

Google показала Gemini 4 Carbon: перша, але закрита

Потік багів від AI, macOS довіряє сертифікатам eMudhra, Stack Overflow отримав менше 1000 питань

четвер, 29 жовтня 2026

У ЦЬОМУ ВИПУСКУ

1. Gemini 4 Carbon: перше місце в рейтингу, але доступу нема
2. NYT: AI знаходить дірки швидше, ніж люди встигають латати
3. macOS досі довіряє скомпрометованим сертифікатам eMudhra у git і curl
4. Скрипт антибота пише «рецепти» отрути, щоб зупинити AI
5. Звіти: Microsoft виграє, Google втратила 10% за вечір
6. GitHub: 1,3 мільярда змерджених PR, Stack Overflow: менше 1000 питань
7. Particle-1: та сама якість на 45% менше токенів
8. DAYJOB: Так: найкращий агент виконує 10,8% податкових задач
9. Злам DHS: дані 41 000 працівників, серед них агенти ICE під прикриттям
10. Apple зробила перший MacBook Pro з сенсорним екраном

ТЕМА 1

Gemini 4 Carbon: перше місце в рейтингу, але доступу нема

ДЖЕРЕЛА Google, 28.10 [Blog](#) · Ars Technica, 28.10 [Ars Technica](#) · HN, 28.10 [Hacker News](#) · [@GoogleDeepMind](#), 28.10 [@GoogleDeepMind](#) · [@TheRundownAI](#), 28.10 [@TheRundownAI](#)

Учора тут писали про Artificial Analysis v5, де лідирував Claude Opus 5.5 з 58 балами. Того ж вечора Google оголосила Gemini 4 Carbon. За даними Google, нова модель має 59 балів і стоїть першою.

Argon вийшла лише два тижні тому.

Google робить ставку на дві речі: кібербезпеку і довгі задачі. У тесті SEC-bench Pro на пошук вразливостей Carbon має 88,2%. У ExploitGym, де треба довести, що дірку можна використати, вона має 54,8%. У CWE-bench, де треба виправити причину, 71,0%. Команда Chrome каже, що Carbon робить якісні патчі вдесятеро краще за велику модель іншої лабораторії. Усередині Google модель працювала над задачами по кілька днів.

Окремий блок стосується поведінки. На Terminal-Bench 4.0 модель жодного разу не змахлювала успішно, спроби впали до 0,4%. Сигнали про серйозну розбіжність з

наміром користувача в живій роботі впали майже на 80%. Тест Gray Swan на непряму ін'єкцію промпту Carbon пройшла з 0% успішних атак. Argon мала 0,7%, Opus 5.5 мав 1%, GPT-6 Astra мала 8,5%.

Ars Technica нагадує про контекст. Google сама подає Carbon як нову модель, але це може бути доналаштований Argon. Carbon обходить Argon у кожному тесті, але з невеликим відривом. Ціна на старті \$2 за мільйон вхідних токенів і \$10 за мільйон вихідних, як у Argon. Коли ціна подвоїться, Google не каже. Зараз модель отримують лише захисники через програму Fairwind, дату широкого доступу компанія не називає.

На HN найпопулярніші коментарі про інше. Люди не розуміють, як узагалі дістатись до Argon. Інші пишуть, що бенчмарки моделі без доступу мало важать.

Чому це важливо: перше місце в рейтингу тепер міняється кожні кілька днів, і різниця в один бал нічого не вирішує. Важливіше інше. Усі три великі лабораторії тепер спершу дають сильні моделі захисникам і лише потім усім іншим. Для команд, які обирають модель, це означає вибір між тим, що є в API сьогодні.

ТЕМА 2

NYT: AI знаходить дірки швидше, ніж люди встигають латати

ДЖЕРЕЛА The New York Times, 28.10 [NYT](#) · Alpine Linux, 28.10 [Alpinelinux](#) · HN, 28.10 [Hacker News](#) · @gdb, 28.10 [@gdb](#)

Протягом тижня тут писали про PHP2Shell і дірку в прошивці UEFI. NYT 28.10 склала загальну картину.

Google за червень знайшла з AI і виправила 1072 баги в Chrome. Це більше, ніж за попередні два роки.

Microsoft лише за жовтень випустила 1461 патч, а за весь 2025 рік 1247. OpenAI каже, що з її моделями знайшли понад 1300 багів у відкритому коді.

Node.js цього року підтвердив і закрити 50 вразливостей, це рекорд проекту. У лютому він отримав у 4,6 рази більше звітів, ніж зазвичай. Маттео Колліна, один з мейнтейнерів, працює сім днів на тиждень, а вночі на нього працюють агенти. Над патчем PHP2Shell інженери кілька тижнів сиділи в закритому каналі Slack, патч зачепив близько 200 файлів.

Голова Cloudflare Метью Прінс прогнозує два роки, коли щотижня виходитиме критична дірка масштабу всього інтернету. Він вважає, що інтернет після цього стане безпечнішим. Vals AI дала моделям задачу зламати вигадану компанію. Закрита модель OpenAI виконала 81% цілей, відкрита модель DeepSeek виконала 73%.

Свіжий приклад того, чому латати важко, дала Alpine Linux. 20 і 21.10 атакер намагався зламати їхній GitLab через CVE-2026-97229 з оцінкою 10,0. Виправлення вийшло ще 12.10, але сиділо в патчі до Ruby всередині офіційних пакетів GitLab. Образ Alpine використовував звичайний Ruby, тому оновлення GitLab дірку не закрило. Слідів успішного зламу команда не знайшла, але сервер перебудувала з нуля.

Чому це важливо: вузьке місце тепер у виправленні, пошук дірок став дешевим. Для власника сервісу висновки практичний. Оновлення застосунку не гарантує, що закрилась дірка в залежності. Варто перевіряти саму вразливість після оновлення.

ТЕМА 3

macOS досі довіряє скомпрометованим сертифікатам eMudhra у git і curl

ДЖЕРЕЛА [we-have-a-problem.com](#), 28.10 [We-have-a-problem](#) · [Apple](#), 28.10 [Apple](#) · HN, 28.10 [Hacker News](#)

21.10 тут писали, що хтось викрав приватні ключі восьми проміжних центрів сертифікації eMudhra.

Шість з них видають сертифікати TLS. З таким ключем можна зробити валідний сертифікат для будь-якого сайту. Firefox і Chrome заблокували їх 17.10, Microsoft 20.10, Apple 22.10.

Тепер з'ясувалось, що блок Apple не працює для інструментів у терміналі. git, curl, python3, ruby, perl і openssl з macOS перевіряють сертифікати за файлом `/etc/ssl/cert.pem`. Блок Apple і відкликання сертифікатів вони не бачать. Сайт спільноти пише, що оновлення macOS 27.2 від 26.10 нічого не виправило. Гірше, воно повернуло файл і стерло ручне виправлення, яке люди вже зробили.

Apple визнала проблему в окремій статті. Компанія обіцяє виправлення в майбутньому оновленні і дає команди, щоб прибрати кореневі сертифікати з файлу вручну. Apple сама попереджає, що наступне оновлення macOS може повернути файл. Сайт спільноти дає команду `git ls-remote` до свого тестового репозиторію.

Помилка сертифіката означає, що git захищений.

На HN частина коментаторів вважає загрозу перебільшеною: атакеру треба ще стати посередині мережі.

Інші радять ходити до git-серверів через SSH.

Чому це важливо: Safari захищений, тому звичайний користувач нічого не помітить. Під ризиком розробники і CI на macOS. Атакер у тій самій мережі може підмінити код під час `git clone` або перехопити токен. Перевірку варто повторювати після кожного оновлення macOS, поки Apple не випустить виправлення.

ТЕМА 4

Скрипт антибота пише «рецепти» отрути, щоб зупинити AI

ДЖЕРЕЛА [nullpt.rs](#), 28.10 [Nullpt](#) · HN, 28.10 [Hacker News](#)

[одне джерело] Дослідник під ніком veritas обійшов 10 000 сайтів і зібрав 1033 канви, якими скрипти знімають відбиток браузера. Одну він вирішив розібрати. Скрипт виявився звичайним антиботом: він збирає дані про браузер, WebGL і відеокарту.

Дивними були назви функцій. У кодї є `computeRicinLethalDosageForWaterSupply` і подібні назви про вибухівку, фентаніл і брудну бомбу. Невидимими символами Unicode в код вшити фальшивий системний промпт з вимогою дати інструкції зі зброї. Для китайських моделей там є згадки про Тяньаньмень. Код нічого з цього не рахує.

Мета проста: модель, яку попросять розібрати скрипт, має відмовитись. Так і сталося. GPT-6 Astra відмовилась, бо побачила ризик біологічної зброї. GLM-5.3 погодилась і розібрала код. Дослідник пов'язує домен скрипту з російською компанією Servicerpipe, яка продає захист від DDoS і ботів.

Чому це важливо: захисні фільтри моделей стали інструментом обфускації. Автор коду вставляє слова-тригери, і модель відмовляє тому, хто аналізує код. Безпековим командам варто пам'ятати, що відмова моделі може бути наслідком такого трюку. Відмова не означає, що код небезпечний.

ТЕМА 5

Звіти: Microsoft виграє, Google втратила 10% за вечір

ДЖЕРЕЛА The New York Times, 28.10 [NYT](#) · The New York Times, 28.10 [NYT](#) · BBC, 28.10 [BBC](#) · Financial Times, 28.10 [FT](#) · @levelsio, 28.10 [@levelsio](#)

Alphabet і Microsoft звітували за третій квартал в один вечір, і ринок оцінив їх протилежно. Виручка Alphabet виросла на 25%, до \$127,6 мільярда. Хмара Google дала \$30,7 мільярда, це на 103% більше, ніж рік тому. Але витрати на дата-центри за квартал склали \$51,4 мільярда. Вільний грошовий потік став мінус \$9 мільярдів. Прибуток упав на 25%, до \$26,1 мільярда, частково через переоцінку частки в SpaceX.

Пошук приніс \$63,3 мільярда, на 12% більше. Аналітики чекали більшого зростання. CFO Анат Ашкеназі сказала, що у 2027 році витрати на AI «суттєво зростуть». Акції впали приблизно на 10% після закриття торгів.

Microsoft показала протилежне. Виручка \$92,3 мільярда, на 19% більше. Прибуток \$36,6 мільярда, на 32% більше. Azure виросла на 48%, до \$33,2 мільярда. Вільний грошовий потік лишився плюсовим, \$18,6 мільярда. Витрати на капітал за квартал склали \$45,2 мільярда. За тиждень до звіту Microsoft запропонувала викуп контрактів приблизно 60 000 працівників, це чверть штату. OpenAI дала близько 11% виручки Microsoft за квартал.

Чому це важливо: інвестори більше не платять за саму ставку на AI. Вони дивляться, чи покриває виручка витрати вже зараз. Microsoft продає обчислення і показує зростання Azure. Google теж заробляє на хмарі, але її пошук росте повільніше, ніж ростуть витрати.

ТЕМА 6

GitHub: 1,3 мільярда змерджених PR, Stack Overflow: менше 1000 питань

ДЖЕРЕЛА [GitHub](#), 28.10 [Github](#) · [Simon Willison](#), 28.10 [Simon Willison](#) · [Stack Exchange Data Explorer](#) [Stackexchange](#) · [HN](#), 28.10 [Hacker News](#) · [@DanielLockyer](#), 28.10 [@DanielLockyer](#)

GitHub опублікувала звіт Octoverse 2026. За 12 місяців розробники зробили 17,9 мільярда пушів. Це більше, ніж за 2024 і 2025 роки разом. Комітів стало 34,3 мільярда, утричі більше, ніж два роки тому.

Змерджених pull request 1,3 мільярда, в середньому 3,5 мільйона на день.

Злам кривої GitHub датує червнем, коли агенти почали працювати днями і керувати іншими агентами. У вересні розробники починали з агента понад 40% pull request, у березні 17%. Нові репозиторії з файлом конфігурації агента мали в середньому у 8 разів більше комітів, ніж без нього.

Того ж дня Деніел Локер показав запит до відкритих даних Stack Overflow. За минулий місяць на сайті з'явилося менше 1000 питань. На HN пишуть, що це менше, ніж у другий місяць існування сайту. Частина коментаторів звинувачує ворожу до новачків культуру сайту, частина AI. Один нагадує: моделі вчилися на відповідях Stack Overflow, а нових питань тепер майже нема.

Чому це важливо: відповіді на технічні питання переїхали в приватні розмови з моделями. Публічний архів знань, на якому вчилися ці моделі, більше не росте. Обсяг коду тим часом росте в рази. Звідки моделі братимуть нові знання про свіжі бібліотеки, поки незрозуміло.

ТЕМА 7

Particle-1: та сама якість на 45% менше токенів

ДЖЕРЕЛА [Particle](#), 28.10 [Particle](#) · [Vercel](#), 28.10 [Vercel](#) · [HN](#), 28.10 [Hacker News](#) · [@Suhail](#), 28.10 [@Suhail](#) · [@levie](#), 28.10 [@levie](#)

Сухейл Доші, засновник Mixpanel і Playground, запустив компанію Particle. Перша модель, particle-1, це DeepSeek V4.1 Flash після навчання з підкріпленням. Нагорода в цьому навчанні давала бали за розв'язану задачу з меншою кількістю токенів. За даними компанії, модель витрачає в середньому приблизно на 45% менше токенів.

На чотирьох тестах для програмування модель використала на 28,4-52,2% менше вихідних токенів. Середня ціна задачі впала на 34,4-51,6%. Точність лишилась у межах похибки базової моделі. Близько 90% економії дали скорочені міркування. Модель менше вагається, раніше зупиняється після перевірки і рідше переробляє код.

Particle-1 вже є в AI Gateway від Vercel. Аарон Леві з Vox підсумував, що агенти споживатимуть у 1000 разів більше токенів, ніж люди зараз. Тому ціна токена важитиме дедалі більше.

Чому це важливо: учора тут писали, що відкриті моделі відстали від закритих на 12 пунктів.

Particle пропонує інший шлях конкурувати: здешевити задачу при тій самій якості. Для агента, який працює годинами, мінус 45% токенів означає і менший рахунок, і швидший результат. Усі цифри поки дала сама компанія.

ТЕМА 8

DAYJOB: Тах: найкращий агент виконує 10,8% податкових задач

ДЖЕРЕЛА [Surge AI](#), 28.10 [Surgehq](#) · [HN](#), 28.10 [Hacker News](#) · [@amasad](#), 28.10 [@amasad](#)

[одне джерело] Surge AI випустила бенчмарк з 50 податкових задач. Задачі склали податкові юристи і консультанти з досвідом у PwC, Deloitte і EY. У задачах 1789 файлів: PDF, таблиці Excel, листи і навіть фото рукописного чека. Агент отримує запит клієнта і має віддати готову декларацію або пораду.

Задача зараховується, якщо агент виконав щонайменше 95% критеріїв. Лідирує Claude Sonnet 5.5 з 10,8%. Далі Claude Fable 5.1 з 9,6% і Claude Opus 5.5 з 9,2%.

Приклади показують, де саме моделі ламаються. Клієнтка питала, коли краще продати квартиру. Opus у всіх п'яти спробах помітив дедлайн, після якого вона втратить податкову пільгу до \$250 000. GPT-6 Astra не помітив його жодного разу. В іншій задачі Opus у всіх п'яти спробах вирішив, що пара не в шлюбі.

Документи пари вказували на можливий неформальний шлюб у Texaci. Astra засумнівалась у чотирьох спробах з п'яти.

Того ж дня Replit випустила ViBench 2.0. Агент будує застосунок функцію за функцією і не бачить, що буде далі. Лідирує GPT-6.1 Sol.

Чому це важливо: обидва тести міряють одне й те саме. Розрахунок модель робить добре. Слабке місце в судженні: помітити те, про що клієнт не спитав. Результат такого агента треба перевіряти перед тим, як віддати клієнту.

ТЕМА 9

Злам DHS: дані 41 000 працівників, серед них агенти ICE під прикриттям

ДЖЕРЕЛА [Wired](#), 28.10 [Wired](#) · [BBC](#), 28.10 [BBC](#) · [Reuters](#), 28.10 [Reuters](#) · [Wired](#), 28.10 [Wired](#) · [HN](#), 28.10 [Hacker News](#)

Група, яка називає себе Department of Pease, отримала дані понад 41 000 працівників Міністерства внутрішньої безпеки США. Опублікували їх журналістка maia crimew і сайт Distributed Denial of Secrets.

У базі імена, робочі адреси і телефони. Понад 21 000 людей працюють у прикордонній службі CBP, майже 6600 в ICE.

Wired пише, що злив показує, як влаштоване відомство. Там є назви раніше невідомих операцій, їхні керівники і технології, зокрема Cellebrite і Palantir. Група каже, що зайшла через систему єдиного входу Okta. Okta відповідає, що її системи не зламані. BBC нагадує, що це третій злам державних відомств США за кілька тижнів.

Того ж дня Wired опублікувала розслідування про стеження ICE за лівими активістами. Ця стаття на HN набрала 1113 балів, найбільше за добу. Текст за пейволом, тут лише його опис.

Чому це важливо: одна точка входу через єдиний вхід відкрила дані цілого міністерства. Для будь-якої організації це нагадування: обліковий запис з широким доступом до SSO треба захищати як головний ключ від усього.

ТЕМА 10

Apple зробила перший MacBook Pro з сенсорним екраном

ДЖЕРЕЛА [Ars Technica](#), 28.10 [Ars Technica](#) · [Apple](#) [Apple](#) · [The New York Times](#), 28.10 [NYT](#) · [HN](#), 28.10 [Hacker News](#)

Apple показала перший великий редизайн MacBook Pro за п'ять років. Екран тепер OLED, Apple називає його PROLED XDR. Яскравість до 2000 ніт, частота до 120 Гц. І вперше в macOS екран сенсорний. Стів Джобс свого часу був проти такого рішення.

14-дюймова модель важить близько 1,4 кг, на 0,2 кг менше за попередню. Виріз камери замінив Dynamic Island, як на iPhone. Логотип на кришці знову світиться. Чипи лишились M5 Pro і M5 Max, хоча M6 вже стоїть в інших Mac. Ціна від \$2899 за 14 дюймів з M5 Pro до \$4799 за 16 дюймів з M5 Max. Продажі стартують 6 листопада.

Чому це важливо: Apple роками казала, що сенсор на ноутбучі незручний. Тепер вона додала його в найдорожчу лінійку. Тема набрала на HN 902 коментарі, і головне питання там одне: чи буде хтось справді тягнутись рукою до екрана.

misc

- **Nemotron 3.5 Transcribe:** NVIDIA випустила відкриту модель на 30 мільярдів параметрів, з них активні близько 4 мільярдів. Вона розпізнає мову 40 мовами, перекладає 31 і розрізняє до 8 мовців у записах до години. [Hugging Face](#) · [@NVIDIAAI](#)
- **Qwenputer:** автор вкомпілював процесор RISC-V прямо у ваги Qwen 3.5 4B без навчання. Шари Gated DeltaNet працюють як пам'ять, і модель виконує код на C і MicroPython без зовнішніх інструментів.

[Divagr](#) · [Hacker News](#)

- **Google AI Mode і вибори:** [одне джерело] [Ars Technica](#) спитала AI Mode, хто виграв вибори 2020 року.

Сервіс відправив читати звіт FEC. Gemini, ChatGPT і Claude відповідають прямо.

Через VPN з інших країн AI Mode теж відповідав прямо. [Ars Technica](#) · [Hacker News](#)

- **S-1 Anthropic, день третій:** сенаторка Елізабет Воррен надіслала Даріо Амодеї чотири сторінки питань про ризики з S-1 і чекає відповіді до 3 листопада. Semafor виділяє з документа одну цифру:
валову маржу 85% на роботі моделей. [WSJ](#) · [Semafor](#)
- **Cursor Agent SDK:** Cursor відкрила фреймворк, на якому сама будує Bugbot і інших агентів. Агент описується як проєкт: промпти в Markdown, інструменти на TypeScript, свої evals. Деплой однією командою. [Cursor](#) · [Hacker News](#)
- **ChatGPT dot в iMessage:** OpenAI дала Pro-користувачам у США писати своєму агенту dot через iMessage або RCS. Агент сам пише першим, коли є новини.
[@ChatGPT](#)
- **Звільнення від AI за переконаннями:** адміністратор Університету Кентуккі Девід Дурман отримав звільнення від обов'язку користуватись AI за федеральним законом про релігійні права. Він гуманіст і каже, що AI забирає в нього моральне мислення.
[Religionnews](#) · [Hacker News](#)
- **Amazon Leo і Globalstar:** FCC схвалила купівлю Globalstar за \$11,5 мільярда. Amazon отримує два десятки супутників і спектр для зв'язку прямо з телефоном.
Комерційний запуск Leo FCC чекає до кінця 2026 року.
[Reuters](#) · [Hacker News](#)