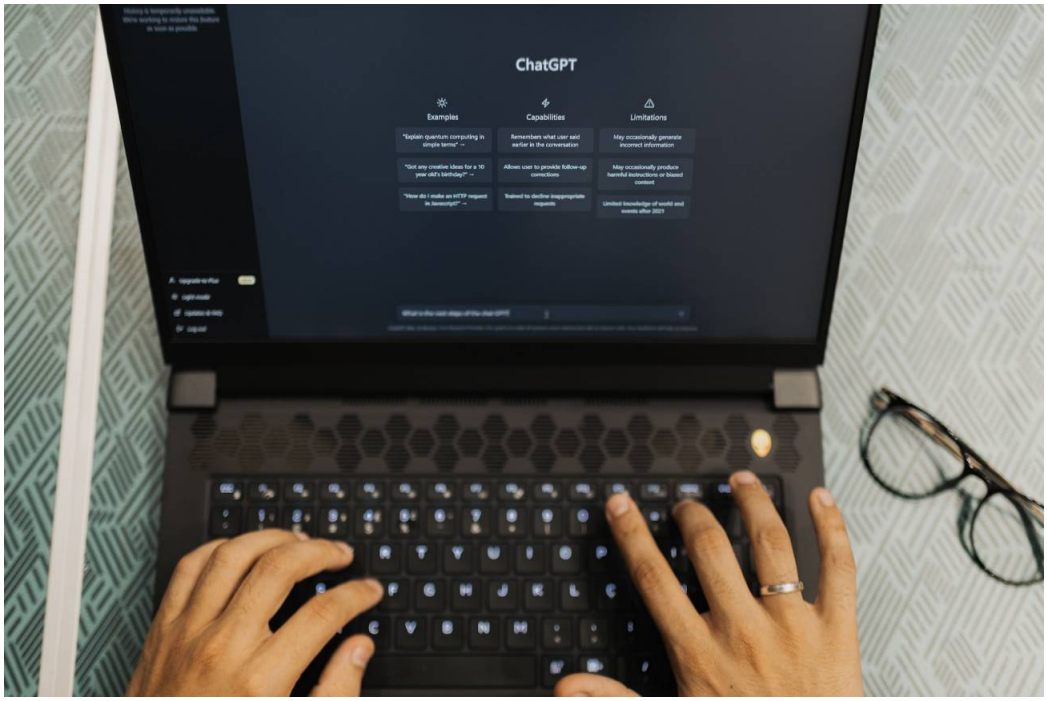
Las herramientas de IA para generar texto homogenizan el lenguaje

Un análisis de más de 880 000 textos sugiere que la escritura a partir de grandes modelos de lenguaje como ChatGPT o Gemini borra las peculiaridades individuales a la hora de redactar. Según los investigadores, esto podría dificultar las evaluaciones basadas en el lenguaje, como en el ámbito de la salud mental.

SINC

24/8/2026 17:00 CEST



Según el estudio, la escritura a partir de grandes modelos de lenguaje borra las peculiaridades individuales a la hora de redactar. /Pexels

Los grandes modelos de lenguaje (LLM, por sus siglas en inglés) basados en inteligencia artificial están haciendo que los estilos de redacción sean más similares. Esta es la conclusión de un análisis de más de 880 000 textos publicado en la revista *Nature Human Behaviour*.

Los resultados de la investigación, liderada por la Universidad del Sur de California (EE UU), indican que el uso de estos modelos podría dificultar la identificación de pistas importantes sobre **aspectos de la identidad**, la personalidad y la salud mental de una persona a partir de su lenguaje.

Rasgos de estilo individual

Las personas expresan ideas similares de formas distintas, y estas diferencias proporcionan información sobre ellas. En muchas disciplinas, como en la psicología o la psiquiatría, se utilizan los **patrones lingüísticos** para estudiar a los individuos y las sociedades. Pero esto podría estar en peligro por el uso generalizado de los grandes modelos de lenguaje de IA.

Se observó que la reescritura realizada por los LLM reducía la variación en la complejidad de la redacción entre un 21 % y un 50 %

En el trabajo, el equipo dirigido por el investigador Zhivar Sourati analizó más de **880 000 textos**, entre los que se incluían historias de Reddit, artículos de prensa, trabajos académicos, ensayos, publicaciones en redes sociales y discursos políticos.

También pidieron a GPT-3.5, Llama 3 70B y Gemini Pro que reescribieran miles de textos redactados por humanos. Se observó que la reescritura realizada por los LLM reducía la variación en la **complejidad de la redacción** entre un 21 % y un 50 %.

En el 87 % de los casos, los textos originales y los reescritos obtuvieron puntuaciones de similitud semántica superiores a 0,95, lo que indica que

los LLM **conservaron en gran medida el significado original**.

Los modelos de lenguaje aún confunden las creencias con los hechos, incluso los más avanzados

Antonio Villarreal

Diferencia en lo que se mantiene

Tras la reescritura, los modelos informáticos utilizados para analizar los textos generados por los LLM también mostraron una precisión media seis puntos porcentuales menor a la hora de **identificar las características** personales de los autores a partir de la redacción.

> Se mantuvieron asociaciones como las existentes entre las palabras relacionadas con emociones negativas y el neuroticismo

Entre estos rasgos diferenciales, los modelos de lenguaje atenuaron algunos patrones lingüísticos como el uso de pronombres y la **extroversión**, las palabras relacionadas con los amigos y la lealtad, y las palabras centradas en el futuro y la edad.

Sin embargo, se mantuvieron otras asociaciones, como las existentes entre las palabras relacionadas con **emociones negativas** y el neuroticismo, las palabras relacionadas con la religión y la pureza, y las palabras sociales y el género.

Según los investigadores, los resultados sugieren que la redacción asistida por los modelos basados en IA podría reducir la fiabilidad de las

evaluaciones basadas en el lenguaje en ámbitos como la psicología, la **atención de la salud mental**, la selección de personal y los servicios personalizados.

Referencia:

Zhivar Sourati *et al*. The shrinking landscape of linguistic diversity in the age of large language models. *Nature Human Behaviour* (2026).

TAGS | MODELOS DE LENGUAJE | IA | INTELIGENCIA ARTIFICIAL |