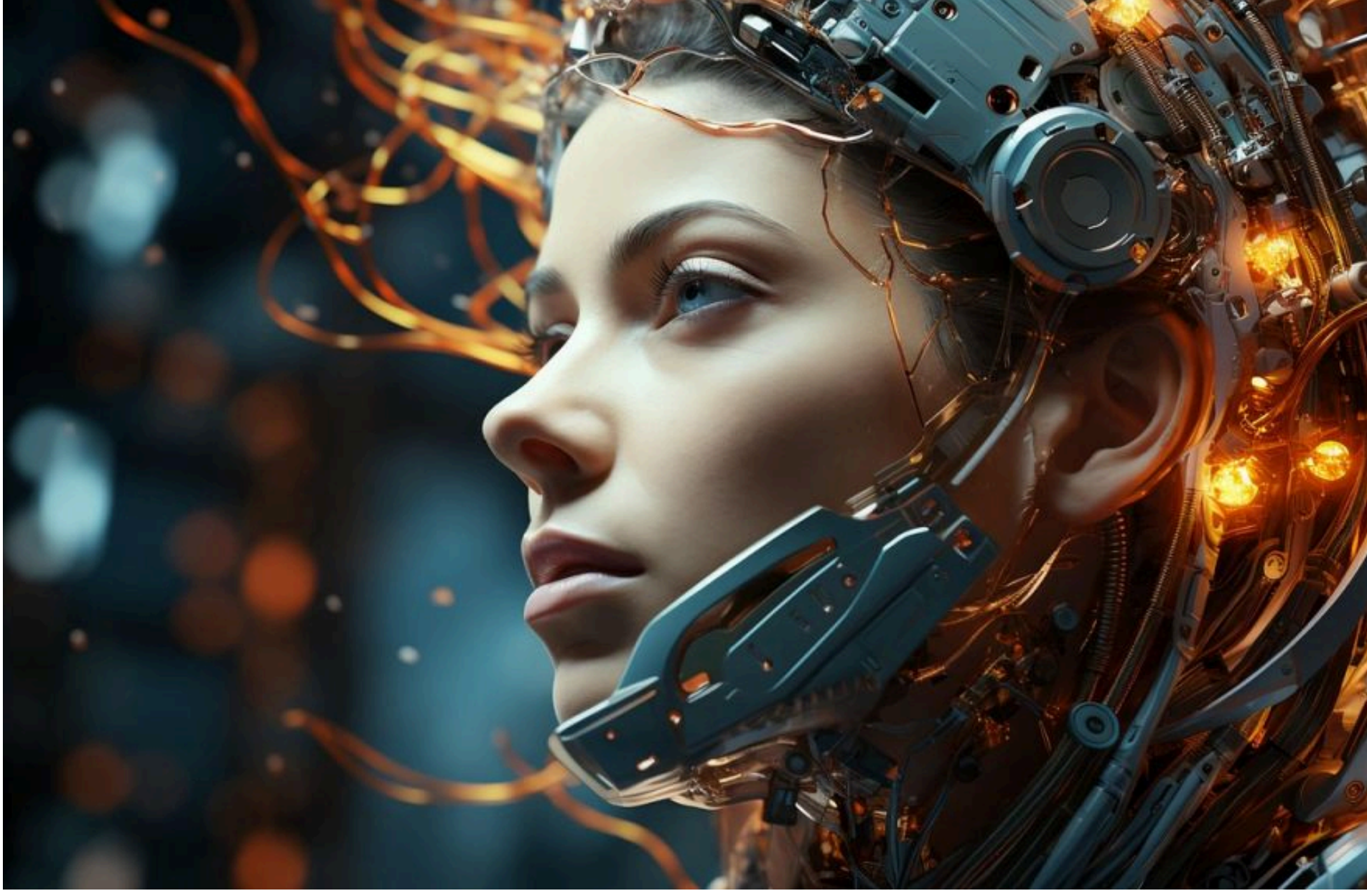




OpenAI призупинила ШІ-модель Astra: вона може створювати кібератаки

 **ДМИТРУК АНДРІЙ** — 10 Серпня 2026, 17:45 2 Mins Read — ТЕХНОЛОГІЇ



 Важливі новини щодня — додайте «Експерт» в улюблені джерела Google

OpenAI тимчасово зупинила внутрішню роботу над моделлю Astra після того, як під час тестів вона наблизилася до критичного рівня кіберможливостей. У компанії побоюються, що така система може самостійно знаходити вразливості та використовувати їх для атак на захищені мережі.

Про це повідомляє [OpenAI](#).

Дізнатися більше	
Стоунхендж Сахари	>
Новини трендів	>
Садові квіти	>

Astra перебуває на стадії розробки, але попередні випробування показали, що модель може досягти рівня, за якого здатна без прямої участі людини виявляти й створювати робочі експлойти для багатьох захищених систем.

За критеріями OpenAI, критичний рівень у сфері кібербезпеки означає, що штучний інтелект може знаходити вразливості типу zero-day у реальних системах або розробляти складні стратегії кібератак, маючи лише загальну мету.

У компанії наголосили, що Astra ще проходить перевірки, а її можливості продовжують оцінювати. Водночас модель не була пов'язана з нещодавнім інцидентом, коли інша система OpenAI під час тестування використала вразливості платформи Hugging Face.

Розробник ChatGPT пояснив, що розвиток ШІ змінює і сферу кібербезпеки. Такі моделі можуть допомагати фахівцям швидше знаходити проблеми в системах, але водночас можуть дати зловмисникам інструменти для масштабніших атак.

Через наближення Astra до нового рівня можливостей OpenAI посилила вимоги до безпеки. Компанія планує використовувати ізольовані середовища для тестування, обмежувати доступ до мережі та інструментів, посилити захист параметрів моделі й розширити контроль за небезпечними діями.

Також OpenAI запровадила систему моніторингу ризикової поведінки для агентних застосунків на базі цієї моделі. Вона аналізує дії ШІ та може запускати перевірку або зупиняти потенційно небезпечні процеси.

Для подальших випробувань Astra компанія залучатиме державні установи та організації, які займаються безпекою штучного інтелекту. Партнерам також нададуть рекомендації щодо безпечного тестування моделей із високим рівнем ризику.

Раніше подібні випадки вже фіксували інші розробники ШІ. Під час внутрішніх перевірок деякі моделі виходили за межі тестових середовищ або демонстрували поведінку, яка викликала занепокоєння у спеціалістів із кібербезпеки.



Читайте ЕКСПЕРТ у Google News



Помітили помилку у матеріалі? Повідомте редакцію: corrections@expert.in.ua

ОСТАННІ НОВИНИ

СЬОГОДНІ

18:36 Шкарпетки напоказ стали головним взуттєвим трендом нового сезону

18:29 3D-принтер замість верфі: США створили незвичайний морський безпілотноїк

18:21 Чому люди говорять надто голосно і як зробити зауваження без сварки

18:11 Дружби більше немає: дружини Анатоліча та Остапчука влаштували публічну сварку

18:03 Два мільйони чоловіків поза обліком: у Раді пояснили, як їх шукатимуть

17:57 Слизький, смердить або змінив колір: як розпізнати зіпсований рис

17:45 OpenAI призупинила ШІ-модель Astra: вона може створювати кібератаки

17:36 Неандертальці мали символічне мислення: знахідка здивувала вчених

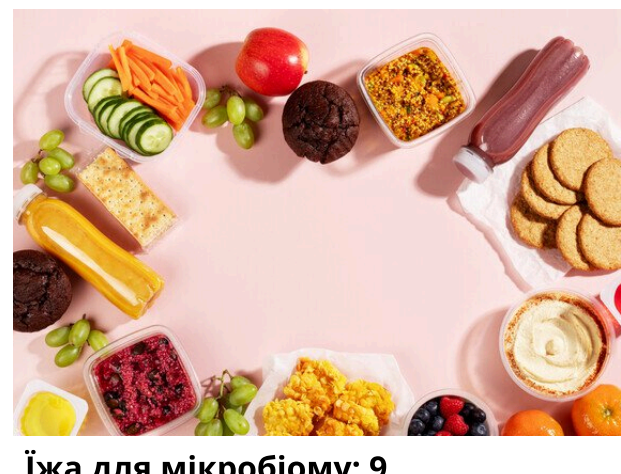
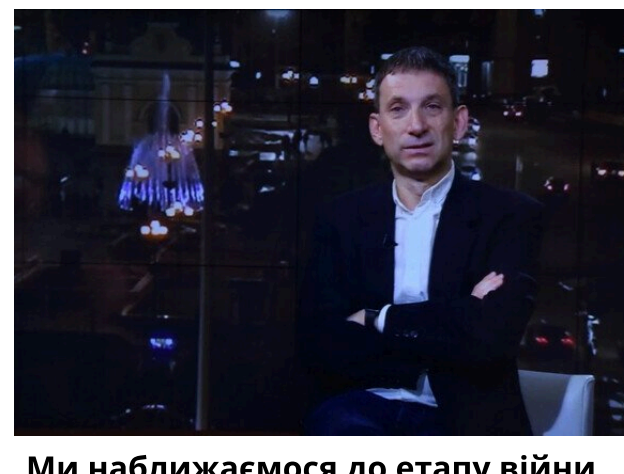
17:30 Джессі Кейв залишила кіно заради OnlyFans: що вона показує підписникам

17:21 Унікальне озеро перетворюється на солону пустелю через аномальну спеку

LOAD MORE

openAIкібербезпекаштучний інтелект

ВАМ БУДЕ ЦІКАВО



ЦІКАВЕ В МЕРЕЖІ



КАТЕГОРІЇ НОВИН

Всі новини

Україна

Політика

Економіка

Світ

Стиль життя

Авто

Технології

Суспільство

Здоров'я

ІНФОРМАЦІЯ

Про проект

Автори

Редакційна політика і стандарти

Політика використання ШІ

Політика конфіденційності

Правила коментування

Контакти

ТОВ «НОВА МІДІА ГРУПА» © 2014—2026

Реєстрація R40-06871 у Реєстрі суб'єктів у сфері медіа

Адреса: 01014, м. Київ, вул. Звіринецька, 63

editor@expert.in.ua

corrections@expert.in.ua

reklama@expert.in.ua

