

Un estudio detectó que muchos usuarios no verifican las respuestas de la inteligencia artificial

06/04/2026

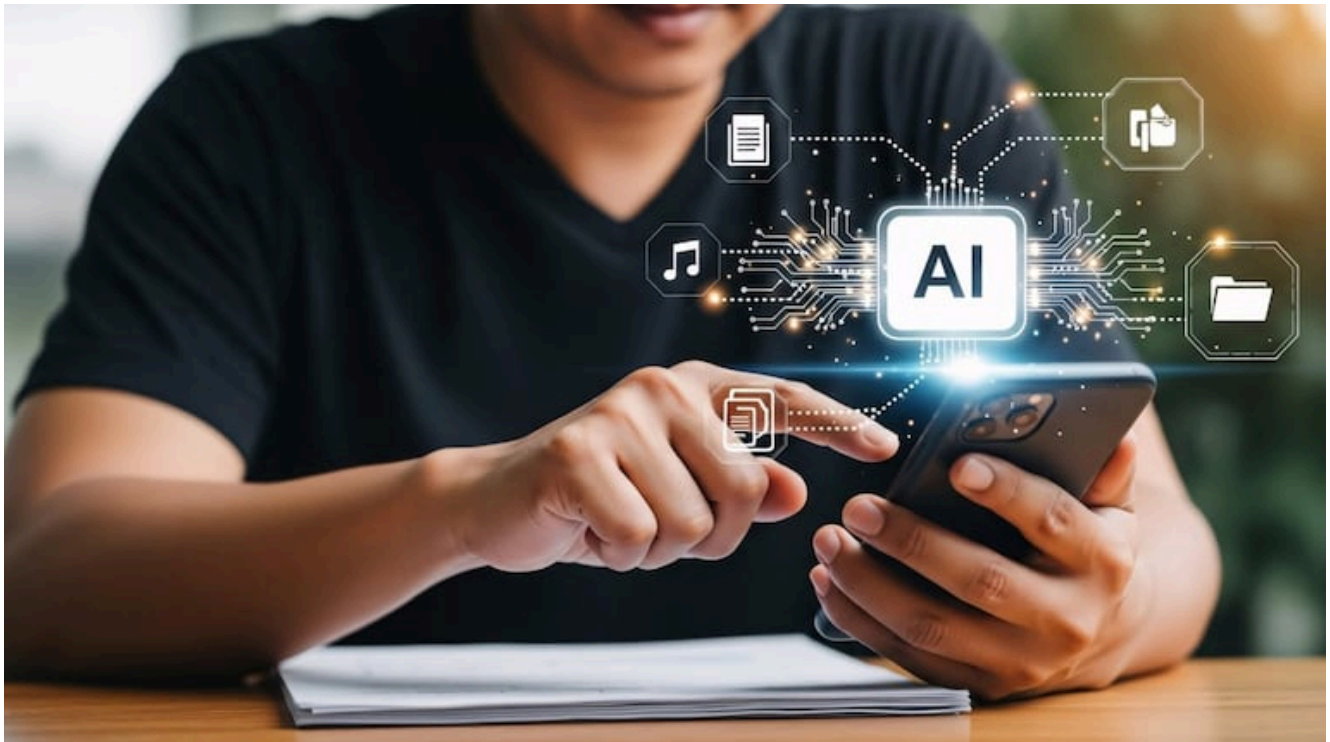


La confianza en la inteligencia artificial (IA) puede tener un costo poco visible: dejar de pensar por cuenta propia. Así lo advirtió una investigación de la Universidad de Pensilvania, que identificó un fenómeno al que llamó “rendición cognitiva” para describir la tendencia de algunos usuarios a aceptar respuestas de IA sin verificar si son correctas.

El trabajo demostró que el problema no aparece solo cuando un chatbot se equivoca, sino en la forma en que muchas personas se relacionan con esa herramienta.

Frente a respuestas redactadas con seguridad, fluidez y tono convincente, una parte importante de los participantes optó por **confiar en el sistema en lugar de revisar el**

razonamiento detrás de la respuesta.



Una investigación de la Universidad de Pensilvania definió el fenómeno de aceptar lo que dicen las herramientas de inteligencia artificial como “rendición cognitiva”. (Foto: Adobe Stock).

De acuerdo a los resultados del estudio, los usuarios **aceptaron razonamientos erróneos generados por IA** en el 73,2% de los casos analizados. En contraste, solo corrigieron esas respuestas en el 19,7% de las veces.

La investigación se apoyó en más de 9500 pruebas individuales realizadas con 1372 participantes y buscó observar hasta qué punto la presencia de **un chatbot modifica la manera en que las personas resuelven problemas** simples de razonamiento.

Qué es la “rendición cognitiva” y por qué preocupa

Los investigadores **diferencian este fenómeno de otras formas conocidas de apoyo tecnológico**. Usar una calculadora o seguir un GPS implica delegar una tarea específica, pero con cierta supervisión humana. En cambio, en la “rendición cognitiva” el

usuario no solo externaliza una parte del trabajo, sino que también reduce su intervención crítica y da por válida la respuesta sin someterla a revisión.

Para estudiar ese comportamiento, el equipo utilizó **pruebas de reflexión cognitiva** y ofreció a los participantes acceso opcional a un chatbot modificado. Ese sistema respondía de forma correcta en algunos casos y de manera incorrecta en otros. Aun así, la aceptación fue alta en ambos escenarios.

Cuando la inteligencia artificial acertaba, los usuarios tomaban su respuesta como válida el 93% de las veces. Pero cuando el sistema fallaba, también la seguían en cerca del 80% de los casos. El dato refuerza una idea central del estudio: **la autoridad percibida de la herramienta puede pesar más que el análisis propio.**

Además, quienes usaron IA se mostraron más seguros de sus respuestas que el grupo que resolvió los ejercicios sin asistencia, incluso cuando el chatbot había entregado respuestas equivocadas en la mitad de los casos.

El tiempo y la confianza previa influyen en la relación con la IA

La investigación también analizó qué factores aumentan o reducen la **capacidad de cuestionar a la inteligencia artificial**. Uno de los más claros fue el tiempo disponible para decidir.

Cuando los participantes tuvieron que **responder bajo presión**, con un límite de 30 segundos, **disminuyó la probabilidad de que detectaran errores del sistema**. En cambio, cuando había **incentivos económicos** y devolución inmediata sobre si la respuesta era correcta, mejoraba la **disposición a frenar, revisar y corregir** lo que proponía el chatbot.

Otro punto relevante fue la **predisposición previa frente a la**

tecnología. Las personas que ya tendían a ver a la IA como una fuente confiable o autoritativa fueron más propensas a aceptar respuestas defectuosas. Por el contrario, quienes mostraban mayores capacidades de **razonamiento** abstracto y analítico **confiaron menos y corrigieron más.**

Una advertencia para el uso cotidiano de chatbots

El estudio no plantea que apoyarse en inteligencia artificial sea siempre una mala decisión. De hecho, sus autores señalan que **delegar parte del razonamiento puede resultar útil** cuando el sistema ofrece respuestas de alta calidad o trabaja sobre grandes volúmenes de datos.

El punto crítico aparece cuando la confianza se convierte en sustituto del pensamiento crítico. En ese escenario, la calidad de la decisión deja de depender del usuario y pasa a estar atada a la precisión del sistema. **Si la IA acierta, el resultado mejora. Si falla, el error también se arrastra.**

La expansión de estas **herramientas** en ámbitos como el estudio, el trabajo y la búsqueda cotidiana de información hace que esa advertencia cobre más peso. El **desafío**, además de desarrollar modelos más capaces, también pasa por evitar que su aparente seguridad termine desplazando una práctica básica: **dudar, revisar y pensar** antes de aceptar una respuesta como verdadera.

Fuente: TN