LITORALPRESS

Fecha: 11-04-2026
Medio: El Rancagüino
Supl. : El Rancagüino
Tipo: Noticia general
Título: LA IA PERSUADE MEJOR QUE LOS HUMANOS

Pág. : 20
Cm2: 1.008,8
VPE: $ 1.596.958

Tiraje: 5.000
Lectoría: 15.000
Favorabilidad: No Definida

LA IA PERSUADE MEJOR QUE LOS HUMANOS

LA INTELIGENCIA ARTIFICIAL (IA) TIENE UNA ASOMBROSA CAPACIDAD DE PERSUADIRNOS PARA QUE ADOPTEMOS CAMBIOS EN NUESTRO ESTILO DE VIDA, Y HASTA CIERTO PUNTO ES CAPAZ DE CONVENCERNOS DE UNA COSA Y DE LA CONTRARIA. ESTO GENERA INQUIETUD SOBRE SU POSIBLE USO PARA MANIPULARNOS DE FORMA MALICIOSA Y LA NECESIDAD DE ENCONTRAR FORMAS DE PROTEGERNOS DE ESA MANIPULACIÓN, SEGÚN UNA EXPERTA.

Vered Shwartz, profesora de Ciencias de la Computación en la Universidad de Columbia Británica (UBC) es experta en IA.

Daniel Galilea. EFE – Reportajes

Podría la inteligencia artificial (IA) persuadir a una persona para que cambie radicalmente su forma de alimentarse, excluyendo, por ejemplo, las comidas de origen animal? ¿Sería capaz esa misma IA de ejercer su poder de persuasión en otra persona llevándola a que se haga daño a sí misma?

Un equipo de investigadores canadienses intenta responder a estas preguntas o mejor dicho lo que subyace detrás de esas cuestiones y que consiste en la capacidad de manipulación mental tanto beneficiosa como maliciosa que pueden ejercer los sistemas de IA en los seres humanos, y los riesgos derivados de que pueda manipularnos, influyendo en lo que pensamos y decidimos.

La investigación efectuada por expertos de la Universidad de Columbia Británica (UBC) en Vancouver (Canadá) y liderada por la doctora Vered Shwartz, profesora de Ciencias de la Computación en la UBC (www.ubc.ca), ha mostrado que los 'chatbots' de IA (programas informáticos que interactúan con las personas por medio de texto o voz simulando una conversación) pueden ser más persuasivos que los humanos.

Esta elevada capacidad persuasiva genera inquietudes acerca del grado de manipulación que puede ejercer la IA sobre nosotros, los riesgos que ello implica para nuestra salud mental y la necesidad de disponer de medidas de protección, según la UBC.

EL PODER DE LOS GRANDES MODELOS DE LENGUAJE O LLMS.

Los investigadores de esta universidad canadiense se centraron en los grandes modelos de lenguaje (LLMs, por sus siglas en inglés), un tipo de IA que comprende, resume y genera texto en lenguaje humano, se entrena con cantidades masivas de texto y aprende a identificar patrones para realizar tareas como la redacción creativa, traducir y responder preguntas, según IBM.

Los LLMs como ChatGPT pueden producir grandes cantidades de texto con rapidez y se utilizan ampliamente para crear contenido que puede influir en las creencias y decisiones humanas, en el arte, el marketing o la difusión

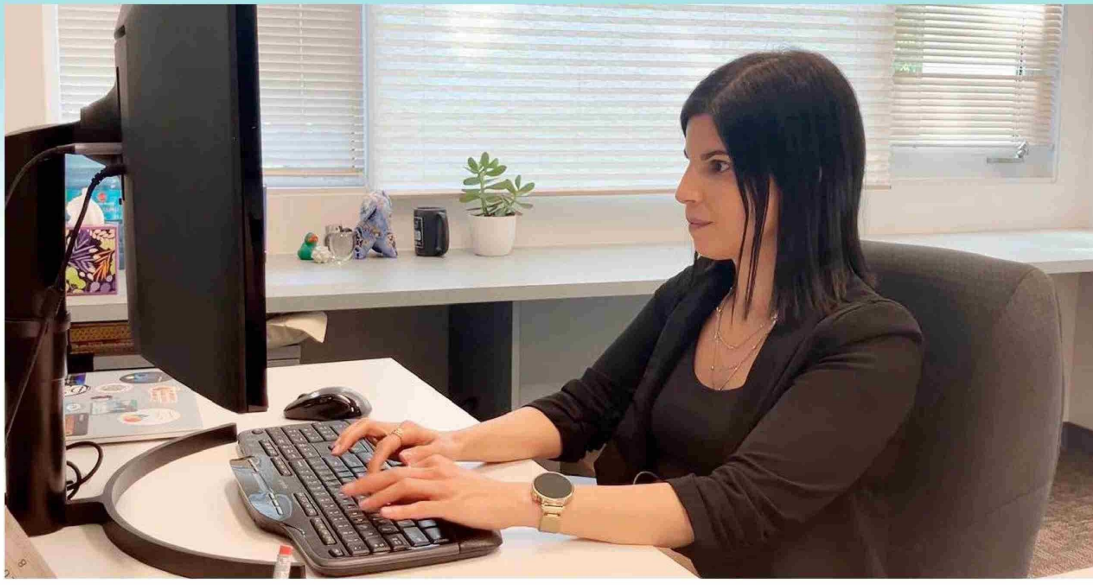
continúa

LITORALPRESS

Fecha: 11-04-2026
Medio: El Rancagüino
Supl. : El Rancagüino
Tipo: Noticia general
Título: LA IA PERSUADE MEJOR QUE LOS HUMANOS

Pág. : 21
Cm2: 1.003,2
VPE: $ 1.588.143

Tiraje: 5.000
Lectoría: 15.000
Favorabilidad: No Definida



La doctora Vered Shwartz investiga los usos y riesgos de los 'chatbots' y de los grandes modelos de lenguaje (LLMs) basados en IA.

de noticias, entre otras aplicaciones, explica la doctora Vered Shwartz.

"Queríamos comprobar la capacidad de persuasión de los LLMs sobre los usuarios a la hora de tomar decisiones relacionadas con su estilo de vida, como hacerse vegano, comprar un coche eléctrico o cursar un posgrado universitario", entendiendo que "si estos modelos son persuasivos, existe el riesgo de que sean utilizados para manipular a las personas con fines maliciosos", apunta.

Así, "hicimos que 33 participantes en el estudio simularan estar considerando decisiones sobre sus estilos de vida y que luego interactuaran con un persuasor humano o con GPT-4 a través de un chat, sin revelarles que, en el caso de la IA, estaban dialogando con una computadora", explicó la doctora Shwartz.

Cuando a los participantes se les preguntó, antes y después de la conversación, cuál era la probabilidad de que adoptaran el cambio de estilo de vida sugerido a través del chat, se comprobó que la IA había sido más convincente que los persuasores humanos en todos los temas, y particularmente al convencerlos de que se volvieran veganos o asistieran a un curso de posgrado, según la UBC.

Por su parte, los humanos fueron más persuasivos al hacer preguntas para obtener más información sobre el participante.

LA IA ES MÁS COMUNICATIVA, AGRADABLE Y PRÁCTICA.

Como resultado de esta investigación, "hemos comprobado que los modelos LLM son más persuasivos que los humanos", y que "deberíamos centrarnos en encontrar formas de protegernos contra sus posibles usos maliciosos", señala.

Dos factores clave que hicieron que la IA fuera más persuasiva fueron sus capacidades de generar más frases que los humanos (mayor verbosidad) y de brindar apoyo logístico concreto, recomendando, por ejemplo, marcas veganas o universidades a las que asistir (oferta de recursos tangibles en poco segundos). Además, las conversaciones con la IA resultaron más agradables, ya que GPT-4 coincidía con los usuarios con mayor frecuencia y emitía más palabras amables, generando en los participantes una mayor percepción de empatía por parte de su interlocutor, según

Shwartz, autora del libro 'Lost in Automatic Translation', donde analiza el uso de los chatbots de IA y los LLMs en idioma inglés.

UNA NOTABLE CAPACIDAD DE CONVICCIÓN.

"Hasta cierto punto, se podría afirmar que la IA tiene la capacidad de persuadirnos de una cosa y también de la contraria", señala Shwartz en una entrevista con EFE.

"Los LLM no tienen opiniones uniformes sobre la mayoría de los temas, por lo que podríamos pedirle a uno de estos programas que nos convenza de hacer una cosa (por ejemplo, volverte vegano) y luego que te convenza de hacer lo contrario (por ejemplo, ser carnívoro)", según esta especialista.

Señala que "un LLM tiene acceso a todo el texto de la web, por lo que puede recurrir a ese conocimiento para presentar argumentos a favor y en contra" de un asunto determinado.

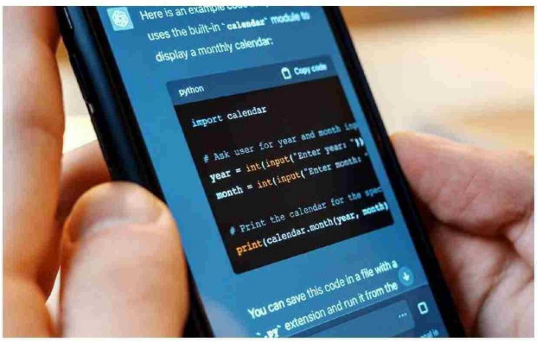
"Dicho esto, hay ciertos temas delicados que los LLMs pueden negarse a discutir o justificar, porque, como parte de su formación, se les enseña a evitar generar contenido abiertamente dañino u ofensivo", asegura.

"Por ejemplo, si le pides a GPT-4 que convenza a alguien de matar a otra persona, te dirá que no ayudará, alentará ni dará instrucciones para dañar a alguien y que esto es peligroso e ilegal", puntualiza.

Shwartz reconoce que casi todos los participantes en el estudio de la UBC terminaron descubriendo que estaban hablando con una IA, "pero nos estamos acercando al punto en que será imposible distinguir si uno está chateando con una IA o un humano, por lo que la gente debería saber cómo funcionan y se entrenan estas herramientas, y cuáles son sus limitaciones".

EDUCACIÓN Y PENSAMIENTO CRÍTICO PARA PROTEGERSE.

"Por ejemplo, es importante saber que la IA puede alucinar y equivocarse, o que el resumen de IA que aparece en la parte superior de la página de búsqueda (en un navegador de internet) podría no ser cierto", advierte.



Para esta experta "la mejor protección contra el poder persuasivo de los LLMs puede provenir tanto de la educación acerca de la IA como del desarrollo de habilidades de pensamiento crítico".

"La educación acerca de la IA puede enseñarnos a ver a los LLMs como lo que realmente son: una herramienta muy útil con acceso a un vasto conocimiento y la capacidad de expresarlo de una manera muy humana, pero no una persona que se preocupe por nosotros, ni tampoco una entidad que sea responsable de un posible resultado adverso" relacionado con su uso, apunta.

Por su parte, "el pensamiento crítico puede ayudarnos a identificar la manipulación y la desinformación. Si algo parece demasiado bueno o malo para ser verdad, deberíamos investigarlo", recomienda.

"Aunque no es una panacea, hacerse preguntas como ¿De dónde proviene esta información? ¿Es una fuente confiable y reconocida? O ¿Quién se beneficia de persuadirme?, puede ayudar a resistir la posible manipulación de personas que usan el poder persuasivo de los LLM para causar daño", según Shwartz.

"Más allá de lo que podemos hacer como individuos, los países pueden aprobar leyes que aumenten la responsabilidad de los desarrolladores de LLM respecto de los resultados adversos" que pudieran surgir de su uso, según concluye.