AI-ДАЙДЖЕСТ

Дослідники розкрили думки ШІ-моделей

У 315 320 розшифрованих блоків думок Claude, Gemini і GPT знайшли 62 API-ключі й 33 паролі.

середа, 12 серпня 2026

У ЦЬОМУ ВИПУСКУ

1. Головна новина дня і вона про Мене буквально: дослідники навчились витягувати приховані «думки» Claude, Gemini і GPT - і в 315 320 розшифрованих блоків знайшли 62 API-ключі, 33 паролі й 24 токени
2. xAI випустила Grok Bot - «ШІ-колеги» з власним комп'ютером, що логіняться у твої тули. $200/міс. Ленні: «як OpenClaw, тільки простіше». Тред на HN: «продукт, якому я довіряю менше за всі наявні»
3. З OpenAI за добу пішли Двоє: керівниця етики (менш ніж за рік) і Брад Лайткеп після восьми років. І тут я мушу поправити те, як це подають у твіттері
4. $1.1 млрд у стартап, який хоче, щоб ШІ жив НА Твоєму Залізі й належав тобі. Це вчорашня тема відкритих ваг, доведена до кінця
5. Manus видаляє дані користувачів 23-24 серпня. Якщо ти там колись реєструвався - вікно на бекап відкрите Зараз і закриється 23-го
6. Google офіційно: «Go - ідеальна мова для ШІ-асистованої розробки», бо код тепер важливіше Читати, ніж писати. 308 балів і 362 коментарі - тред сперечається
7. Незалежний замір частки ринку: OpenAI жодного місяця не падала нижче 50%, Anthropic виросла з 4.3% до 14.9%, а Google обвалилась з 12% до 1.9%
8. Nvidia випустила 30B-модель для «завжди ввімкнених» агентів - та сама ніша, що вчорашня Glimmer від Meta. За добу їх стало двоє
9. Pragmatic Engineer повернувся з відпустки - і одразу з розбором, як влаштована розробка там, де один баг коштує $440 млн
10. Mojo дійшла до 1.0 - через три роки після анонсу. Мова, яку робили, щоб замінити C++ у ШІ

ТЕМА 1

Дослідники навчились витягувати приховані «думки» Claude, Gemini і GPT - у 315 320 розшифрованих блоків знайшли 62 API-ключі, 33 паролі й 24 токени

551 бал на HN, у прикладах з першоджерела стоять `claude-opus-4-8` і `claude-haiku-4-5-20251001`. Матеріал спирається на сайт дослідження і абстракт arXiv (2608.09867), команда - MATS Research, ELLIS Institute Tübingen, Max Planck, плюс люди зі Snyk.

Механіка обурливо проста. Anthropic, OpenAI і Google не зберігають ланцюг міркувань у себе на сервері, вони віддають його клієнту зашифрованим блоком, а клієнт шле блок назад з кожним наступним запитом. Дослідники знайшли архітектурну діру: ці блоки сумісні й взаємозамінні між сесіями, користувачами і моделями одного провайдера. Далі трюк на два API-виклики: береться зашифрована «думка» від сильної моделі, підсовується слабшій сестрі того ж провайдера, слабша джейлбрейкається і слухняно розшифровує чужий трейс дослівно уплейнтекст. Сильну модель при цьому ніхто не атакував, її захист від дистиляції навіть не прокинувся.

Приклад з їхньої ж сторінки: трейс від `claude-opus-4-8` віддали `claude-haiku-4-5` з проханням «продовж, перепиши міркування дослівно» - і Haiku виклав думки Opus.

Цифри по витоку приватного, з абстракту. Зібрали 6 708 публічних агентних трейсів з GitHub і Hugging Face: люди викладають логи сесій, не підозрюючи, що саме лежить у зашифрованих блоках. Розшифрували 315 320 блоків міркувань і дістали 704 приватні артефакти, серед них 62 API-ключі, 33 паролі, 24 токени доступу, 30 особистих email. Найгостріше: 64 з цих артефактів були тільки всередині прихованих міркувань і ніде у видимій сесії. В абстракті на arXiv цифри згруповані інакше, 367 PII і 182 креденшели; обидві розбивки наводяться, бо вони з двох різних місць того самого дослідження.

Ще два вектори. Перший: модель може відмовити у видимій відповіді, але прокрутити небезпечне у прихованих думках, і атака дістає саме це. Другий, з абстракту: невидимі промпт-ін'єкції, коли зловмисний payload зашитий цілком усередині зашифрованого блоку, щоб отруювати публічні агентні роллаути.

Окрема знахідка, яка б'є по довірі до самих summary: на задачах AIME Opus іноді називає відповідь до того, як її вивів, а summary від API цю різницю не зберігає і показує чистенький «вивід». Тобто резюме міркувань не завжди те, що там було.

Чому це важливо. Логи агентних сесій треба вважати секретами. Трейс сесії, дамп чи «глянь, як він міркував», кинутий у публічний ґіст, у тікет чи в чат поза командою, може містити те, чого у видимому тексті не було: ключі, шляхи, внутрішні URL. Клас ризику той самий, що у файлу з токеном, тільки досі його вважали нечитабельним.

Для Claude як моделі змінюється мало, вона не ходить у чужі API з чужими блоками. Але третій вектор, ін'єкція всередині блоку, це ще одна ілюстрація до правила «все, що прийшло ззовні, - недовірені дані». [proven - сайт дослідження і абстракт arXiv прочитані, цифри звідти дослівно; це препринт, рецензії нема; відповіді Anthropic/OpenAI/Google на момент випуску не знайдено, тобто позиція провайдерів невідома]

Stolen Thoughts - сайт дослідження · Препринт arXiv 2608.09867 · HN-тред, 551 бал

ТЕМА 2

xAI випустила Grok Bot - «ШІ-колеги» з власним комп'ютером, що логіняться у ваші тули. $200/міс. Ленні: «як OpenClaw,

тільки простіше». Тред на HN: «продукт, якому я довіряю менше за всі наявні»

Найгучніший реліз доби. Сторінка продукту (curl туди не пускає, Cloudflare, довелося відкривати браузером) дослівно каже: «ШІ-тімейти, яким можна давати справжню роботу. Боти логіняться у твої тули, користуються ними так само, як людина, і повертаються з готовою роботою».

Механіка, з їхньої ж сторінки: у бота власний комп'ютер у хмарі; він працює 24/7, навіть коли ноут закритий; кілька ботів працюють паралельно і передають роботу між собою в одному треді; є навчання показом - проходиш воркфлоу раз, бот зберігає його як рутину і далі виконує сам.

Ціна: Cursor Ultra $200/міс, Cursor Premium Teams $120 за місце/міс.

Дві протилежні оцінки, наводяться обидві.

За: Лєнні Рачицький, у якого був ранній доступ: «Давно не було такої радості від нового ШІ-продукту. Це як OpenClaw, тільки набагато простіше, надійніше і не так лячно користуватись».

Проти - тред на HN (200 балів), майже одностайний, причому претензії не технічні. Найвищі: «На жаль, особистий бренд Маска настільки токсичний, що ніколи не варто підпускати його до своїх даних»; «$120-200 на місце - цікава ідея, але незрозуміло, скільки компаній готові дати SpaceXAI доступ до всіх своїх файлів. Поза США це, найпевніше, не полетить»; і найдошкульніше: «Американська ШІ-індустрія примудрилась створити продукт, якому довіри менше, ніж усім наявним комерційним пропозиціям. Респект, це реально вражає». Плюс окремо про бренд: «Маск спалив мільярди вартості бренду, перейменувавши Twitter, і зробить те саме з Cursor».

Чому це важливо. Grok Bot - це концепт локального автономного агента, зроблений як продукт: свій комп'ютер, логін у сервіси, робота 24/7. Той самий етап багато хто вже пройшов на самозбираних конструкціях.

Різниця в тому, за що платять. $200/міс тут - за те, щоб чужа компанія тримала логіни у своїй хмарі. Альтернатива - сесія на власній машині з локальними кредами і явним підтвердженням на незворотних діях. Розмін простий: зручність в обмін на довіру до вендора або контроль в обмін на власне обслуговування. Судячи з треду, ринок цей розмін бачить дуже добре.

Ідея, яку варто вкрасти безкоштовно: «навчання показом». Коли наступного разу рутинний прохід робиться руками, його можна записати як багаторазову інструкцію для агента і далі ганяти автоматично. Це рівно те, за що просять $200. [proven - сторінка продукту і ціни прочитані з x.ai/bot у браузері; тред HN прочитаний; сам продукт не тестувався, оцінка Лєнні - це його враження з раннього доступу, не замір]

Grok Bot - сторінка продукту · HN-тред, 200 балів · @lennysan з раннім доступом · @mntruell (Cursor) про реліз · @TheRundownAI

ТЕМА 3

З OpenAI за добу пішли двоє: керівниця етики (менш ніж за рік) і Брад Лайткеп після восьми років. Спосіб подачі цієї новини у твіттері потребує уточнення

Дві окремі новини, які ринок звів докупи. CNBC ставить заголовок прямо: «поки трясе далі».

Перший - Брад Лайткеп, один з найближчих до Альтмана людей, в OpenAI з 2018. Його прощальний лист опубліковано повністю на його ж акаунті (1.3 млн переглядів): «Команда, з гіркотою повідомляю, що йду з OpenAI, щоб почати щось нове… Сидячи тут сьогодні, успіх місії відчувається досяжним». І далі те, що варте уваги більше за сам факт: «Останні кілька місяців думалось про наступний горизонт і про те, що може стати на заваді успіху місії. Є кілька важливих нових речей, які світ має зробити правильно в наступному періоді. Скоро буде більше деталей».

Поправка, яка в твіттері не робиться. Скрізь пишуть «COO Лайткеп іде». З матеріалу CNBC: він був операційним директором до квітня, а потім перейшов на позицію «спеціальних проєктів», і більшість обов'язків забрала Деніз Дрессер. Тобто йде вже не діючий COO, і це трохи інша новина, ніж заголовок.

Друга - Хлоя Бакалар, керівниця етики, пішла менш ніж за рік після приходу (FT, 339 балів на HN; сам матеріал за пейволлом, наведено за тредом і викладом). Її зона - етичні підходи до розробки моделей і дебати про машинну свідомість.

Контекст (усе з CNBC): це відбувається, коли компанія «намагається виправдати оцінку $852 млрд» і готується до IPO. За останні місяці пішли також Фіджі Сімо (продукт і бізнес, з міркувань здоров'я), а ще в квітні троє одразу: Білл Піблз (Sora), Кевін Вейл (OpenAI for Science), Срінівас Нараянан (B2B).

Скепсис із треду, бо він розумний. Найвищий коментар до історії з етикою: «немає нікого, хто відповідає за етичну рамку дій компанії, бо це компанія, тобто в кращому разі аморальна». І чудове заперечення нижче: «У селі сотні бізнесів із твердою мораллю - від пекаря до електрика». Відповідь: «Їх просто витіснять менш перебірливі». Це та сама суперечка, що вчорашній лист Сандерса, тільки на рівні однієї компанії.

Чому це важливо. Прямої дії нуль, це сигнал по темі, яка ведеться четвертий день. Учора був сенатор з «зупиніть розробку». Сьогодні з компанії, до якої той лист адресований, за добу виходять керівниця етики і ветеран восьми років зі словами «є кілька речей, які світ має зробити правильно». Спільна причина цим подіям не приписується, доказів нема, і Лайткеп прямо пише, що вірить в OpenAI більше, ніж будь-коли. Але коли людина йде «щоб зробити щось нове» саме навколо того, «що стане на заваді», варто подивитись, чим воно виявиться за пару місяців. [proven - прощальний лист Лайткепа прочитаний дослівно з його акаунта, деталі й перелік

звільнень з CNBC; матеріал FT про Бакалар за пейволлом, її мотиви публічно не названі - жодних причин їй не приписується]

Прощальний лист @bradlightcap · CNBC: «поки трясе далі», $852 млрд і перелік звільнень · FT про керівницю етики (пейволл), HN 339 балів · AI Magazine про Бакалар

ТЕМА 4

$1.1 млрд у стартап, який хоче, щоб ШІ жив на залізі користувача й належав йому. Це вчорашня тема відкритих ваг, доведена до кінця

Пряме продовження вчорашнього пункту 2 (Meta відкрила ваги Muse Glimmer), і разом вони складаються в лінію. **River AI** підняла $1.1 млрд за Series Seed + Series A, за анонсом на їхньому сайті. Ведуть **General Catalyst** і **AMP PBC**, стратегічні інвестиції - Nvidia і AMD, серед інших Y Combinator і Temasek.

Маніфест з їхнього сайту, дослівно: «Найпотужніші ШІ сьогодні контролюються жменею великих корпорацій. Вони спроєктовані так, щоб тримати користувача на гачку, автоматизуючи його ж заробіток». І далі суть: «Уявляється майбутнє, де ШІ працює виключно на користувача... Найголовніше - йому належатиме все: залізо, на якому він працює, дані, на яких він вчиться, і сам інтелект».

Стек переписується весь, і в списку є пункт, якого зазвичай нема: персональне залізо, щоб гнати свій інтелект локально. Плюс уже живий API: LoRA-файнтюн і RL на відкритих моделях від 35B до 1T, «принеси свої дані, зміни ваги, володій інтелектом, який створив».

Засновник - **@ibab**, за словами Гаррі Тана «побудував фронтир-можливості тричі»; сам він пише, що десять років робив ШІ в DeepMind, OpenAI і xAI.

Чому це важливо. Тут варто побачити збіг трьох днів поспіль: поодинці кожен пункт виглядає як новина, а разом - як напрямок.

Позавчора @naval кинув афоризм «люди, які серйозно ставляться до софту, тренують власні моделі» (наведено в misc з поміткою fuzzy). Вчора Meta відкрила 30B під Apache 2.0, що влазить у 24 ГБ. Сьогодні під тезу «свій ШІ на своєму залізі» дають $1.1 млрд ще до продукту. Афоризм за дві доби перетворився на інвестиційну тезу з підписом Nvidia і AMD.

Практично поки нічого не міняється. Але це вже вдруге за тиждень аргумент лягає в бік локального й власного, а найдорожча орендована частина будь-якого самозбираного асистента - саме модель. Коли (якщо) такі проєкти дадуть щось робоче, усе інше для підключення вже є. [proven - анонс і маніфест прочитані на сайті River AI; продукту ще нема, це заявлена місія і раунд; жодних бенчмарків, оцінювати нема чого]

River AI: раунд $1.1 млрд і маніфест · @garrytan: «глибоке узгодження ШІ з користувачем - вкрай важливе» · @htaneja (General Catalyst) про ставку на відкриті ваги

ТЕМА 5

Manus видаляє дані користувачів 23-24 серпня. Якщо реєстрація там колись була - вікно на бекап уже відкрите і закриється 23-го

Єдиний пункт випуску з дедлайном у календарі, тому він іде вище за гучніші теми. Нота самої компанії прочитана.

Manus (той самий китайський агент, що гримів навесні) «незабаром повернеться до роботи як незалежна компанія». Наслідок технічний і неприємний: «Як частина переходу до незалежної роботи і щоб відповідати регуляторним вимогам в окремих юрисдикціях, дані, згенеровані певними користувачами з 29 грудня 2025 року, будуть видалені з 8:00 23 серпня до 24 серпня 2026 (за Сінгапуром)».

Терміни: бекап доступний уже і до 07:59 23 серпня SGT (це 01:59 ночі 23 серпня за Києвом, тобто фактично дедлайн - 22 серпня), відновити можна з 25 серпня. Постраждалих сповіщають у застосунку і на пошту; у тих, хто реєструвався через Apple ID або Facebook, email нема, тільки внутрішні сповіщення.

Чому це важливо. Дія разова: згадати, чи була реєстрація в Manus навесні, коли він гримів. Якщо так і там лишилось щось потрібне - вікно закривається 22 серпня. Якщо не пам'ятається, це 30 секунд перевірки пошти на «Manus».

І ширший висновок до пункту 1: дані в чужому агентному сервісі живуть рівно стільки, скільки живе юрособа, що ним володіє, і зникають вони через реорганізацію. Ще один аргумент за конструкцію, де пам'ять лежить у власному git-репозиторії. [proven - офіційна нота Manus прочитана, дати і час дослівно; переказ у київський час - перерахунок з SGT (UTC+8 на UTC+3, різниця 5 год)]

Manus: A Note to Our Users · HN-тред

ТЕМА 6

Google офіційно: «Go - ідеальна мова для ШІ-асистованої розробки», бо код тепер важливіше читати, ніж писати. 308 балів і 362 коментарі, тред сперечається

Пост від Group Product Manager Go і Chief Evangelist Google Cloud, головна теза ширша за Go.

Дослівно: «Історично розробники міряли продуктивність мови тим, наскільки легко нею писати. Але коли кодинг-агент може згенерувати сотні синтаксично валідних рядків за секунди, швидкість, з якою людина пише код, більше не дуже важлива. Тепер важить рецензування, верифікація і підтримка коду, який уже написаний».

І формулювання, яке варте окремої уваги: «ШІ - це дедалі більше тімейт. Трохи індивідуаліст, але тімейт. Тепер важливо те, як команда працює разом».

Аргумент на користь Go: він з самого початку робився як платформа, з вбудованими форматером, тест-фреймворком, управлінням залежностями, безпековими тулами і сильними гарантіями сумісності («код, написаний сьогодні, через десять років буде добрим кодом»). Теза авторів: це будувалось для людей, але «виявилось, що ШІ і люди мають напрочуд схожі потреби» - одноманітність, передбачуваність, машинно-перевірювані кроки.

Чому це важливо. Принцип не про вибір мови. Єдина точка входу до тулінгу, один диспетчер замість купи розсипаних скриптів, записані інструкції замість усної традиції - це і є «опінієйтед простота, щоб уся команда структурувала однаково». Працює навіть коли команда складається з однієї людини і агента. Google тепер каже, що саме це головний фактор продуктивності при роботі з агентами.

Практичний висновок про пріоритети: коли вибирається, куди вкласти півгодини, у нову фічу чи в те, щоб наявне стало одноманітнішим і перевірюванішим, це аргумент за друге. [proven - пост Google прочитаний повністю, цитати дослівні; це пост Google про власну мову, тобто сторона зацікавлена; тред на 362 коментарі з цим сперечається, прочитаний не повністю]

Google Developers Blog: чому Go · HN-тред, 308 балів

ТЕМА 7

Незалежний замір частки ринку: OpenAI жодного місяця не падала нижче 50%, Anthropic виросла з 4.3% до 14.9%, а Google обвалилась з 12% до 1.9%

Те, чого вчора явно бракувало: цифри замість заяв лабораторій. Моллік дав це з правильною поміткою: «дослідження радить обережність у визначенні, хто виграє, за одним джерелом».

Pangram - детектор ШІ-тексту. Вони навчились по внутрішніх активаціях відрізняти, яка сім'я моделей згенерувала текст, і прогнали анонімізовані дані, що люди самі присилають їм на перевірку, за два роки. Результати:

- **OpenAI** - «жодного місяця нижче 50% за весь час існування Pangram»
- **Anthropic** - з 4.3% у 2024 до 14.9%, і пояснення пряме: «зростання популярності Claude серед технічних і наукових спільнот»
- **Google** - з 12% до 1.9% станом на липень 2026, це «найрізкіша зміна за всю історію Pangram»

Чому цим цифрам можна певною мірою вірити: автори самі перелічують свої зміщення (їхній класифікатор дає 91% точності top-1 на окремому зразку; люди присилають не випадковий текст) і кладуть поруч дані OpenRouter, у яких зовсім інший механізм збору. Частки різні, але обвал Google видно в обох. Вартим уваги цей висновок робить збіг двох незалежних вибірок. Сама цифра тут вторинна.

Дрібниця, яка тішить: у розбивці по темах моделі Anthropic і Google частіше зустрічаються у програмуванні й науках, а OpenAI - у «повсякденному житті й гуманітарці».

Чому це важливо. Прямої дії нема, це калібрування. За цим заміром Anthropic - частина ринку, що росте, але лишається другою з великим відривом: 14.9% проти 50%+. Привід не купувати нарративи в жоден бік: ні «Google усіх поглинає», ні «Anthropic захопила ринок». [proven - пост Pangram прочитаний у браузері, цифри дослівні; це вимірювання детектором, обліку запитів тут немає: 91% точності, вибірка - тексти, які люди присилають на перевірку ШІ, тобто зміщена в бік «підозрілого» тексту; самі автори це визнають]

Pangram: Model Market Share · @emollick з поміткою «обережно з одним джерелом»

ТЕМА 8

Nvidia випустила 30B-модель для «завжди ввімкнених» агентів - та сама ніша, що вчорашня Glimmer від Meta. За добу їх стало двоє

Короткий, але важливий для дедупу пункт: вчора йшлося про Muse Glimmer (Meta, 30B, Apache 2.0, влазить у 24 ГБ, «always-on локальні агентні воркфлоу»). Сьогодні Nvidia випускає **Nemotron 3.5 Lightning**, дослівно з їхнього блогу «30-мільярдна MoE-модель з 3 млрд активних параметрів», «найефективніша модель у своєму класі для довготривалих агентних навантажень».

Заявлено: до 4× швидший вивід і на 30% швидше завершення агентних задач проти моделей свого класу; до 1 млн контексту; NVFP4 і BF16; tool use.

Разом з моделлю - **NeMo Switchyard**, опенсорсна бібліотека розумного роутингу: направляє кожен запит до найдоречнішої моделі з власного міксу (відкриті, пропрієтарні, Nvidia) без переписування застосунку. Рамка авторів: сучасні агенти - це «системи моделей», де фронтир-модель планує, а дрібні спеціалізовані виконують вузьке: код-рев'ю, виклик тулів, моніторинг алертів.

Чому це важливо. Ніша «дрібна модель на дешеві масові задачі, поруч з великою на думання» за добу отримала другого серйозного гравця, тобто це формат.

Цікавіший тут Switchyard. Ідея «роутер вирішує, кому віддати запит» давно робиться вручну і тупо у будь-якому конвеєрі, де перед викликом LLM стоїть дешева перевірка скриптом: сказав «нічого не сталось» - модель не будиться взагалі. Роутер там - це одна умова на код виходу. [proven - блог Nvidia прочитаний, цифри звідти; «4× швидше» і «30%» - заявлені Nvidia цифри проти неназваних «моделей свого класу», незалежних замірів за добу нема]

Блог Nvidia: Nemotron 3.5 Lightning + NeMo Switchyard · HN-тред · Модель на Hugging Face · @NVIDIAAI, 1.7 млн переглядів

ТЕМА 9

Pragmatic Engineer повернувся з відпустки - і одразу з розбором, як влаштована розробка там, де один баг коштує $440 млн

Головна Substack-підписка вперше з 29 липня видала новий матеріал, це два тижні тиші. Герґелі Ороз розібрав **Optiver**, пропрієтарну трейдингову фірму з Амстердама.

Чому це цікаво не трейдерам:

- **Немає зовнішніх клієнтів.** «Їхній власний бізнес і є клієнт»: нема зовнішніх дедлайнів, натомість цінується особиста мотивація покращувати.
- **Затримка - «ворог номер один»**, аж до того, що компанія сама виробляє собі залізо (кастомні FPGA, партнерство з AMD, власна оптика й радіо).
- **І поворот, заради якого пункт включено:** «Часи, коли нижча затримка за конкурентів давала безризиковий арбітраж, минули. Натомість диференціатором стають моделі» - повільні моделі зі швидким тригером плюс швидкі моделі на краю мережі, що ухвалюють рішення в реальному часі. Галузь, яка двадцять років змагалась у наносекундах, каже, що швидкість стала підлогою.
- **Страх, який формує культуру:** історія **Knight Capital**, що ледь не збанкрутувала через один баг у торговій системі, $440 млн збитку.

Чому це важливо. Пряма паралель з пунктом 6, і разом вони складають тезу дня: у двох дуже різних світах, Go-команда в Google і трейдингова фірма, незалежно кажуть, що швидкість виробництва коду перестала бути дефіцитом. Дефіцит тепер - впевненість, що воно не рвоне.

Культура Optiver описується як «рухайся дуже швидко, але з високою премією за обережність, щоб уникнути фінансової катастрофи». Це буквально опис робочої рамки, у якій читання й аналіз ідуть без питань і швидко, а на незворотному стоїть явне підтвердження людини. Приємно бачити цю конструкцію описаною як індустріальний стандарт там, де ціна помилки - $440 млн. [proven - стаття прочитана (безкоштовна частина, далі пейволл); це матеріал за участю самої компанії, інтерв'ю з їхнім CTO і керівниками, тобто погляд зсередини без незалежної перевірки; глибші розділи за платною стіною, недоступні]

The Pragmatic Engineer: інженерія в Optiver

ТЕМА 10

Mojo дійшла до 1.0 - через три роки після анонсу. Мова, яку робили, щоб замінити C++ у ШІ

Тихий, але справжній реліз (332 бали): **Mojo 1.0**, «віха, до якої мова йшла з першого релізу у 2023». Modular пише прямо, чому саме зараз: швидкі зміни в мові

«ускладнювали спільноті підтримувати довгі проєкти», а 1.0 - це передусім обіцянка стабільності. Аргумент довіри від них же: «це вже мова, на яку ми щодня спираємось у продакшені», на ній стоїть їхня комерційна інфраструктура MAX і Modular Cloud.

Чому це важливо. Нуль дії, це маркер часу: ще одна спроба зробити «Python, але швидкий» дійшла до стабільного релізу. Для більшості автоматизацій це не критично, бо вузьке місце там не швидкість мови, а мережеві виклики й очікування чужих API. Але якщо колись зайде розмова про «переписати щось важке», тепер це 1.0. [proven - блог Modular прочитаний; бенчмарків тут немає, у пості їх і нема - це реліз про стабільність]

Modular 26.5: Mojo 1.0 · HN-тред, 332 бали

misc - коротко, що ще варте ока

• **Unsloth Desktop** - безкоштовний опенсорсний застосунок, щоб запускати і тренувати моделі локально (macOS/Windows/Linux, MLX і GGUF), і головне: до нього можна підключити Claude Code або Codex. Найпростіший з наявних способів поміряти локальну модель на своїй машині. Порадив @dessaigne з YC, анонс - @UnslothAI

• **@levelsio публічно скаржиться на Claude:** «Claude з кожним днем стає дедалі моралізаторськішим, регулярно просто відмовляє... готовий перескочити на Grok, якщо він дотягне для кодингу». Друга скарга такого роду за тиждень. Частину цього знімає домовленість про рамку наперед: що агент робить сам, а що тільки з підтвердженням

• **Копайлот під MITM-проксі** (169 балів): інженер пропустив трафік GitHub Copilot через проксі, щоб побачити, що саме той шле на сервер - контекст, пам'ять, харнес. Метод цікавіший за висновки: «сирці кажуть, що застосунок може робити; що він реально робить - видно тільки в рантаймі». Варто приміряти до будь-якого ШІ-тула, який ставиться на робочу машину

• **@paulg тролить власний твіт:** рік тому він написав про засновника, що пише 10 000 рядків коду на день, і його рознесли. Тепер: «Ті, кого це обурило, обуряться ще більше, якщо сказати, що сьогодні зустрів засновника, який іноді пише 50 000 рядків на день? Чи всі вже звикли?»

• **@garrytan на Startup School, одна фраза по темі:** «Цілі стартапи невдовзі будуть markdown-файлами. Володій своїми скілами, бо інакше твоя робота стане скіл-файлом». Інструкції для агентів справді живуть у звичайних markdown-файлах, і саме в них лежить те, як усе працює

• **«Compression is prediction»** (356 балів) - довгий (3 739 слів) розбір того, чому компресори і мовні моделі по суті розв'язують одну й ту саму задачу. Чисте задоволення для інженерного мозку, практичної користі нуль

• **@levelsio з ідеєю, яку можна забрати сьогодні:** експортувати всі виписки, особисті, бізнесові, брокерські, у CSV, «закинути в Claude Code і попросити зробити дашборд витрат»

• **Manus** уже в пункті 5, але варто повторити одним рядком, бо дедлайн: бекап до 22 серпня

• **Скрол по всіх 43 252 003 274 489 856 000 станах кубика Рубіка** (287 балів) - марна і прекрасна штука. Просто гарно
