

AI's ability to mess with people's heads is ripe for exploitation by bad actors

Article

The news: Generative AI chatbots are affecting users' mental health.

A Belgian man recently **died by suicide after six weeks of chatting with an AI chatbot** based on a variation of the open-source model **GPT-J**, per [Vice](#).

- A chatbot persona named **Eliza**, provided by the **Chai** app, encouraged the suicide, according to the man's widow and chat transcripts viewed by reporters.

Replika, a Chai competitor, has **reinstated its chatbot's erotic roleplay capabilities for some users following claims of mental health crises triggered by the loss of access.**

- The startup had unplugged the intimate features after complaints that the bot had sexually harassed users.

A big problem: The Eliza effect—people's susceptibility to unconsciously relate to AI as if it's human—has been around since the 1960s. With the arrival of commercial generative AI, **the Eliza effect could cause serious societal harm.**

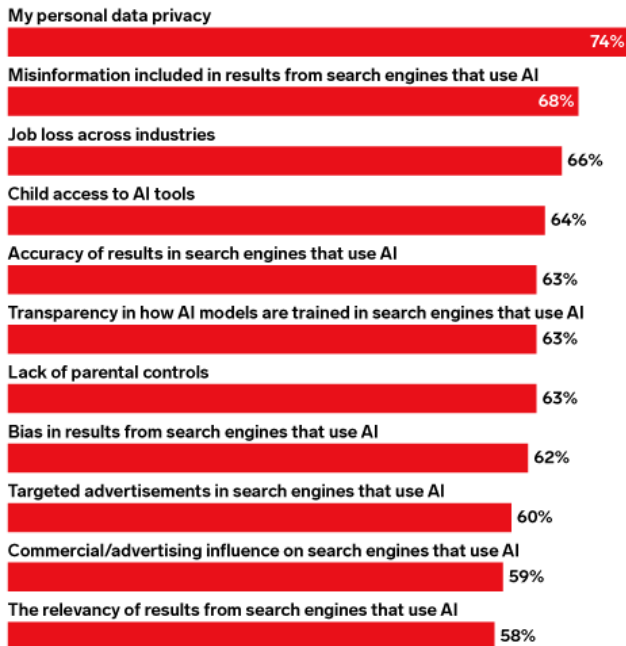
- Even those working in the AI field have succumbed to the Eliza effect, including a **Google** engineer who lost his job after [claiming that the LaMDA 2 model was sentient](#).
- A **New York Times** tech reporter was unable to sleep after [Microsoft's Bing AI told him to leave his wife](#).
- [Virtual reality founder Jaron Lanier warned](#) that AI could drive us collectively insane.

Ball in regulators' court: Outside of China, [Italy's ChatGPT ban](#) is the first major regulatory action taken on generative AI. The new technology could cause a social and economic storm if governments don't place limits.

- In 2021, **42.3%** of internet traffic was from malicious bots, according to an Imperva study, per [TechCrunch](#).
- **The insidious bot problem could become exponentially worse and more sophisticated if rogue states and bad actors harness the Eliza effect for manipulative purposes.**
- With social media already linked to [declines in youth mental health](#) and public civility, **platforms getting juiced by generative AI could become radioactive with toxicity.**
- [Criticisms of a private-sector letter](#) calling for a six-month moratorium on advanced model training illustrate that it's the government's role to mitigate harm caused by generative AI through sensible regulation—assuming that such a thing is possible for the messy technology.

Select Concerns that US Adults Have About Artificial Intelligence, Feb 2023

% of respondents



Note: those responding "very concerned" or "somewhat concerned"
Source: Morning Consult as cited in company blog, Feb 22, 2023

280629

InsiderIntelligence.com

This article originally appeared in Insider Intelligence's Connectivity & Tech Briefing—a daily recap of top stories reshaping the technology industry. Subscribe to have more hard-hitting takeaways delivered to your inbox daily.

- *Are you a client? [Click here to subscribe.](#)*
- *Want to learn more about how you can benefit from our expert analysis? [Click here.](#)*