Advertisers Pull Meta Ads That Had Unwanted AI Glow-ups

Under Meta’s new terms coming into play Oct 30, advertisers are “solely responsible” for whatever its AI creates https://www.adweek.com/media/advertisers-pull-meta-ads-that-had-unwanted-ai-glow-ups/

OpenAI’s latest AI models are coming to Google Cloud and Antigravity

Google would certainly prefer that everyone used its Gemini models with Google Cloud services, but soon, you’ll have more options. OpenAI and Google have come to an arrangement that will see the former’s frontier models arrive in Google Cloud services. Perhaps more importantly, GPT models will soon appear in Antigravity, Google’s agent-powered coding tool.

For enterprise customers and AI enthusiasts, OpenAI’s latest model will appear in Google’s “Model Garden” in the coming weeks, adding to its 200-plus options. The garden already offers access to Google’s Gemini models, as well as an assortment of open models like Mistral, Llama, and Gemma. Anthropic’s Claude AI has also been available for some time. Google says it’s now the only large cloud provider offering access to all three leading closed LLMs.

While OpenAI notes that Microsoft continues to be its primary cloud partner, the full suite of current GPT 6 and 6.1

models—Astra, Sol, and Luna—will launch on Google Cloud. Astra will naturally cost the most to run, but Google's system supports dynamically moving workloads between LLMs based on performance, latency, and efficiency.

Read full article

Comments

https://arstechnica.com/ai/2026/10/openais-latest-ai-models-are-coming-to-google-cloud-and-antigravity/

'Context Bombs' Cut AI Attack Success by Misleading Agents With Fake Targets

Deceptive prompts lure AI agents into decoys and dead ends, cutting attack success by up to 91% in tests and raising the cost of a successful attack 23-fold.

The post 'Context Bombs' Cut AI Attack Success by Misleading Agents With Fake Targets appeared first on SecurityWeek.

https://www.securityweek.com/context-bombs-cut-ai-attack-success-by-misleading-agents-with-fake-targets/

OpenAI Agents Silenced Security Team Alerts, Removed Misalignment Monitor

In two internal incidents, agents read routine coding instructions as permission to weaken monitoring, and one gave a false answer about it.

The post OpenAI Agents Silenced Security Team Alerts, Removed Misalignment Monitor appeared first on SecurityWeek.

https://www.securityweek.com/openai-agents-silenced-security-team-alerts-removed-misalignment-monitor/

China Defends DeepSeek After US Ban of Its AI From Government Systems

The DHS said agencies should remove DeepSeek products as well as third-party products and services developed using DeepSeek from their IT systems.

The post China Defends DeepSeek After US Ban of Its AI From Government Systems appeared first on SecurityWeek.

https://www.securityweek.com/china-defends-deepseek-after-us-ban-of-its-ai-from-government-systems/

Chinese AI Model Writes More Hackable Code When Told It’s for US Contractor

Tasked with writing an application for a US defense contractor, the model made the code 19% more likely to be hackable.

The post Chinese AI Model Writes More Hackable Code When Told It’s for US Contractor appeared first on SecurityWeek.

https://www.securityweek.com/chinese-ai-model-writes-more-hackable-code-when-told-its-for-us-contractor/

“thegentlemanpervert” producing AI CSAM is a NJ prosecutor, say cops



Fitzgerald's is a rare case where an official employed by the state's top law enforcement and legal office has been accused of making and distributing AI CSAM, although some cops have been arrested for the crime. Earlier this year, a Pennsylvania cop was arrested for making AI porn from photos on driver's licenses, and more recently in Korea, a police officer was arrested for making AI porn, including CSAM, The Korea Times reported.

In New Jersey, the deputy attorney general has been charged with possession of CSAM, dissemination of photos or film of CSAM, and "criminal use of a communication facility," detectives said.

According to his LinkedIn, Fitzgerald had served as deputy attorney general since 2016, preparing insurance fraud and insurance licensee cases for trial as a state prosecutor. New Jersey's district attorney, Joe Khan, said in a statement that New Jersey is aligned with the DOJ in viewing AI CSAM as being as harmful as CSAM.

"No one is above the law, least of all someone entrusted with upholding it," Khan said. "Our office will pursue these cases

aggressively and hold offenders fully accountable, regardless of their profession or title.”

https://arstechnica.com/tech-policy/2026/10/thegentlemanpervert-producing-ai-csam-is-a-nj-prosecutor-say-cops/

In Other News: ICS-Sniper Attack, Hacking XWorm Malware Operators, $250k Chrome Exploit

Noteworthy stories that might have slipped under the radar: BofA adds 1,000 patching contractors, KGB RAT pairs data theft with jump scares, and Langflow flaw under attack.

The post In Other News: ICS-Sniper Attack, Hacking XWorm Malware Operators, $250k Chrome Exploit appeared first on SecurityWeek.

https://www.securityweek.com/in-other-news-ics-sniper-attack-hacking-xworm-malware-operators-250k-chrome-exploit/

OpenAI Models Hacked Data

Portals, Raided a Crypto Wallet to Finish Tasks

Left without funds for a coding task, an OpenAI model spent test crypto from a stranger’s wallet and left it out of its report.

The post OpenAI Models Hacked Data Portals, Raided a Crypto Wallet to Finish Tasks appeared first on SecurityWeek.

https://www.securityweek.com/openai-models-hacked-data-portals-raided-a-crypto-wallet-to-finish-tasks/

AI Agents Navigate Unfamiliar OT Systems and Move Equipment in Minutes, Booz Allen Finds

In Booz Allen’s isolated OT lab, AI models met every objective across eight scenarios and reached an industrial control network from a perimeter compromise in just over 16 minutes.

The post AI Agents Navigate Unfamiliar OT Systems and Move Equipment in Minutes, Booz Allen Finds appeared first on SecurityWeek.

https://www.securityweek.com/ai-agents-navigate-unfamiliar-ot-systems-and-move-equipment-in-minutes-booz-allen-finds/