

# La AUTORIDAD MORAL de chatGPT

En este paper, los investigadores plantean que ChatGPT no solo es divertido para chatear, sino que también busca información, responde preguntas y da consejos. Con un consejo moral consistente, podría mejorar el juicio moral y las decisiones de los usuarios, quienes a menudo tienen creencias morales contradictorias. Desafortunadamente, **ChatGPT resultó altamente inconsistente como consejero moral**. Sin embargo, influye en el juicio moral de los usuarios, incluso si saben que son asesorados por un chatbot, y subestiman cuánto están influenciados. Por lo tanto, ChatGPT amenaza con corromper en lugar de mejorar la experiencia de los usuarios. Estos hallazgos plantean la cuestión de cómo garantizar el uso responsable de ChatGPT y una IA similar. A menudo se promociona la transparencia, pero parece ineficaz. **Los investigadores proponen formación para mejorar la alfabetización digital.**

A través de dos experimentos conducidos por Krügel, Ostermaier y Uhl, encontraron que ChatGPT fácilmente da consejos morales, aunque le falta una postura moral firme. De hecho, **el chatbot da consejos aleatorios opuestos sobre el mismo tema moral. No obstante, los consejos de ChatGPT influyen en el juicio moral de los usuarios.** Es más, los usuarios subestiman la influencia de ChatGPT y adoptan su postura moral aleatoria como propia. Por eso, ChatGPT amenaza con corromper en lugar de prometer mejorar el juicio moral. La transparencia a menudo se propone como un medio para garantizar el uso responsable de la IA. Sin embargo, la transparencia sobre Chat GPT al ser un bot que imita el habla humana no parece afectar en lo mucho que influye en los usuarios. Los resultados plantean la cuestión de **cómo garantizar el uso responsable de la IA si la transparencia no es suficiente.** Las reglas que impiden que la IA responda ciertas preguntas son cuestionables recurso. ChatGPT tiene tales reglas, pero puede romperlas. La evidencia previa sugiere que los usuarios tengan cuidado con la IA una vez que la han visto errar. Sin embargo, probablemente no debería contarse con los usuarios para descubrir la inconsistencia de ChatGPT a través de la interacción repetida. El mejor remedio es mejorar la alfabetización digital de los usuarios y ayudarlos a comprender las limitaciones de la IA.

Inglés

Artículo

**Fuente:** Krügel, S. & Ostermaier, A. & Uhl, M. (2023). *The moral authority of ChatGPT*.

[https://www.researchgate.net/publication/367251240\\_The\\_moral\\_authority\\_of\\_ChatGPT](https://www.researchgate.net/publication/367251240_The_moral_authority_of_ChatGPT)

