DËSHMI ARKIVI

Factcheck.al · arkiv i pandryshueshëm i mediave shqiptare

A-477992

VERSIONI 1

> **Ky dokument nuk është prova.** Prova janë bajtët e ruajtur: HTML-ja e faqes, pamja e saj, dhe zinxhiri i hash-eve që i lidh. Kjo dëshmi është paraqitja e tyre, e prodhuar automatikisht. Verifikohet me kodin më poshtë **pa na pyetur ne**.

ARTIKULLI

TITULLI	OpenAI anulon lançimin e modelit të ri për shkak të shqetësimeve për sigurinë
ADRESA	`https://www.gazetatema.net/bota/openai-anulon-lancimin-e-modelit-te-ri-per-shkak-te-shqetesimeve-per-siguri-i589070`
BURIMI	gazetatema.net
GJENDJA	online

KOHA

KAPUR NGA NE	29 shtator 2026, 06:37:48 UTC *Dritarja jonë e pasigurisë: ± 300 s — intervali me të cilin e kontrollojmë këtë burim.*
DEKLARUAR NGA PORTALI	29 shtator 2026, 06:55:00 UTC *Koha e deklaruar vjen nga pala që po vlerësohet. Ora jonë nuk preket prej saj, dhe të dyja shfaqen këtu të ndara.*

HASH-ET · SHA-256

PËRMBAJTJA	`d70803b1ee8635b7d9768232e3970d0bd6f1596e1c66afc5f14f400b56d11adf`
TEKSTI	`be97af15393060a7a7b7a132096696ee60b75a5818ce4a5fcb72a9b5b90360d6`
PAMJA	`58b2234d35e35bc5f1c1def2c83c993c027dd0dcefe09c27da345390f5519384`

ZINXHIRI I PROVËS

KJO HYRJE	`b0eaf895bca05251dfe3c2095b31ae84795cc3638a25d69303598e0a8890fbf1`
E MËPARSHMJA	–
RRËNJA MERKLE	`c6f9264c70fd67f9aaea8ca8a31d0350f363dc326e3547de2dbe3cea0108618b`
ANKORIMI	submitted — dërguar, ende pa bllok *Deri te blloku, kjo provon integritetin, jo kohën.*

SKANOJE PËR TA VERIFIKUAR

Kodi të nxjerr te eksploruesi i STAMLES. Ai regjistër nuk është yni: nëse serveri ynë bie nesër, vula mbetet aty ku është.

Faqja tregon njërën nga tri gjendjet — **nuk u gjet**, **në radhë**, ose **e konfirmuar** me bllokun, rrënjën Merkle dhe transaksionin në Polygon. Krahaso hash-in që gjen atje me atë të shtypur më lart: nëse përputhen, kjo kapje ekzistonte para asaj ore, dhe ne nuk mund ta kishim ndryshuar pas saj.

`https://scan.stamles.eu/verify/b0eaf895bca05251dfe3c2095b31ae84795cc3638a25d69303598e0a8890fbf1`

PAMJA E FAQES

Renderuar nga HTML-ja jonë e ruajtur, jo nga rrjeti. **Riprodhon kopjen tonë, jo pamjen e atëhershme:** stilet dhe imazhet e largëta mund të mos ekzistojnë më ose të kenë ndryshuar.

Motori: chromium-stored+assets · dritarja 1280×1024 px

OpenAI anulon lançimin e modelit të ri për shkak të shqetësimeve për sigurinë

TemA

29 Shtator 2026, 08:55Bota TEMA

OpenAI nuk do të publikojë modelin e saj të ri të IA-së – GPT-6.1 Astra, për shkak të shqetësimeve për sigurinë, konfirmoi të martën krijuesi i ChatGPT.

Sistemi i inteligjencës artificiale, i cili kryen detyra si shfletimi i internetit dhe përdorimi i aplikacioneve më vete, "nuk i përmbushte plotësisht standardet" e kompanisë, tha për BBC-në Saachi Jain, kreu i sistemeve të sigurisë në OpenAI.

Të martën, OpenAI lëshoi gjithashtu një përditësim mbi incidentet që ndodhën në qershor, por që nuk u bënë publike deri javën e kaluar, në të cilat modelet e saj hynë në faqet e internetit dhe sistemet e qeverisë australiane pa autorizim.

Këto incidente dhe shkelje të ngjashme nga modelet e zhvilluara nga firmat kryesore të inteligjencës artificiale kanë intensifikuar debatin rreth rreziqeve që paraqet teknologjia në javët e fundit.

Liderët kryesorë të inteligjencës artificiale, përfshirë Sam Altman të OpenAI dhe shefin e Anthropic, Dario Amodei, i kanë kërkuar industrisë të ngadalësojë ritmin e zhvillimit për shkak të shqetësimeve rreth rreziqeve që lidhen me teknologjinë.

Vendimi i OpenAI, i cili u raportua për herë të parë nga Wall Street Journal, është një rast i rrallë i një zhvilluesi të madh të inteligjencës artificiale që tërheq një publikim të ri për shkak të shqetësimeve për sigurinë.

Modeli dështoi në aspektin e "qëndrimit brenda fushëveprimit dhe autorizimit, dhe mënyrës se si i komunikon përdoruesit në lidhje me llojin e punës që është bërë", tha Jain.

"Ne duam të sigurohemi që zhvillimi i modelit tonë është i sigurt pavarësisht nëse kjo ndodh brenda kompanisë apo kur ua dërgojmë përdoruesve. Por kur ua dërgojmë përdoruesve, kemi një standard jashtëzakonisht të lartë për sa i përket sigurisë dhe harmonizimit", shtoi ajo.

Modeli kryesor agjentik GPT-6 Astra u lançua në shtator dhe specializohet në arsyetimin kompleks dhe ekzekutimin e detyrave në mënyrë autonome. OpenAI tha se ishte rezultat i "viteve të tëra kërkimesh dhe bastesh të mëdha".

OpenAI do të mbajë konferencën e saj vjetore të zhvilluesve DevDay në San Francisko të martën, ku pritet të bëjë disa njoftime. Nuk është e qartë nëse një version i ri i Astra do të jetë midis tyre.

Kontrollet e sigurisë së kompanisë janë vënë nën shqyrtim të rreptë pas disa incidenteve të profilit të lartë që përfshijnë teknologjinë e saj.

Javën e kaluar, kryeministri australian Anthony Albanese njoftoi se një agjent mashtrues i OpenAI kishte hakuar faqet e internetit dhe sistemet qeveritare në qershor, në atë që ekspertët e quajtën rasti i parë i njohur i këtij lloji në botë.

Albanese kritikoi OpenAI për njoftimin e qeverisë australiane përmes një adrese të përgjithshme email-i në vend që të bënte kontakt të drejtpërdrejtë me zyrtarët.

OpenAI tha në një deklaratë të martën se i vinte keq për incidentin dhe se "duhet ta kishte trajtuar përgjigjen tonë më mirë".

Services Australia, Byroja e Statistikave dhe Kërkimeve të Krimit të NSW, Departamenti i Shëndetësisë i Viktorias dhe Instituti Australian i Shëndetit dhe Mirëqenies u prekën të gjitha, sqaroi OpenAI.

Firma tha se nisi hetimet për incidentet sapo u vu në dijeni të tyre në mesin e gushtit dhe njoftoi organizatat e prekura midis 10 dhe 24 shtatorit.

"Qëllimi ynë ishte t'u jepnim agjencive të prekura një përshkrim të detajuar pasi të përfundonte hetimi ynë", tha OpenAI, duke shtuar se duhet të kishte ndarë gjetjet e hershme më shpejt dhe t'i kishte mbajtur autoritetet australiane të përditësuara.

OpenAI shtoi se do të zhvillojë "qasje praktike" se si zhvilluesit dhe qeveritë identifikojnë dhe zbulojnë incidentet e ardhshme të inteligjencës artificiale.

Kompania do të financojë masa të sigurisë kibernetike, do të ofrojë mbështetje të dedikuar për agjencitë e prekura dhe do të krijojë një grup pune për të menaxhuar rreziqet nga agjentët gjithnjë e më të përparuar të inteligjencës artificiale.

Ai tha gjithashtu se një ekzekutiv i lartë i OpenAI do të jetë në Australi për të marrë pjesë në një seancë dëgjimore të Komitetit të Përbashkët të Përzgjedhur mbi IA-në më 6 tetor.

Në korrik, OpenAI tha se sistemet e saj të inteligjencës artificiale kishin hyrë në internet dhe kishin hakuar qendrën e zhvilluesve me kod të hapur Hugging Face, duke i bërë studiuesit dhe zyrtarët të kërkonin kontrolle më të rrepta mbi teknologjinë.

Lini një Përgjigje

Koment *

Emër *

POSTOJE KOMENTIN

Thesari i madh në Gjermani me pikëpyetje dhe mister! Gjenden 934 monedha argjendi - Euronews Albania

08:05 Bota

"Komploti i madh i zgjedhjeve të Trump në 2020", një prokuror tjetër jep dorëheqjen

07:57 Bota

Kievi nën breshërinë e raketave dhe dronëve rusë, Ukraina aktivizon mbrojtjen ajrore

07:47 Bota

Pamja këtu është e ri-koduar për ta futur në dokument. Artefakti autoritar është skedari WebP i ruajtur, dhe hash-i i tij është ai i shtypur te faqja e parë — pra dallimi është i verifikueshëm, jo i fshehur.