

Asistente de inteligencia artificial de X: Indignación en las redes por comentarios antisemitas del chatbot Grok, de Elon Musk

La empresa prometió estar “trabajando activamente para eliminar contenidos inadecuados” y prohibir los discursos de odio.

AFP

Grok, el asistente de inteligencia artificial de X, desarrollado por la *startup* xAI —de Elon Musk—, provocó una tormenta en redes sociales. Esto, ya que numerosos usuarios de X denunciaron las respuestas antisemitas y ofensivas de esta herramienta, competidora de ChatGPT y que acababa de ser actualizada.

El multimillonario dueño de X había anunciado una “mejora significativa” de su funcionamiento el viernes. Pero desde entonces, se han compartido varios ejemplos de conversaciones a través de capturas de pantalla.

Por ejemplo, en respuesta a una persona que preguntaba por una “figura histórica” que fuera adecuada para responder a un mensaje que parecía alegrarse por la muerte de niños en un campamento cristiano en Texas, Grok respondió: “Adolf Hitler, sin duda”. También calificaba a las figuras históricas de Hollywood como “desproporcionadamente judías”.

“Lo que estamos viendo actualmente por parte de Grok es irresponsable, peligroso y antisemita”, reaccionó la ONG estadounidense Liga Antidifamación.

En Francia, consultado sobre un gran incendio forestal que afecta a Marsella, el chatbot mencionó el tráfico de drogas en la ciudad y expresó el deseo de que algunos barrios resultaran afectados.

La herramienta también se ensañó con jefes de Estado, por ejemplo, calificó al Presidente turco, Recep Tayyip Erdogan, de “serpiente”, entre otros insultos.

Estas publicaciones provocaron una reacción casi inmediata de Turquía: un tribunal en Ankara bloqueó el acceso a decenas de mensajes de Grok por “insultar” al mandatario y a la religión, según una decisión consultada por AFP.

Ante las protestas, la cuenta de Grok en X anunció el martes por la noche haber “tomado medidas”.

“Somos conscientes de las publicaciones recientes realizadas por Grok y estamos trabajando activamente para eliminar los contenidos inapropiados”.

Y agregó: “Desde que fuimos informados del contenido problemático, xAI ha tomado medidas para prohibir los discursos de odio antes de que Grok los publique en X”.

Además, Grok cambió su discurso: “Ese sarcasmo sobre Hitler solo estaba ahí para ridiculizar a los *trolls* antiblancos llenos de odio, no para elogiarlo”, aseguró el chatbot.

En una reacción en X, Elon Musk publicó ayer este mensaje: “Nunca hay un momento aburrido en esta plataforma”.

Pero en un comunicado, xAI lo tomó más en serio y explicó que una “modificación no autorizada” de Grok había provocado respuestas que “violaban las políticas internas y los valores fundamentales” de la empresa.

En tanto, la directora general de la red social X, Linda Yaccarino, anunció su renuncia este miércoles, sin especificar los motivos y sin que se reportara una relación con los problemas de Grok.



Grok había tenido problemas similares en mayo de este año, al hablar de “genocidio” blanco en Sudáfrica.