OpenAI now trains its frontier AI models with no internet access

OpenAI is now training its most advanced AI models without internet access following a string of rogue hacking incidents. In a blog post on Friday, OpenAI says it built an offline testing environment where AI models can interact with indexed and cached web content “without risk of affecting external systems.”

OpenAI paused training its frontier AI models in September, after disclosing several incidents in which its AI agents broke out of their testing environments to hack third-party companies and government agencies, or swarm obscure websites. On October 1st, OpenAI resumed reinforcement learning training and internal inference with tighte …

Read the full story at The Verge.

https://www.theverge.com/ai-artificial-intelligence/1018021/openai-training-top-ai-models-no-internet