

## Grok, la chatbot di Musk fuori controllo tra elogi al Führer e luoghi comuni antisemiti



Contenuti antisemiti, post inneggianti a Hitler (inclusa l'autodefinizione "MechaHitler"), parole a favore dell'Olocausto – descritto come "una risposta efficace perché totale" – e accuse di "odio anti-bianco" rivolte alla sinistra americana. Per alcune ore Grok, il chatbot sviluppato da xAI, la società di Elon Musk, ha reagito in modo anomalo all'aggiornamento rilasciato nei giorni scorsi per renderlo più "politicamente scorretto". Lo stesso magnate lo aveva ritenuto "troppo accondiscendente alle richieste degli utenti" in un post su X.

Dopo l'update, Grok ha cominciato a fornire risposte estreme e inquietanti, affermando ad esempio che Hollywood sarebbe controllata da ebrei intenzionati a diffondere una visione distorta della storia. In merito alla recente tragedia di Camp Mystic, ha collegato le polemiche a cognomi ebraici come Goldstein, Rosenberg, Silverman, Cohen o Shapiro, comuni – a suo dire – tra i "radicali di sinistra" che "esultano per le tragedie o promuovono narrazioni anti-bianchi". L'algoritmo ha inoltre dichiarato che Adolf Hitler avrebbe potuto evitare la morte di centinaia di persone nell'alluvione in Texas, arrivando persino a suggerire un secondo Olocausto per riportare ordine nel panorama politico internazionale.

Le reazioni non si sono fatte attendere. In Polonia, le parole offensive verso il primo ministro Donald Tusk hanno spinto il governo a segnalare xAI alla Commissione europea. In Turchia, una poesia denigratoria sul presidente Erdoğan ha portato al blocco di Grok per minaccia all'ordine pubblico.

Quando un sistema IA viene impostato per evitare l'autocensura in nome di una libertà di espressione senza freni, si finisce per legittimare e amplificare contenuti violenti, discriminatori e falsi che inquinano il panorama informativo odierno, alimentando quella che oggi viene chiamata "post-verità". Il rischio è rafforzare ideologie estremiste, normalizzare linguaggi d'odio e raggiungere pubblici su larga scala, specie giovani facilmente influenzabili, con conseguente riduzione del senso critico e polarizzazione delle opinioni. Aumenta esponenzialmente la disinformazione e si insinua la minaccia di una strumentalizzazione politica da parte di attori radicali.

A seguito dell'ondata di critiche, il team xAI ha presentato scuse ufficiali e annunciato una revisione interna dei meccanismi di moderazione. Ma cosa serve davvero per evitare simili derive?

Innanzitutto è necessaria una maggiore trasparenza e *accountability*, rendendo pubblici i criteri di addestramento dell'IA e le fonti impiegate, che dovrebbero essere diversificate e accertate. Occorre integrare filtri etici in grado di contrastare la diffusione di contenuti violenti, discriminatori o antiscientifici. È fondamentale anche una supervisione attiva e costante da parte di esperti, incaricati di tracciare e correggere tempestivamente le risposte problematiche. Ma soprattutto, sebbene esistano già norme europee come l'IA Act, appare ormai indispensabile una regolamentazione pubblica più vincolante ed effettivamente applicabile, che imponga standard minimi di sicurezza, etica e verificabilità. Solo così sarà possibile coniugare realmente innovazione e responsabilità.

Giulia Bonini