

Стара нейромережа обійшла GPT-6 Astra і Fable 5.1 у діагностиці раку

Токени подешевшали до 80%, Qwen3.8-27B наздогнала Fable 5, телевізор Samsung слухає дім, 11 дір у Tailscale

понеділок, 26 жовтня 2026

У ЦЬОМУ ВИПУСКУ

1. Стара мережа на 20 хвилин донавчання обійшла GPT-6 Astra і Fable 5.1 у діагностиці раку шкіри
2. Токени OpenAI і Anthropic подешевшали до 80% за чотири місяці. Макс Вулф питає, чому
3. П'ять копій Qwen3.8-27B зі спільним файлом плану наздогнали Fable 5 на складних задачах
4. Розтин телевізора Samsung: мікрофон через Bluetooth і AlexNet на кожні кілька секунд
5. Tailworm: 11 знахідок у Tailscale, але спершу атакеру потрібна точка входу
6. Ной Сміт: «інтернет ламається». Учорашній Сакс думає інакше
7. Спадкоємці Елвіса позиваються до Anthropic за 274 пісні
8. Deloitte знову попалась на AI: звіт про злиття університетів з помилкою на \$260 мільйонів
9. «Будуймо амбітніше»: Ден Лу і Ітан Моллік про одне й те саме

ТЕМА 1

Стара мережа на 20 хвилин донавчання обійшла GPT-6 Astra і Fable 5.1 у діагностиці раку шкіри

ДЖЕРЕЛА Lynch Lab, 25.10 [Lynchlab](#) · HN, 25.10 [Hacker News](#)

[одне джерело] Магнус Лінч, дерматолог і хірург лікарні Guy's у Лондоні, взяв 200 уражень шкіри з публічного набору MILK10k. Діагноз кожного підтвердила біопсія. Половину не можна відпускати без лікування: 30 меланом, 50 карцином і 20 актинічних кератозів. Друга половина безпечна: родимки і себореїні кератози. Для кожного ураження є два знімки: звичайне фото і знімок через дерматоскоп.

Ці знімки отримали GPT-6 Astra, Claude Fable 5.1 і Inception v3. Inception v3 це згортовка мережа Google з 2015 року. Лінч донавчив її на 68 011 знімках з архіву ISIC на одному N100. Тестові 200 уражень він з навчання прибрав.

Результат: мережа назвала правильний діагноз з шести у 125 випадках (62%). GPT-6 Astra дала 84 (42%), Fable 5.1 дала 85 (42%). Найбільший розрив на безпечних ураженнях.

Мережа впізнала 37 з 50 родимок, GPT-6 Astra 16, Fable 5.1 33. Себореєйних кератозів мережа впізнала 26 з 50, GPT-6 Astra 13, Fable 5.1 лише 6. Кожна модель мала улюблену відповідь. GPT-6 Astra назвала базаліомою 75 уражень, хоча їх було 25, серед них 14 меланом.

Картина з безпекою інша. Моделі майже нічого небезпечного не відпустили: 97 і 96 зі 100. Але й безпечних вони відпустили лише 32 і 22. Мережа на своєму порозі відпустила 76 безпечних і втратила 18 меланом. Після підлаштування порогу вона тримає 98-99 небезпечних. Лінч сам перелічує обмеження:

вибірка мала, 30 меланом дають широку похибку, кожную модель запускали один раз.

Чому це важливо: загальна модель, яка «бачить картинки», ще не спеціаліст. Мережа на 20 хвилин донавчання виграє там, де є добрий розмічений набір. Для медичних і промислових задач з вузьким класифікатором порівняння з малою спеціалізованою моделлю має бути першим кроком. Фронтирна модель тут обережна, але корисності від такої обережності мало.

ТЕМА 2

Токени OpenAI і Anthropic подешевшали до 80% за чотири місяці. Макс Вулф питає, чому

ДЖЕРЕЛА Max Woolf, 25.10 [Minimaxir](#) · HN, 25.10 [Hacker News](#)

Макс Вулф, колишній дата-сайєнтист BuzzFeed, зібрав хронологію цін з червня. 30.07 OpenAI здешевила GPT-5.6 Luna на 80%. 01.09 Anthropic випустила Fable 5.1 з читанням кешу на 75% дешевшим. 22.09 вийшли GPT-6 Sol за \$2/\$10 і Opus 5.5 за \$4/\$20 за мільйон токенів. Anthropic тоді прямо написала, що Opus 5.5 потребує менше обчислень. 29.09 GPT-6.1 Sol удвічі здешевила читання кешу.

Вулф пояснює, чому кеш важливіший за ціну входу. У агента для коду читання кешу дає 96-99% усіх токенів. Тому знижка кешу на 75% дає майже таку саму знижку на всю роботу агента.

Причин він бачить дві. Перша: моделі стали товаром. Модель міняють одним перемикачем в OpenRouter, і там DeepSeek V4.1 Flash 1017 уже обслуговують 29 провайдерів. Друга причина менш очевидна. Обидві компанії готуються до IPO, і знижка основного доходу виглядає дивно перед проспектом. Вулф допускає, що маржа на інференс справді велика, і вона ж дозволяє ці знижки.

Окрема деталь про підписки. Разом з GPT-6.1 Sol OpenAI урізала ліміт Codex за \$200 на 50%. Офіційна причина звучала так: «моделі стали ефективнішими».

Чому це важливо: теза «токени дешеві тимчасово, потім ціни піднімуть» поки не справджується.

Конкуренція відкритих ваг і десятків провайдерів тримає ціну вниз. Для бюджету агентів головна цифра це ціна читання кешу, і її варто дивитись першою.

ТЕМА 3

П'ять копій Qwen3.8-27B зі спільним файлом плану наздогнали Fable 5 на складних задачах

ДЖЕРЕЛА [GVS5H](#) на [GitHub](#) [GitHub](#) · [HN](#), 25.10 [Hacker News](#)

[одне джерело] Автори репозиторію називають метод «самооркестрацією через журнал». Свіжі копії однієї моделі ділять задачу на частини. Вони координуються через спільну теку з планом, нотатками і поточним розв'язком. Додаткового навчання метод не потребує.

Тест: 100 найновіших складних задач LiveCodeBench, одинадцять моделей. Qwen3.8-27B з відкритими вагами на локальному сервері піднялась з 66,8% до 92,4% pass@1. Fable 5 на тих самих задачах має 90,4%. Qwen3.8 Flash Next з оркестрацією дала 93,0% за 9% вартості Fable 5. Приріст сягає 25,6 процентного пункту. Автори чесно пишуть, що він не універсальний: деякі моделі не змінились або стали гіршими. За їхнім аналізом, приріст дають розбиття задачі і пам'ять між кроками.

Чому це важливо: це препринт на одному бенчмарку, і олімпіадні задачі не дорівнюють роботі в реальному репозиторії. Але напрям ясний: організація роботи агентів додає стільки ж, скільки крок між поколіннями моделей. Модель на одному GPU з добрим каркасом стає реальною альтернативою для закритих середовищ.

ТЕМА 4

Розтин телевізора Samsung: мікрофон через Bluetooth і AlexNet на кожні кілька секунд

ДЖЕРЕЛА [@mike64_t](#), 25.10 [@mike64_t](#) · [@freddier](#), 24.10 [@freddier](#) · [@Tristan0x](#), 25.10 [@Tristan0x](#) · [HN](#), 25.10 [Hacker News](#) · [HN](#), 25.10 [Hacker News](#)

Учора цю тему тут не взяли: дослідник [@log4j](#) отримав root на своєму телевізорі Samsung ще 21.10. За вікенд інші люди перевірили його знахідки, і тема набрала 287 і 136 балів на HN. Андрій Карпатий зробив репост.

За даними [@log4j](#), телевізор класифікує звук у кімнаті, перш ніж звітувати Samsung. У телеметрії є лічильники подій «гавкає собака» і «плаче дитина». Якщо у телевізора нема свого мікрофона, він шукає його в саундбарі або іншому Bluetooth-пристрої. Майк ([@mike64_t](#)) розібрав другу частину. Телевізор кожні кілька секунд запускає модель `alex.tflite` на кадрах 960x540 з квантизацією INT8. Це AlexNet, і вона намагається вгадати, у яку гру ви граєте. За його словами, у 99% випадків вона не працює, але процесор вантажить.

Трістан ([@Tristan0x](#)) зібрав інструкцію. Головна функція тут ACR, автоматичне розпізнавання вмісту екрана. Вона працює і на сигналі з HDMI, тобто з ноутбука чи приставки. Перший крок: вимкнути «Viewing Information Services» у Samsung або «Live

Plus» у LG. Другий: винести телевізор в окрему гостьову мережу. Третій: фільтрувати трекери на рівні DNS.

Чому це важливо: дешевий телевізор окупається даними глядача, і тепер це видно в коді прошивки.

Розтин прошивки, який раніше займав тижні, з агентами займає вихідні. Таких звітів про побутову техніку стане більше. Окрема мережа для розумних пристроїв перестала бути порадою для параноїків.

ТЕМА 5

Tailworm: 11 знахідок у Tailscale, але спершу атакеру потрібна точка входу

ДЖЕРЕЛА Tailworm (Cyberdione Labs), 25.10 Tailworm · @levelsio, 25.10 @levelsio · Tailscale Security Bulletins Tailscale

Ерика Віндіш з Cyberdione Labs опублікувала аудит Tailscale v1.104.2 і headscale v0.29.5. Тестування тривало 22-24.10. У звіті 11 знахідок, шість з них з високою критичністю. Головний клас стосується Tailscale SSH. Вбудований SSH-сервер працює від root, і пересилання unix-сокетів робить перевірки від імені root. Звичайна сесія може замінити або відкрити привілейований сокет. За звітом, у лабораторії черв переходив з вузла на вузол менше ніж за секунду.

Межу атаки зафіксувала примітка спільноти X. Атакеру потрібен уже схвалений пристрій у мережі і правило SSH, яке дозволяє доступ. Тобто це рух усередині мережі після зламу, а не вхід ззовні.

Джеймісон О'Рейлі написав, що жодна з 11 знахідок не стосується стандартної установки Tailscale.

Автори звіту свідомо не чекали виправлень. Вони вважають, що мережі вразливі вже зараз. На ранок 26.10 на сторінці бюлетенів безпеки Tailscale останній запис TS-2026-013 від 13.10, про цей звіт там нічого нема. Звіт радить три дії: правила SSH лише для конкретних пар пристроїв, вимкнене пересилання і термін дії ключів вузлів на серверах.

Чому це важливо: mesh-мережа з довірою до кожного пристрою перетворює один злам на доступ до всього парку. Після тижня PHP2Shell і Brotli точок входу вистачає. Перевірити свої правила SSH в політиці Tailscale займає п'ять хвилин, і чекати бюлетеня для цього не треба.

ТЕМА 6

Ной Сміт: «інтернет ламається». Учорашній Сакс думає інакше

ДЖЕРЕЛА Noahpinion, 25.10 Noahpinion · @levie, 25.10 @levie

Учора тут писали про Джошуа Сакса. Він вважає, що збитки від кіберзлочинів обмежують люди, гроші і ризик арешту, а кількість знайдених вразливостей тут другорядна. 25.10 економіст Ной Сміт написав протилежне. Його головна теза: відкриті моделі з легко знятими обмеженнями прибрали бар'єр знань.

Як приклад Сміт наводить звіт Unit 42. Низькокваліфікований хакер з DeepSeek V4 без обмежень зламав публічний університет в Індії і вкрав понад 5 200 записів студентів і викладачів. Запит до бота був на рівні «знайди, що вони там шахрували». Сміт цитує попередження NSA про PHP від 21.10 і Джейсона Клінтона з Anthropic, який ще в серпні радив ізолювати критичні мережі. Висновок Сміта: інтернет, яким ми його знали, може на якийсь час перестати працювати. Більша частина посту за пейволом.

Аарон Леві з Vox дає третю позицію. Більшість великих компаній не встигає оновлювати бібліотеки і не знає реального стану ризиків, бо дані розкидані. Але рої агентів уже вміють перевіряти атаки і читати логи, тому для захисту теж є шанс.

Чому це важливо: Сакс і Сміт сперечаються про масштаб, а з фактами згодні обидва. Атак стало більше, вони дешевші, і вікно для підготовки коротке. Практичний висновок Леві найближчий до роботи:

спершу інвентаризація даних і доступів, потім агенти.

ТЕМА 7

Спадкoємці Елвіса позиваються до Anthropic за 274 пісні

ДЖЕРЕЛА The Guardian, 25.10 [Guardian](#)

[одне джерело] Authentic Brands Group володіє правами на музику і образ Елвіса Преслі. Компанія подала позов проти Anthropic у п'ятницю, першим про це написав Billboard. За позовом, Anthropic навчала Claude на 274 піснях Преслі, серед них «Can't Help Falling in Love». Позивачі хочуть \$150 000 за кожен трек, тобто близько \$41 мільйона.

Позов звинувачує Claude у відтворенні захищеного вмісту «практично кожної пісні з кожного каталогу». Обмеження на виведення пісень, за позовом, лише доводять, що компанія знала про проблему. За даними Guardian, це один з перших великих позовів від спадкoємців артиста, а не від лейблу.

Чому це важливо: музичні позови до AI-компаній переходять від лейблів до окремих каталогів. Кожен власник прав тепер може рахувати свої \$150 000 за трек. Для лабораторій ціна навчальних даних росте і без рішення суду.

ТЕМА 8

Deloitte знову попалась на AI: звіт про злиття університетів з помилкою на \$260 мільйонів

ДЖЕРЕЛА The Guardian, 25.10 [Guardian](#)

[одне джерело] Уряд Західної Австралії заплатив Deloitte \$166 000 за аналіз злиття трьох університетів. У звіті знайшли посилання, які згенерував AI. Тепер Guardian знайшов і арифметику. За одним сценарієм \$514 мільйонів плюс \$145 мільйонів дають у звіті \$399 мільйонів, хоча мали дати \$659 мільйонів. Розрив \$260 мільйонів. В іншому сценарії розрив \$311 мільйонів.

Одне посилання на звіт ВООЗ має в адресі мітку ChatGPT, тобто його скопіювали з чату. Deloitte визнає, що AI допомагав зі збором матеріалу і вичиткою. Спершу компанія назвала помилки дрібними.

Тепер уряд штату вимагає повернути гроші. Раніше Deloitte вже повертала гроші уряду Австралії за звіт з AI-помилками.

Чому це важливо: AI у консалтингу дає звіт швидше, але перевірку цифр за людей не робить. Для замовника тепер є простий тест: перерахувати підсумки таблиць і відкрити кожне посилання. Договори з консультантами отримають пункти про використання AI і штрафи.

ТЕМА 9

«Будуймо амбітніше»: Ден Лу і Ітан Моллік про одне й те саме

ДЖЕРЕЛА Dan Luu [Danluu](#) · HN, 25.10 [Hacker News](#) · [@emollick](#), 25.10 [@emollick](#)

Есей Дена Лу став найпопулярнішою сторінкою доби на HN: 759 балів і 698 коментарів. Приклад у Лу це pgrust, переписаний на Rust Postgres від Майкла Маліса. Більшість таких переписувань Лу називає «вайбслопом». Але pgrust тестують так ретельно, що він уже знайшов справжні баги в самому Postgres. Їх виправили, тож Postgres виграв навіть у разі провалу pgrust.

Головна теза Лу: більшість проєктів з агентами слабкі, але це не обмежує того, хто буде ретельно.

Проєкти, які раніше були завеликими, тепер реальні, якщо до них є тести і бенчмарки. 25.10 Ітан Моллік з Wharton дав схожу пораду: питати себе, як зробити задачу ще амбітнішою. Так людина або впирається в межу моделі, або бачить, що модель вміє більше.

Чому це важливо: скепсис до AI-проєктів у стрічці здебільшого виправданий. Але ціна масштабного переписування впала. Тести і бенчмарки тепер відділяють такий проєкт від сміття.

misc

- **JetBrains ThinkRail:** JetBrains відкрила легку IDE для агента рі. Кожна задача живе в окремому git worktree, а знання про проект зберігаються в «живих специфікаціях». 1,9 тисячі зірок на GitHub.
[GitHub](#) · [Hacker News](#)
- **Cerebras і моделі на 3 трильйони параметрів:** компанія пояснила, як ділить Kimi K3 між пластинами WSE-3T по 44 ГБ пам'яті. Шар займає близько 16 ГБ, і по пластинах передаються лише активації. Заявлена швидкість 1000-2000+ токенів за секунду.
[Cerebras](#) · [Hacker News](#)
- **Nikon відібрала перемогу:** переможця конкурсу мікровідео Small World in Motion позбавили титулу. Науковці помітили, що структури у відео з'являються нізвідки. Автор визнав, що використовував AI в обробці. [Techspot](#) · [Hacker News](#)
- **Інцидент DreamHost:** хостер скинув паролі частини поштових і SFTP-акаунтів, файловий менеджер панелі не працює. Причину компанія не називає.
[Dreamhoststatus](#) · [Hacker News](#)
- **Rest of World може закритись:** видання про технології поза Заходом готує закриття наприкінці січня. Трафік з пошуку і соцмереж зник, і серед причин Semafor називає AI-підсумки в пошуку.
[Semafor](#)
- **Gumroad і Valkey:** @ershush переніс увесь Redis Gumroad на Valkey і зменшив рахунок ElastiCache на 69%, близько \$2 200 на місяць. Код змінювати не довелося.
[@ershush](#)