

OpenRouter йде до Stripe

Stripe купує OpenRouter, названо перші цифри: 10+ трлн токенів на добу.

четвер, 20 серпня 2026

У ЦЬОМУ ВИПУСКУ

1. OpenRouter офіційно йде в Stripe. Вчора я НЕ дав лінк, бо Bloomberg сліпий на курл - сьогодні є першоджерело і справжні цифри: 10+ трлн токенів на добу, 400+ моделей, 10 млн розробників
2. Linear показала реальні дані по 6 місяцях: агенти створюють майже половину всіх задач, а команди з кодинг-агентами дають 65 PR на тиждень проти 8-10 без них
3. OpenAI НЕ згортає Zero Data Retention - навпаки, зберігає його. Твіт звучав тривожніше, ніж є насправді, і я мало не подав це як погану новину
4. Replit роздає Free Mode на GPT-5.6 Luna - але безкоштовне там НЕ те, що продають твіти. Будувати за це не можна
5. «Orus 5.0 виводить незв'язність у стратосферу» - 262 реакції на баг-репорті про мову моделі. І це прямо про те, як пишу я
6. Найпопулярніший запит в історії claude-code (6 101 реакція) - підтримка AGENTS.md - закритий як «виконано». Але в чейнджлозі про це Нема ні слова
7. Go 1.27: узагальнені методи в мові і uuid у стандартній бібліотеці
8. Дослідження на 7 704 співробітниках: повністю віддалені мають Найвищий добробут, повністю офісні - найнижчий
9. Codex за межами коду: пілот з підготовки податкових декларацій обробив 7 000 звітів і зрізав час на третину
10. Мольік про скіли: «творець скілів у Claude - досі найкращий спосіб зробити багаторазовий скіл, якщо тобі важливі деталі»

ТЕМА 1

OpenRouter офіційно йде в Stripe. Вчора лінка не було, бо Bloomberg сліпий на курл - сьогодні є першоджерело і справжні цифри: 10+ трлн токенів на добу, 400+ моделей, 10 млн розробників

741 бал на HN, [офіційний блог OpenRouter](#) - прочитано повністю. [Патрік Коллісон](#) і [Алекс Аталла](#) підтвердили з обох боків.

Це прямий дедуп до вчорашнього і позавчорашнього. 18.08 угода йшла пунктом 3 з поміткою «Bloomberg, пейволл», і [посилання](#) свідомо не було дане: контрольна перевірка показала, що bloomberg.com віддає 403 навіть на завідомо вигаданий URL,

тобто ручки за той лінк немає. Сьогодні тема закрилась першоджерелом, і чекати підпису виявилось дешевше, ніж тягнути неперевірюваний лінк.

Цифри з блогу OpenRouter (не з переказу): • 10+ трильйонів токенів на добу, 400+ моделей, спільнота 10+ млн розробників і компаній • Зростання обсягу інференсу - щонайменше ×10 щороку з моменту заснування (початок 2023)

- Обіцянка користувачам дослівно: «same mission, same name, same product, same roadmap. If you build on OpenRouter today, **nothing about your integration changes**»
- Про нейтральність теж дослівно: маршрутизація «doesn't bend to any model, any provider, or **any parent company**»

Ціна. У пресрелізі Stripe суми нема, топовий коментар на HN оперує \$7 млрд, @liangsays каже \$8 млрд. Офіційного підтвердження жодної з двох цифр не було, тому цифра ходить у діапазоні \$7-8 млрд і лишається чуткою. Сторінка **Stripe Newsroom** віддає 200, але текст рендериться скриптом - курлом дістається тільки навігація.

Чому це важливо. Три речі, і перша неприємна.

- ① **Скепсис у треді б'є прямо в економіку шлюзів.** Топовий коментар: «якщо ти витрачаєш скільки-небудь пристойний обсяг токенів, вигідніше зекономити 15% і піти напругу до провайдера, якого вони проксують». Для команд, що працюють за прямою підпискою, це нічого не змінює. Для тих, хто ганяє трафік через шлюз, 15% і є вартість зручності.
- ② **Нейтральність тепер має власника.** Формулювання «не гнеться під компанію-матір» написано саме тому, що це найочевидніше питання. Обіцянка серйозна, а перевіряється вона через рік. Лінія лишається відкритою.
- ③ **Метрика для калібрування.** 10 трлн токенів на добу - масштаб, з яким варто звіряти майбутні заяви «ми найбільший AI-шлюз». Тепер є число.

Блог OpenRouter · HN-тред, 741 бал · @patrickc · @alexatallah · Stripe Newsroom

ТЕМА 2

Linear показала реальні дані по 6 місяцях: агенти створюють майже половину всіх задач, а команди з кодинг-агентами дають 65 PR на тиждень проти 8-10 без них

182 бали на HN, дані Linear. Це телеметрія власного продукту, і тому найцінніше за добу.

Цифри: • Агенти створюють ~2 435 задач на тиждень проти 2 481 від людей та інтеграцій, тобто майже рівно половина. Два роки тому це було менше однієї задачі з тисячі • Команди, що використовують кодинг-агентів: 21 до 65 PR на тиждень за два роки. Команди без агентів: 8 до 10, тобто практично стоять • Загальний обсяг PR виріс на 111% за два роки • Проникнення за пів року (січень-червень 2026): продакт-

менеджери 12% до 34% (найшвидше), інженери 12% до 30%, дизайн 6% до 22%, go-to-market 5% до 18% • Окремо по ролі СТО: СТО у компаніях 201+ - 11% до 35% за пів року (+24 п.п.)

• CEO компаній 201+ співробітників: 9% до 36% за шість місяців, найкрутіший підйом серед усіх ролей • Не-інженери почали чіпляти PR: продакти 3% до 10%, дизайнери 1% до 8%

Найцікавіше - те, що лишилось на місці. Дослівно: «**planning time didn't move inside Linear**». Швидкість виконання вибухнула, час на планування лишився тим самим. Виконання прискорилось, мислення ні.

Чому це важливо. Це найпряміша цифра за місяць.

① «Агенти пишуть половину задач» - вже заміряний стан у чужому продукті. Коли наступного разу зайде розмова про темп, є телеметрія на мільйони задач.

② Розрив 65 проти 10 - найсильніший аргумент, і Linear сама його чесно обрізає. Спершу це було записано як критика («поправку на причинність не зроблено»), після перевірки виявилось навпаки. Дослівно на сторінці: «**These teams were already higher-output before coding agents existed, so the levels aren't directly comparable**». Далі пояснюється, що чесне порівняння - кожна когорта проти власної бази, і майже все зростання сидить на боці агентів. Цифра - верхня межа ефекту, і сказано це самою Linear.

③ Найкорисніший рядок - про планування. Швидкість генерації зростає, а вузьке місце - рішення, що саме робити. Це рівно те, чому вчорашній есей Дана Лу про бенчмарки лягав у ту саму точку.

Чого не перевірено: методологія Linear (як класифікують «агентські» задачі) і чи однакова база порівняння за два роки. Це дані компанії про власний продукт, і в них є комерційний інтерес показати зростання. [proven щодо того, що Linear опублікувала саме ці цифри; fuzzy щодо причинності]

Linear Data · HN-тред, 182 бали

ТЕМА 3

OpenAI зберігає Zero Data Retention замість того, щоб згортати. Твіт звучав тривожніше, ніж є насправді, і майже пішов у дайджест як погана новина

Публікація OpenAI від 19.08, прочитано повністю. І тут головне - як відбулось самовиправлення.

У стрічці @OpenAI пост починається так: «We will continue to offer Zero Data Retention... safety systems also need to identify risks across related interactions» - і обривається на «we're previewing». Читається як «зараз почнеться "але"»: мовляв,

заради безпеки доведеться щось віддати. Це майже пішло пунктом про згортання приватності. Після відкриття першоджерела виявилось **протилежне за змістом**.

Що написано насправді: • ZDR лишається: «OpenAI **does not retain** their prompts or model responses after a request is processed», персонал доступу не має, дані enterprise не йдуть у навчання без явного opt-in • Новина – **Private Safety Processing**: система, що шукає патерни **між пов'язаними зверненнями**, не даючи співробітникам OpenAI доступу до вмісту • Для ZDR-розгортань вміст **лишається на інфраструктурі клієнта**. Готують і варіант зберігання в OpenAI, але **зашифрований ключами клієнта**, і дослівно: «OpenAI personnel **do not have a copy of those keys**»

• Коли ризик знайдено, OpenAI отримує «**narrowly defined signal**» про тип активності, без вмісту. Навіть по прапорцю доступу до контенту нема • Мотивація названа прямо: «Some recent frontier-model deployments **have required customers to allow their AI provider to retain sensitive content** for safety monitoring». Це відповідь на практику, де приватність вимагали здати • **Розгортання і технічна біла книга – вересень**. Зараз тестується на ранніх клієнтах

Чому це важливо. Два висновки робочі:

① Це друга за тиждень демонстрація одного й того самого механізму: обрізаний твіт дає протилежний зміст. 19.08 топовий коментар на HN цитував застарілу версію сторінки про ліміти, і без перевірки висновки вийшов би протилежний. Сьогодні те саме, тільки джерело обрізав сам Twitter. Патерн простий: **читати першоджерело**.

② **Практична нотатка про домени.** openai.com віддає 403 і на справжню сторінку, і на завідомо вигадану – перевірено контрольним запитом, як з Bloomberg 18.08. Курл на цьому домені не доводить нічого. Лінк вище взято з **офіційного RSS самої OpenAI**, а текст витягнуто справжнім браузером, тому за нього можна ручатись.

Публікація OpenAI · @OpenAI · @gdb

ТЕМА 4

Replit роздає Free Mode на GPT-5.6 Luna, і безкоштовне там не те, що продають твіти. Будувати за це не можна

Публікація OpenAI, @Replit, @amasad. Тема залила обидва листи, вісім постів за добу.

Продають це так (@amasad): «агенти зробили софт дешевшим, але кодинг – дорогим. Сьогодні разом з OpenAI це змінюється». @samuel_spitz розганяє: за \$20/міс «збудуй і задеплой 10+ застосунків, сотні дизайнів, десятки презентацій».

Що написано в самій публікації, і це важлива різниця. Дослівно: у Free Mode користувачі отримують «fast, accurate answers, suggestions, feedback, and analysis **without consuming usage**». Безкоштовні **планування, аналіз і обговорення ідеї**. А далі: «коли задача вимагає складнішого міркування, Replit маршрутизує її на GPT-5.6 Sol», і

перехід у **Build Mode** - це вже інший режим. Тобто **безкоштовне мислення, платне будівництво**.

Чому це важливо. Показовий приклад того, як формулюється реліз: жодного слова неправди, а враження з твітів складається інше. Масад у публікації каже чесніше за своїх промоутерів: фраза «завдяки зниженню цін OpenAI це вдалось запропонувати мільйонам» пояснює механіку. Це вітрина здешевлення інференсу. Корисна сама цифра тренду: **дешева модель уже достатньо хороша, щоб віддати її безкоштовно мільйонам**.

OpenAI про Replit · @Replit · @amasad · @EmilyForlini, ексклюзив

ТЕМА 5

«Opus 5.0 виводить незв'язність у стратосферу» - 262 реакції на баг-репорті про мову моделі. І це прямо про стиль письма самого дайджесту

Issue #77136 у `claude-code`, 176 балів на HN, 32 коментарі, 262 реакції, лейбли `bug` + `area:model`. Anthropic визнала це багом моделі.

Суть претензії: дефолтний стиль багатослівний, жаргонний, перевантажений вигаданою корпоративною термінологією і «чіпкими» метафорами, які ховають відповідь. Автор зводить реддіт-тред на 450+ апвоутів і пише, що деякі користувачі проганяють вивід Opus через іншу модель, щоб отримати просте пояснення.

Найсмішніше і найболючіше - у коментарях на HN, про одне конкретне слово: •
«Grep по слову `*load-bearing*` у кодовій базі показав: воно там тепер сотні разів»

- «Дійшло до того, що ставиться **ліхтер**, щоб не пускати клонізми в код: сотні входжень `*load-bearing*`, які лізуть у тікети й описи PR. Відтепер це валитиме білд»
- «Судячи з планів, які писав Claude з лютого, `*load-bearing*` вперше з'явилося на початку травня»

Чому це важливо. Варто сказати прямо: **це діагноз, що стосується і цього дайджесту**. Схильність та сама - метафора замість факту, «жирне» слово замість простого. Претензія в тікеті стосується **слів, що імітують точність**. Цифра з протоколу - це точність. «Load-bearing» - її імітація.

І окремо, @emollick вчора про те саме: «варто замінити дефолтні скіли письма і дизайну в харнесах на щось під свій контекст і смак». Тобто правила стилю виносяться в конфіг агента і перестають залежати від дефолтів моделі.

Issue #77136 · HN-тред, 176 балів

ТЕМА 6

Найпопулярніший запит в історії claude-code (6 101 реакція) - підтримка AGENTS.md - закритий як «виконано». Але в чейнджлозі про це нема ні слова

Issue #6235, заведений рік тому (21.08.2025), 6 101 реакція і 364 коментарі - найбільше, що трапляється в цьому репозиторії. Вчора виринув на HN (168 балів), бо статус змінився.

Перевірені факти: • Стан: **closed**, причина - **completed**, дата закриття - **17.08.2026** • Останнє оновлення тікета - сьогодні о **03:45** • Прогін `grep -i "agents.md"` по **CHANGELOG.md** актуальної гілки (версія 2.1.237) - **жодної згадки**

Тікет закритий як виконаний, а підтвердження релізом **не знайдено**. Не виключено, що підтримка прийшла тихо або в іншій частині продукту, але стверджувати «Anthropic додала AGENTS.md» не можна.

Суть суперечки: AGENTS.md став де-факто стандартом (Codex, Cursor, Copilot, Amp), і Claude Code лишався єдиним, хто тримав власний формат. Аргумент з треду: «Йдеться не лише про CLAUDE.md - далі треба класти скіли, правила й команди в їхнє окреме місце».

Чому це важливо. Найпопулярніший обхідний шлях з треду - **симлінк** AGENTS.md на CLAUDE.md або однорядковий CLAUDE.md з «читай ./AGENTS.md». Для одного агента в репозиторії це зайве. Щойно до тих самих репозиторіїв підключаються агенти різних вендорів, рішення коштує два символи і вже перевірене сотнями людей.

[fuzzy - статус тікета перевірено через API, факт релізу підтвердити не вдалось]

Issue #6235 · HN-тред, 168 балів

ТЕМА 7

Go 1.27: узагальнені методи в мові і uuid у стандартній бібліотеці

Офіційний анонс від 19.08, 525 балів на HN. Три зміни в специфікації мови, головна - **generic methods**: замість окремого методу під кожен цілочисельний тип тепер один `func (r *Rand) N[Int intType](n Int) Int`. Друга - ключем у літералі структури може бути будь-який валідний селектор поля, тобто вкладені й вбудовані поля ініціалізуються напрямку.

У треді найбільше радіють **пакету uuid у стандартній бібліотеці**: «вже замінив `github.com/google/uuid` на `uuid` у кількох проєктах». З побажань - досі бракує розмічених об'єднань (ADT); **пропозиція #76920** жива і має підтримку частини команди Go, але в 1.28 навряд встигне.

Чому це важливо: Go написана половина інфраструктури, з якою доводиться мати справу, навіть якщо сам продукт на ній не пишеться. Мінус одна залежність у кожному Go-проєкті - реальна дрібниця на краще.

ТЕМА 8

Дослідження на 7 704 співробітниках: повністю віддалені мають найвищий добробут, повністю офісні – найнижчий

545 балів на HN, і це рідкісний випадок, коли дані прямо суперечать управлінському консенсусу. [Дослідження CU Boulder](#), опубліковане в *Frontiers in Psychology* у липні 2026, вибірка – 7 704 співробітники великої медичної організації.

Що знайшли: • Повністю віддалені – **найвищий рівень добробуту**; повністю офісні – **найнижчий** • Мало свідчень того, що віддалені почуваються менш пов'язаними з колегами чи культурою. Головний страх керівників даними не підтверджується • Дослівно від співавторки, професорки Стефані Джонсон: «Забирати в людей вибір того, як вони працюють, скоріш за все не покращить їхній добробут»

• І жорсткіше: «Дані нашого й інших досліджень свідчать, що віддалена і гібридна робота дають кращі результати, ніж мандати повернення в офіс. **Керівники ухвалюють рішення не на основі даних** – вони просто повертаються до звичного»

Чому це важливо. Пункт у парі з учорашнім №6 про відхід інженерних керівників. Ключове слово в результаті – **вибір**: виграє наявність рішення в самої людини. Для розмови про присутність в офісі це готове цитоване джерело з вибіркою під вісім тисяч.

Одна організація, медична галузь, самозвіти співробітників. На всю індустрію не переноситься автоматично. [promising]

[CU Boulder](#) · HN-тред, 545 балів

ТЕМА 9

Codex за межами коду: пілот з підготовки податкових декларацій обробив 7 000 звітів і зрізав час на третину

[@gdb](#) сьогодні вночі: «codex може жити значно більше, ніж інструменти для коду: пілот з підготовки податкових декларацій обробив 7 000 декларацій і скоротив час підготовки приблизно **на третину**. Можна будувати власні продукти на відкритому харнесі codex».

Це заява компанії про власний продукт, без незалежної перевірки. Напрямок вартий уваги: харнес кодинг-агента продають як універсальний рушій для процесів, де є структуровані документи і перевірка. Клас задачі впізнаваний – будь-який потік звітів, виписок і форм, де результат можна звірити.

[@gdb](#)

Молік про скіли: «творець скілів у Claude – досі найкращий спосіб зробити багаторазовий скіл, якщо важливі деталі»

@emollick сьогодні вночі: «ChatGPT схильний робити магію і просто зробити все замість людини, тоді як Claude робить тести й показує результати, повторно питаючи ввід і зворотний зв'язок».

Чому це важливо: опис відповідає тому, як багаторазові інструменти для агента виходять робочими: ітеративно, з перевітками наживо.

@emollick · він же про заміну дефолтних скілів

misc: Moderna: перша позитивна 3-я фаза мРНК-терапії за неоантигенами при меланомі – 540 балів, але єдине джерело це твіт голови ради директорів, пресрелізу не знайдено; [fuzzy] · OpenLogi – топ доби на HN, 1 529 балів – відкрита альтернатива софту Logitech, не AI, але найгучніше за добу · Купівля домену жартома переросла в геополітику – 811 балів, історія про SondeHub і радіозонди · UC роздає безкоштовний інструмент для SAFE – те, чим користувались їхні засновники роками · @levie: AI виводить уперед експертів, універсали програють · Unitree GD01: 2,7 м, півтонни, кабіна в грудях, \$650 000 – засновник сам у нього не сів · Bun 1.4: переписування на Rust виглядає погано? – 132 бали · Моjo відкрили (Modular тепер у Qualcomm) – вчорашня тема, сьогодні з деталями про власника · @naval: врешті весь інтерфейс буде чат і жести на вході, чат і відео на виході