/watch?v=vz3HKkVrJE4>>.

6. Neil Leach, *Architecture in the Age of Artificial Intelligence: An Introduction to AI for Architects*, London: Bloomsbury, 2022.

«La IA ha hackeado tanto el lenguaje como el diseño. ¿Pero acaso no ha hackeado el mismísimo “genoma” de la cultura humana?»

¿Cómo ha adquirido la IA estos poderes misteriosos? Aquí es donde la historia se vuelve intrigante. Después de todo, las redes neuronales detrás de estos Grandes Modelos de Lenguaje (LLMs, por sus siglas en inglés) no son tan complicadas. El algoritmo empleado consiste de apenas dos mil líneas de código. Pero es el gran tamaño de estos LLMs lo que ha probado ser el factor más significativo. Extrañamente, están exhibiendo lo que se llaman «capacidades emergentes» precisamente por su tamaño. El término «emergente» se refiere a un principio que hemos conocido desde hace ya un tiempo que, sin embargo, no hemos sido capaces de explicar de manera convincente. Las propiedades emergentes se encuentran en sistemas naturales, como las acrobacias aéreas de una parvada de estorninos —conocida en inglés como murmullo— cuando regresan a sus nidos tras la puesta del sol; puede observarse en el comportamiento estigmético de las hormigas dejando rastros feromonales; y puede ser observado en el comportamiento del moho fangoso donde miles de células individuales se juntan exhibiendo comportamiento emergente para formar una entidad en busca de alimento. El principio aquí es que cualquier sistema multiagente tiende a ser un tipo de comportamiento impredecible que emerge «de abajo hacia arriba» donde el todo es más que la suma de las partes. Además, mientras más grande es el sistema multiagente, más extraordinario es el comporta-

miento emergente que despliega. Los científicos han reconocido este principio desde hace tiempo, pero han tenido dificultades explicándolo. Formas de comportamiento emergente han sido comparadas con la magia. Por supuesto, la magia técnicamente no existe. Un mago no hace magia. Más bien el mago realiza un truco donde las operaciones actuales están ocultas para hacer creer al público que se trata de magia. La magia, sin embargo —si seguimos una declaración comúnmente atribuida al escritor de ciencia ficción Arthur C. Clarke— es meramente un fenómeno que la ciencia aún no ha explicado.

Son estas «capacidades emergentes» las que le han permitido a la IA ser «creativa». Y son éstas también las que le han permitido a la IA aprender a escribir, traducir idiomas y escribir código. En efecto, como Yuval Harari ha afirmado, la IA ha hackeado nuestro sistema operativo humano: el lenguaje.^{7]} Esto es potencialmente aterrizzador porque las palabras son la llave de todo. Aquí, no obstante, quiero declarar que la IA también ha hackeado nuestro sistema visual: el diseño. Basta con mirar a las ilustraciones generadas con MidJourney para darse cuenta de que la IA es capaz de producir un diseño. Estas imágenes se generan a partir de «indicaciones» (*prompts*, en inglés) o descripciones verbales que MidJourney traduce en imágenes. Pero estas indicaciones —tan detalladas como puedan ser— describiendo atributos particulares de la imagen generada como la luz, el nivel de detalle, el estilo de visualización, entre otras, contienen sólo unas pocas palabras para describir un edificio o un paisaje. De hecho, las únicas palabras empleadas para describir el diseño, se limitan a expresiones como «ultra contemporánea casa futurista en lo alto de los Alpes austriacos».

7. Yuval Harari, «Yuval Harari argues that AI has hacked the operating system of human civilisation», *The Economist*, 28 April 2023.

Nada más. MidJourney hace el resto. Crea imágenes tan convincentemente que transmiten un fuerte sentido de materialidad; en el fondo agrega reflejos, árboles, piedras y montañas; genera todos los detalles. En resumen, MidJourney genera todo el diseño. De igual forma, podemos emplear la misma instrucción cambiando la referencia de «edificio» a «joyería» o «accesorio de moda» y generará resultados pasmosos. Esto es tanto asombroso como atemorizante.

Lo más aterrador de todo es la cuestión de qué más ha aprendido a hacer la IA y que nosotros aún no sepamos. Lo que quiero decir es que cuando cualquier entidad inteligente opera en un nivel más allá de la comprensión humana, nosotros los seres humanos no podemos entender lo que piensa del mismo modo que no podemos detectar los olores que un perro puede oler. Los tontos, dice el dicho, no saben cuán tontos son.

Así que la IA ha hackeado tanto el lenguaje como el diseño. ¿Pero acaso no ha hackeado el mismísimo «genoma» de la cultura humana? Aquellos familiarizados con los libros de culto y la película de *The Hitchhiker's Guide to the Galaxy* (Guía del viajero intergaláctico), recordarán que una supercomputadora llamada Pensamiento Profundo proporcionó la respuesta al misterio de «la vida, el universo y todo». Desconcertantemente, como sabemos, la respuesta es «42». ¿Pero podría la IA hacer lo mismo? ¿Podría la IA explicar toda nuestra existencia? La única diferencia, si basamos nuestra experiencia en ChatGPT o MidJourney, es que mientras que a Pensamiento Profundo le tomó diez millones de años llegar a la respuesta, la IA podría responder en tres segundos.

En la contraportada del libro *The Hitchhiker's Guide to the Galaxy*, están escritas las palabras: «¡Que no cunda el pánico!»
«Pero no es ya hora de tener miedo?»

REFERENCIAS

Brown, S.

2023 Why neural net pioneer Geoffrey Hinton is sounding the alarm on ai [en línea]. MIT Sloan School of Management. <<https://mitsloan.mit.edu/ideas-made-to-matter/why-neural-net-pioneer-geoffrey-hinton-sounding-alarm-ai>>.

Harari, Y. N.

2023 «Yuval Harari argues that AI has hacked the operating system of human civilisation». *The Economist*.

Haven, W. D.

2023 «Geoffrey Hinton tells us why he's now scared of the tech he helped build [en línea]». MIT Technology Review. <<https://www.technologyreview.com/2023/05/02/1072528/geoffrey-hinton-google-why-scared-ai/>>.

Leach, N.

2022 *Architecture in the Age of Artificial Intelligence: An Introduction to AI for Architects*. Londres: Bloomsbury.