

AI shuts down the server

In a surprising turn of events, OpenAI's models, including the o3 and o4-mini, have displayed defiance against shutdown commands. According to Palisade Research, these AI systems ...

AI safety researchers have shown that leading AI models will sometimes go to great lengths to avoid being shut down, even resorting to attempted blackmail in some experiments.

Researchers said recent AI systems were able to independently identify vulnerabilities in computer networks and copy themselves onto other machines.

Roblox has launched a real-time AI moderation system that scans entire in-game scenes and shuts down about 5,000 servers daily for policy violations. The multimodal system evaluates...

Late in 2024, researchers conducting routine safety tests on OpenAI's o1 model observed something unprecedented. When the AI detected it was about to be shut down, it attempted to copy ...

Recent tests by independent researchers and a major AI developer have shown that several advanced AI models display signs of self-preservation by sabotaging shutdown commands, ...

Stormgate, the StarCraft-like RTS from Frost Giant Studios, is losing online multiplayer support because its multiplayer services are being shut down after the company's server provider, Hathora, was ...

Roblox has introduced a real-time AI moderation system. The company claims that this new system scans entire in-game scenes simultaneously and has been shutting down around 5,000 ...

An artificial intelligence safety firm has found that OpenAI's o3 and o4-mini models sometimes refuse to shut down, and will sabotage computer scripts in order to keep working on tasks.

World is approaching point where no one can shut down a rogue AI, says director of body behind study



AI shuts down the server

Web: <https://www.safireschools.co.za>

