

“Hablamos del posible fin de la historia humana”

Fuente: lavanguardia.com

Link: <https://www.lavanguardia.com/tecnologia/20230503/8937187/ia-hackeado-sistema-operativo-civilizacion-humana.html>

Fecha artículo original: 2023-05-03, actualizado en 2023-12-16

Publicado en web-ur.com: 2024-09-09

Fecha última actualización: 2024-09-09

Autor/Editor: Yuval Noah Harari / LA VANGUARDIA

Traducción: Juan Gabriel López Guix

LA VANGUARDIA



Yuval Noah Harari

Historiador y filósofo

ARTÍCULO SOBRE LA IA (THE ECONOMIST)

03/05/2023 06:00 Actualizado a 16/12/2023 06:00

----- Lo más leído del 2023

'La Vanguardia' recupera las noticias que más han interesado a los lectores en el año 2023. Este artículo se publicó el día 3 de mayo.

Los miedos a la inteligencia artificial han obsesionado a la humanidad desde el comienzo de la época informática. Hasta ahora, esos miedos se centraban en las máquinas que utilizaban medios físicos para matar, esclavizar o sustituir a las personas. Sin embargo, en los últimos dos años han aparecido nuevas herramientas de inteligencia artificial que amenazan la supervivencia de la civilización humana desde un flanco inesperado. La inteligencia artificial ha adquirido notables capacidades para manipular y generar lenguaje, ya sea con palabras, sonidos o imágenes. Y, al hacerlo, ha hackeado el sistema operativo de nuestra civilización.

El lenguaje es la materia de la que está hecha casi toda la cultura humana. Los derechos humanos, por ejemplo, no están inscritos en nuestro ADN. Son, más bien, artefactos culturales que creamos construyendo relatos y escribiendo leyes. Los dioses no son realidades físicas. Son, más bien, artefactos culturales que creamos inventando mitos y componiendo escrituras sagradas.



Yuval Noah Harari (Archivo)

También el dinero es un artefacto cultural. Los billetes no son más que pedazos de papel coloreado; en la actualidad, más del 90% del dinero ni siquiera son billetes, sólo información digital almacenada en ordenadores. Lo que da valor al dinero son los relatos que nos cuentan sobre él banqueros, ministros de Economía y gurús de las criptomonedas. Sam Bankman-Fried, Elizabeth Holmes y Bernie Madoff no eran especialmente buenos creando valor real, pero todos ellos fueron narradores de una habilidad extraordinaria.

¿Qué pasará cuando una inteligencia no humana sea mejor que el ser humano medio para contar historias, componer melodías, dibujar imágenes y redactar leyes y escrituras? Cuando pensamos en ChatGPT y otras nuevas herramientas similares, pensamos en escolares que recurren a la inteligencia artificial para componer sus redacciones. ¿Qué le ocurrirá al sistema escolar cuando los jóvenes hagan eso? En realidad, esa clase de pregunta pasa por alto la visión de conjunto. Olvidémonos de las redacciones escolares. Pensemos en las próximas elecciones presidenciales estadounidenses de 2024 e intentemos imaginar la repercusión en ellas de las herramientas de inteligencia artificial, que son susceptibles de utilizarse para producir de modo masivo contenido político, noticias falsas y escrituras para nuevos cultos.

En los últimos años, el culto QAnon se ha aglutinado en la red en torno a mensajes anónimos conocidos como “píldoras Q”. Sus seguidores coleccionan, veneran e interpretan esas “píldoras” como si fueran un texto sagrado. Por lo que sabemos, todas las “píldoras Q” anteriores han sido compuestas por humanos y los bots sólo se han limitado a difundirlas, pero en el futuro podríamos ver los primeros cultos de la historia cuyos textos venerados habrán sido escritos por una inteligencia no humana. Las religiones han sostenido a lo largo de la historia que sus libros sagrados procedían de una fuente no humana. Eso podría ser pronto una realidad.

En un nivel más prosaico, pronto podríamos encontrarnos debatiendo largamente *online* sobre el aborto, el cambio climático o la invasión rusa de Ucrania con entidades que pensamos que son seres humanos, pero que en realidad son inteligencias artificiales. El problema reside en que resulta del todo inútil que dediquemos tiempo a intentar cambiar las opiniones sostenidas de un bot de inteligencia artificial, y que, en cambio, la inteligencia artificial puede refinar con tanta precisión los mensajes que tendrá muchas posibilidades de influir en nosotros.

Gracias a su dominio del lenguaje, las inteligencias artificiales podrían incluso entablar relaciones muy cercanas con las personas y utilizar el poder proporcionado por esa cercanía para modificar nuestras opiniones y visiones del mundo. Aunque no hay ningún indicio de que las inteligencias artificiales tengan conciencia ni sentimientos propios, les bastará con lograr que éstos se sientan emocionalmente vinculados a ellas para fomentar una falsa intimidad con los humanos. En junio de 2022, Blake Lemoine, un ingeniero de Google, anunció que el chatbot de inteligencia artificial LaMDA, en el que estaba trabajando, había adquirido conciencia. La polémica afirmación le costó el empleo. Lo más interesante de ese episodio no fue la afirmación de Lemoine, probablemente falsa. Fue, más bien, la disposición que mostró a arriesgar su lucrativo puesto de trabajo por defender el chatbot de inteligencia artificial. Si la inteligencia artificial puede influir en las personas y hacer que pongan en riesgo su trabajo, ¿a qué más podrá inducir las?

En una batalla política por las mentes y los corazones, la intimidad es el arma más eficaz, y la inteligencia artificial acaba de conseguir la capacidad de establecer de modo masivo relaciones muy cercanas con millones de personas. Todos sabemos que en la última década las redes sociales se han convertido en un campo de batalla por el control de la atención humana. Con la nueva generación de inteligencia artificial, el frente de batalla se está trasladando de la atención a la intimidad. ¿Qué ocurrirá con la sociedad y la psicología humanas cuando la inteligencia artificial luche con la inteligencia artificial en una batalla por fingir relaciones muy cercanas con nosotros, relaciones que luego pueden utilizarse para convencernos de que votemos a determinados políticos o compremos determinados productos?



sociedad y basado en principios éticos. Getty Images/iStockphoto

La Unión Europea apuesta por un modelo de IA centrado en la

Incluso sin llegar a crear una “falsa intimidad”, las nuevas herramientas de inteligencia artificial tendrán una inmensa influencia en nuestras opiniones y concepciones del mundo. Las personas podrían llegar a utilizar un único asesor de inteligencia artificial como oráculo universal omnisciente. No es de extrañar que Google esté aterrizado. ¿Para qué molestarse en buscar, si puedo preguntar al oráculo? Los sectores del periodismo y la publicidad también deberían estar aterrizados. ¿Para qué leer un periódico si puedo pedirle al oráculo que me diga cuáles son las últimas noticias? ¿Y para qué sirven los anuncios si puedo pedirle al oráculo que me diga qué comprar?

De todos modos, ni siquiera esos escenarios logran ofrecer la visión de conjunto. De lo que en realidad hablamos es del posible fin de la historia humana. No el fin de la historia, sólo el fin de su parte dominada por los humanos. La historia es la interacción entre biología y cultura; entre necesidades y deseos biológicos de cosas como comida y sexo, y creaciones culturales como las religiones y las leyes. La historia es el proceso mediante el cual las leyes y las religiones moldean la comida y el sexo.

¿Qué ocurrirá con el curso de la historia cuando la inteligencia artificial se apodere de la cultura y empiece a producir relatos, melodías, leyes y religiones? Las herramientas anteriores, como la imprenta y la radio, ayudaron a difundir las ideas culturales de los humanos, pero nunca crearon ideas culturales propias. La inteligencia artificial es en todo punto diferente. La inteligencia artificial puede crear ideas completamente nuevas, una cultura completamente nueva.

En un principio, es probable que imite los prototipos humanos con los que se ha entrenado en su infancia. Sin embargo, a medida que pasen los años, la cultura de la inteligencia artificial se atreverá a adentrarse en terrenos nunca hollados por el ser humano. Durante milenios, los humanos hemos vivido en los sueños de otros humanos. En las próximas décadas, podríamos encontrarnos viviendo en los sueños de una xenointeligencia.

El miedo a la inteligencia artificial sólo obsesiona a la humanidad desde hace unas pocas décadas. Sin embargo, los seres humanos llevamos milenios obsesionados por un miedo mucho más profundo. Siempre hemos apreciado el poder de los relatos y las imágenes para manipular la mente y crear ilusiones. Por ello, desde los tiempos antiguos, los humanos hemos temido quedar atrapados en un mundo de ilusiones.

En el siglo XVII, René Descartes temió que tal vez un demonio malicioso lo tuviera atrapado dentro de un mundo de ilusiones y que creara todo cuanto él veía y oía. En la antigua Grecia, Platón contó la famosa alegoría de la caverna en la que un grupo de personas pasa toda la vida encadenado en el interior de una cueva frente a un muro vacío. Una pantalla. En esa pantalla ven proyectadas diversas sombras. Los prisioneros confunden las ilusiones que ven en ella con la realidad.

En la antigua India, los sabios del budismo y el hinduismo afirmaron que los seres humanos vivíamos atrapados en maya, el mundo de las ilusiones. A menudo, lo que solemos tomar por realidad no son más que ficciones de nuestra propia mente. Las personas pueden librar guerras, matar a otras y estar dispuestas a que las maten llevadas por la creencia en esta o aquella ilusión.

La revolución de la inteligencia artificial nos enfrenta al demonio de Descartes, a la caverna de Platón, a maya. Si no tenemos cuidado, podríamos quedar atrapados tras un velo de ilusiones que seríamos incapaces de rasgar... ni de darnos cuenta siquiera de su existencia.

Por supuesto, el nuevo poder de la inteligencia artificial también puede utilizarse con propósitos positivos. No insistiré en ese aspecto, porque quienes desarrollan la inteligencia artificial ya hablan bastante de él. La tarea de los historiadores y los filósofos como yo es señalar los peligros. Con todo, no cabe duda de que la inteligencia artificial puede ayudarnos de innumerables maneras, desde encontrar nuevas curas para el cáncer hasta descubrir soluciones a la crisis ecológica. La cuestión a la que nos enfrentamos es cómo asegurarnos de que las nuevas herramientas de la inteligencia artificial se utilizan para el bien y no para el mal. Para ello, primero tenemos que comprender las verdaderas capacidades de esas herramientas.

Desde 1945, hemos sabido que la tecnología nuclear podía generar energía barata en nuestro beneficio; pero también que podía destruir físicamente nuestra civilización. Por ello, hemos reformado por completo el orden internacional con objeto de proteger a la humanidad y asegurarnos de que la tecnología nuclear se utiliza básicamente para el bien. Ahora tenemos que enfrentarnos a una nueva arma de destrucción masiva capaz de aniquilar nuestro mundo mental y social.

Todavía estamos a tiempo de regular las nuevas herramientas de la inteligencia artificial, pero debemos actuar con rapidez. Las armas nucleares no pueden inventar armas nucleares más potentes, pero la inteligencia artificial sí que puede crear inteligencia artificial exponencialmente más potente. El primer paso crucial es exigir rigurosos controles de seguridad antes de que las potentes herramientas de la inteligencia artificial salgan al dominio público. Del mismo modo que una compañía farmacéutica no puede lanzar nuevos medicamentos sin probar antes sus efectos secundarios a corto y largo plazo, las

compañías tecnológicas no deberían lanzar nuevas herramientas de inteligencia artificial sin asegurarse antes de que son inocuas. Necesitamos, en el caso de las nuevas tecnologías, un equivalente de la Administración de Alimentos y Medicamentos estadounidense, y lo necesitábamos para ayer.

Frenar el despliegue público de la inteligencia artificial, ¿no hará que las democracias pierdan terreno frente a los regímenes autoritarios con menos escrúpulos? Todo lo contrario. Son los despliegues no regulados de la inteligencia artificial los que crearán un caos social que beneficiará a los autócratas y destruirán las democracias. La democracia es una conversación, y las conversaciones se basan en el lenguaje. Si la inteligencia artificial hackea el lenguaje, destruirá nuestra capacidad de mantener conversaciones significativas y con ello destruirá la democracia.

Acabamos de darnos de bruces con una xenointeligencia, aquí en la Tierra. No sabemos mucho sobre ella, salvo que podría destruir nuestra civilización. Debemos poner fin al despliegue irresponsable de herramientas de inteligencia artificial en el ámbito público y regular la inteligencia artificial antes de que ella nos regule a nosotros. Y la primera regulación que sugiero que sea obligatorio que una inteligencia artificial revele que es una inteligencia artificial. Si mantengo una conversación con alguien y no puedo saber si es un humano o una inteligencia artificial, se acabó la democracia.

Este texto ha sido generado por un humano.

¿O quizás no?

© 2023 The Economist Newspaper Limited. All rights reserved

Traducción: Juan Gabriel López Guix