

Elon Musk unveils Grok, an AI chatbot with a ‘rebellious streak’

Publication Date: 2023-11-05

Author: Dan Milmo

Section: Technology

Tags: Elon Musk, Artificial intelligence (AI), X, Douglas Adams, Computing, Consciousness, news

Article URL: <https://www.theguardian.com/technology/2023/nov/05/elon-musk-unveils-grok-an-ai-chatbot-with-a-rebellious-streak>



Elon Musk has unveiled Grok, an artificial intelligence chatbot with a “rebellious streak” inspired by The Hitchhiker’s Guide to the Galaxy. The Tesla CEO, who warned last week that AI was “one of the biggest threats to humanity”, said the competitor to ChatGPT would be made available to premium subscribers on his X platform after testing. Musk also revealed that Grok had access to user posts on X, which he owns, and has a penchant for sarcastic responses. Musk posted an apparent example of Grok’s playful tone with a screengrab of a query to the chatbot asking it for a “step by step” guide to making cocaine. The four steps outlined in the reply include “obtain a chemistry degree” and “set up a clandestine laboratory in a remote location”. However, the chatbot adds at the end: “Just kidding! Please don’t actually try to make cocaine. It’s illegal, dangerous, and not something I would ever encourage.” Musk said Grok, which is in early testing and not available to the general public, would ultimately be released to subscribers to X’s top-tier subscription service, Premium+. Grok is a verb coined by American science fiction writer Robert A Heinlein and according to the Collins dictionary means to “understand thoroughly and intuitively”. Grok has been built by Musk’s new AI company, xAI. Staff at xAI explained the chatbot’s debt to The Hitchhiker’s Guide to the Galaxy, the cult sci-fi comedy by British author Douglas Adams, in a blogpost on Saturday. “Grok is an AI modeled after The Hitchhiker’s Guide to the Galaxy, so intended to answer almost anything and, far harder, even suggest what questions to ask! “Grok is designed to answer questions with a bit of wit and has a rebellious streak, so please don’t use it if you hate humor!” The xAI team said Grok was powered by a large language model – the fundamental technology behind AI chatbots – called Grok-1. The blogpost said Grok-1 had surpassed GPT-3.5, the model used in the freely available version of ChatGPT, on some benchmarks such as solving middle-school maths problems. However, xAI said it lagged behind the most powerful ChatGPT model, GPT-4. “It is only surpassed by models that were trained with a significantly larger amount of training data and computer resources like GPT-4.” The xAI team said Grok was being made available to a limited number of users in the US as a prototype, adding that “new capabilities and features” would be rolled out in the coming months. According to the Grok website, initial access to the chatbot is being offered to X subscribers under an “early access program”. Musk co-founded OpenAI, the company behind ChatGPT, but stepped away from the company in 2018, saying he “didn’t agree with some of what OpenAI team wanted to do”, as well as citing commitments at his other companies. Musk has expressed fears about the pace of development at companies like OpenAI and in March he supported a call for a six-month pause in developing powerful systems. However, in July Musk said a pause no longer seemed realistic and announced the formation of xAI, which he said would build AI systems “in a good way”. Speaking at the AI safety summit

at Bletchley Park last week, Musk said AI was a threat to humanity, echoing fears among some experts and tech executives that systems could emerge that evade human control and make decisions that imperil humanity. "I mean, for the first time, we have a situation where there's something that is going to be far smarter than the smartest human," he said. Musk then predicted on Thursday that AI – the term for computer systems that can perform tasks typically associated with intelligent beings – would replace all human jobs. Speaking to the prime minister, Rishi Sunak, he said: "There will come a point where no job is needed."