AI-ДАЙДЖЕСТ

SpaceX уперше зловила баштою Starship з орбіти

SpaceX уперше зловила баштою Starship, а WSJ пише, що Nvidia фінансувала клієнтів, які купували її GPU.

неділя, 15 листопада 2026

У ЦЬОМУ ВИПУСКУ

1. SpaceX уперше зловила баштою Starship, що повернувся з орбіти
2. Nvidia: WSJ про штучний попит, продаж акцій директора і 13 місяців мовчання Хуанга
3. Project Zero: вразливості розкриватимуть через 90 днів, з 2028 року через 30
4. Lovable: у тисячах застосунків база даних відкрита без входу
5. SatBench: 63% з 173 бенчмарків уже «вирішені», незалежно лише 24%
6. Fable 5.2 проти Opus 5.5 у Claude Code: однаковий результат, у 2,8 раза дорожче
7. Google /teamwork: агенти формалізували гіпотези Вейля в 1,8 млн рядків Lean
8. Claude Code Projects: координатор роздає роботу паралельним тредам
9. Італія оштрафувала Meta на €805 млн за доступ до фото на Android
10. Зупинка вильотів у США: L3Harris підтвердила, що переводить зв'язок FAA з міді на оптику

ТЕМА 1

SpaceX уперше зловила баштою Starship, що повернувся з орбіти

ДЖЕРЕЛА SpaceX, 14.11 Spacex · Ars Technica, 15.11 Ars Technica · @zebulgar, 14.11 @zebulgar · HN, 15.11 Hacker News

13.11 Starship стартував з бази Starbase у Техасі в 15-й політ. Корабель вийшов на орбіту і випустив 26 супутників Starlink V3. Понад 20 годин він був у космосі, потім увімкнув один двигун Raptor і зійшов з орбіти.

Через 22,5 години після старту корабель повернувся до тієї самої стартової вежі. Три двигуни розвернули його вертикально, і механічні руки вежі (Mechazilla) зловили корабель. Це перше таке захоплення верхнього ступеня.

Прискорювач Super Heavy цього разу не повернувся. Маневр повернення закінчився раніше через малий запас пального, причину SpaceX ще шукає. Автоматика відправила прискорювач на посадку в Мексиканську затоку, після посадкового маневру спрацювала система самоліквідації.

Ars Technica дає масштаб. Space Shuttle за 30 років зробив 135 польотів, а вартість виводу була близько $25 000 за фунт вантажу. SpaceX зараз будує один Starship кожні три тижні, а завод розрахований на один корабель кожні три дні. Головна відкрита проблема - теплозахист, який треба готувати до наступного польоту швидко і дешево.

Чому це важливо: до цього повторно використовувались лише перші ступені ракет. Тепер SpaceX повернула на вежу обидві частини Starship, хоча й у різних польотах. Повністю багаторазова ракета знижує ціну виводу на орбіту в рази. Від цієї ціни залежить, які проєкти в космосі взагалі мають економічний сенс.

ТЕМА 2

Nvidia: WSJ про штучний попит, продаж акцій директора і 13 місяців мовчання Хуанга

ДЖЕРЕЛА TechSpot, 14.11 Techspot · TechPowerUp, 14.11 Techpowerup · The New York Times, 14.11 NYT · Yahoo Finance, 14.11 Yahoo · @garrytan, 14.11 @garrytan · HN, 14.11 Hacker News

Вчора ми писали про кампанію за повернення Хуанга. За добу з'явились три нові лінії.

Перша - бізнес. TechSpot і TechPowerUp переказують профіль тимчасової CEO Колетт Кресс у WSJ. За цим профілем, Nvidia інвестувала в стартапи і малих хмарних провайдерів, які потім купували її GPU. Мета: продати надлишок пластин, замовлених у TSMC, і тримати ціни високими. У серпні Кресс пропонувала скоротити такі угоди, зокрема угоду на $250 млрд з OpenAI. За словами джерела WSJ, Хуанг тоді назвав її «deer in the headlights». Самого тексту WSJ ми не читали, тільки переказ.

Друга - акції. Директор Nvidia Марк Стівенс продав у серпні і вересні близько 4,2 млн акцій приблизно на $946 млн, пише Yahoo Finance з посиланням на звіти SEC. Заздалегідь оформленого плану продажу (10b5-1) у нього не було. Само по собі це не доказ порушення, але питання до ради додасться.

Третя - хронологія. NYT опублікувала розслідування на основі розмов з понад 100 людьми. Газета вперше спитала Nvidia про Еліс Хуанг у вересні 2025 року. 13 місяців Хуанг заперечував близькі стосунки і зізнався юристу компанії лише минулого вікенду. Після першої публікації NYT Nvidia перестала показувати Сінгапур окремо у фінансовій звітності. До того на Сінгапур припадало $10 млрд, тобто 22% квартальних продажів.

Кампанія Stand with Jensen тим часом триває. Гаррі Тан з Y Combinator підписав лист і пише про «due process».

Чому це важливо: сюжет вийшов за межі особистого скандалу. Якщо модель «фінансуємо клієнта, щоб він купив наші чипи» підтвердиться документами, питання постане до всієї AI-економіки, бо частина попиту на GPU могла бути штучною. Звіт Nvidia за квартал 17.11 буде першою точкою, де компанія мусить відповідати на ці питання публічно.

ТЕМА 3

Project Zero: вразливості розкриватимуть через 90 днів, з 2028 року через 30

ДЖЕРЕЛА Google Project Zero, 14.11 Projectzero · HN, 14.11 Hacker News

З 2021 року команда Google Project Zero давала виробнику 90 днів на виправлення плюс 30 днів на розгортання патча. Тепер правило стає простим: деталі вразливості виходять рівно через 90 днів після звіту, незалежно від статусу патча. 02.02.2027 термін скоротиться до 60 днів, 03.01.2028 - до 30 днів. Пільгові періоди й інші особливі випадки команда прибирає.

Причину Project Zero пише прямо: LLM змінили швидкість атак. Зловмисник з моделлю знаходить ту саму помилку швидше, а з опублікованого патча відновлює деталі вразливості майже миттєво. Тому команда вважає час «від патча до експлойта» практично нульовим. Окрема цифра: там, де команда раніше знаходила кілька вразливостей на рік, тепер знаходить кілька на тиждень. Переговори з кожним виробником про статус кожної помилки перестали масштабуватись.

Виключень лишилось три: вразливість уже експлуатують, її вже опублікував хтось інший, або виробник сам просить розкрити деталі.

Чому це важливо: Project Zero задає неформальний стандарт для всієї галузі. Якщо найвідоміша команда пошуку вразливостей офіційно визнає, що 90 днів забагато, тиск на виробників зросте. Для команд, які роблять ПЗ, висновок практичний: час від звіту до виправлення стає метрикою, яку видно ззовні. Варто заздалегідь мати процес на випадок, коли патча ще нема, а деталі вже публічні.

ТЕМА 4

Lovable: у тисячах застосунків база даних відкрита без входу

ДЖЕРЕЛА Crissy Field, 14.11 Crissyfield · HN, 14.11 Hacker News

Німецька компанія Crissy Field разом з Handelsblatt перевірила застосунки, які люди зібрали на платформі Lovable.

Lovable робить вебзастосунок з опису звичайною мовою. Дослідники знайшли в архіві Common Crawl 159 041 проєкт Lovable за грудень 2025 - вересень 2026 року. 34 545 з них використовують базу даних Supabase.

24 664 сайти віддавали записи з бази одразу при завантаженні, без входу. Додатковими запитами лише на читання дослідники отримали більше. Таблиці з адресами пошти знайшлись на 7 403 сайтах, тобто на 30%. Автори наголошують, що це не підтверджена частка витоків: частина адрес публічна або тестова.

Серйозні випадки такі. Німецький лайфстайл-форум віддавав понад 80 000 профілів і понад 200 000 чат-повідомлень.

Сервісна компанія віддавала понад 1 000 записів з даними паспортів. Ще були квитки й платіжні дані організаторів подій, договори, рахунки і 280 000 листів одного провайдера клієнтам. Чи хтось уже скористався цими даними, дослідники не встановили.

Чому це важливо: помилка в правах доступу буває в будь-якому коді. Але на платформі, де застосунки будуються за одним шаблоном, одна помилка повторюється в тисячах місць. Зловмиснику вистачає одного способу пошуку. Lovable в умовах використання перекладає відповідальність за дані на клієнта. Тому правило просте: перед тим як застосунок отримає реальні дані клієнтів, права доступу має перевірити людина, яка в цьому розбирається.

ТЕМА 5

SatBench: 63% з 173 бенчмарків уже «вирішені», незалежно лише 24%

ДЖЕРЕЛА @emollick, 14.11 @emollick · SatBench на GitHub GitHub

Ітан Моллік з Wharton зібрав SatBench на даних Benchmark Almanac. Він взяв 173 бенчмарки, які лабораторії найчастіше згадують з 2023 року. Бенчмарк вважається вирішеним, коли найкращий результат моделі досягає 85%.

Станом на 14.11 вирішені 109 з 173, тобто 63%. До початку 2025 року частка трималась біля 30%, за останні 12 місяців вона виросла на 20 пунктів.

Бенчмарки падають щороку швидше. За перший рік після виходу були вирішені 17% бенчмарків 2023 року, 30% - 2024 року і 43% - 2025 року. Медіанний час до 85% скоротився приблизно з 20 до 14 місяців. Бенчмарки 2023 року (MMLU, HumanEval, GSM8K та інші) у 2026 році згадали лише 2 з 134 анонсів моделей.

Найважливіше застереження автор пише сам. Більшість «вирішень» спираються на заяви самих лабораторій. Якщо рахувати лише незалежні заміри, SatBench стоїть на 24%. Три з чотирьох перших переходів через 85% взяті з анонсу лабораторії, і незалежно підтвердили лише 10 з 75.

Чому це важливо: цифра в анонсі моделі дедалі менше каже про реальну різницю між моделями. Бенчмарк швидко насичується, а результат часто отримано в налаштуваннях самої лабораторії. Для вибору моделі під свою задачу корисніше мати власний невеликий набір тестів і повторювати його на кожній новій версії. Наступний пункт - приклад.

ТЕМА 6

Fable 5.2 проти Opus 5.5 у Claude Code: однаковий результат, у 2,8 раза дорожче

ДЖЕРЕЛА @garrytan, 14.11 @garrytan · звіт g-harness-evals, 12.11 GitHub

Гаррі Тан, президент Y Combinator, веде відкритий репозиторій тестів для агентних середовищ. Свіжий прогін порівнює Fable 5.2, Fable 5.1 і Opus 5.5 у Claude Code 2.1.322 на 52 задачах з програмування, по дві спроби на задачу. Гіпотези і правило рішення автори записали до запуску.

Результат: Fable 5.2 вирішила 82,7% задач, Opus 5.5 - 84,6%. Різниця -1,9 пункту з інтервалом від -9,6 до +5,8, тобто статистично її нема. Ціна за вирішену задачу різна: $1,83 у Fable 5.2 проти $0,65 в Opus 5.5, у 2,8 раза дорожче. Висновок звіту: Opus 5.5 лишається моделлю за замовчуванням для коду.

Звіт чесно показує межі. Попередній прогін на Claude Code 2.1.312 не знав Fable 5.2, тому дав інші цифри. Цього разу розрив зменшився здебільшого тому, що Opus показав нижчий результат. Opus уже вирішує більшість задач, тож для чіткої різниці потрібні складніші задачі. Лише 52 задачі дають мінімальну помітну різницю 11,5 пункту.

Чому це важливо: нова флагманська модель не завжди краща для щоденної роботи. На типових задачах різниця в якості в межах шуму, а різниця в ціні майже втричі. Цей звіт також показує, як виглядає чесне порівняння:

попередньо записані правила, інтервали похибки і окремий рядок про те, що тест не здатен розрізнити.

ТЕМА 7

Google /teamwork: агенти формалізували гіпотези Вейля в 1,8 млн рядків Lean

ДЖЕРЕЛА Google Research, 13.11 Research · HN, 14.11 Hacker News

Google описала результати /teamwork, багатоагентного режиму свого середовища Antigravity на моделях Gemini 4. Ідея у двох принципах. Перший: багато паралельних спроб плюс окремі перевіряльники (доказник Lean 4, незмінні набори тестів, симулятори). Другий: кожен агент бачить лише потрібні йому файли, а критики перевіряють результат з чистого контексту.

Найбільший результат - автономна машинно перевірена формалізація гіпотез Вейля, які П'єр Делінь довів у 1974 році. Це близько 1,8 млн рядків Lean і 28 800 викликів субагентів. Ще в списку: формалізація гіпотези Біркгофа (відкрита з 1927 року) і доведення гіпотези Куртада-Кумара з теорії інформації. В інженерії агенти за два тижні переписали решту мікроядра Zircon на безпечний Rust, близько 800 000 рядків, що проходять тести ядра.

Окремо Google дала інструмент рецензії Deep Review авторам 578 статей до конференції STOC 2027. За опитуванням 70 авторів, понад 30% кажуть, що інструмент знайшов суттєві помилки в доведеннях. Усі цифри - від самої Google.

Чому це важливо: спільна риса всіх результатів - формальна перевірка. Модель пропонує, а Lean, тести або симулятор вирішують, чи це правда. Це та сама схема, що й

у історії з нотатками Геделя вчора, тільки в масштабі мільйонів рядків. Там, де перевіряльника нема, такі цифри поки недосяжні.

ТЕМА 8

Claude Code Projects: координатор роздає роботу паралельним тредам

ДЖЕРЕЛА Anthropic, 13.11 Anthropic · Anthropic, 13.11 Anthropic · HN, 14.11 Hacker News

Anthropic переробила проєкти в Claude Code. Користувач ставить ціль і підключає репозиторій. Координатор ділить роботу на треди, і кожен тред - окрема хмарна сесія Claude Code на своїй гілці. Треди відкривають PR і запускають тести. Конфлікти між ними вирішуються як звичайні конфлікти злиття. Проєкт має спільну пам'ять і бібліотеку файлів.

Бета доступна частині підписників Pro і Max, ширше розгортання обіцяють протягом тижня. Anthropic попереджає, що кілька тредів одночасно швидше вичерпують ліміти.

Того ж дня вийшла бета Claude Meetings для планів Team і Enterprise. Claude приєднується до дзвінка в Google Meet як видимий учасник, веде нотатки і створює документи під час розмови. Аудіо він не зберігає, лише транскрипт. Якщо в дзвінку є людина ззовні організації, Claude за замовчуванням не приєднується.

Чому це важливо: Google /teamwork і Claude Code Projects за один тиждень прийшли до однієї схеми: координатор, паралельні виконавці, окремий контекст для кожного. Схема переходить з дослідних звітів у продукти. Обмеження теж однакове: паралельна робота множить витрати, тому варто дивитись на ціну за результат.

ТЕМА 9

Італія оштрафувала Meta на €805 млн за доступ до фото на Android

ДЖЕРЕЛА BBC, 14.11 BBC · HN, 14.11 Hacker News

Італійський регулятор захисту даних оштрафував Meta на €805 млн ($887 млн). Функція Facebook пропонувала фото і відео з галереї телефону як ідеї для постів. Доступ до галереї застосунок отримував одним загальним дозволом. За висновком регулятора, близько 20 млн користувачів не знали, що цей дозвіл дає Facebook аналізувати їхні фото.

Meta має 60 днів, щоб вимкнути функцію за замовчуванням і просити окрему згоду. Компанія подасть апеляцію.

Чому це важливо: регулятор оцінював мету, для якої Meta використала дозвіл. Один загальний дозвіл на галерею регулятор не вважає згодою на AI-аналіз фото. Це правило зачепить будь-який застосунок, що додає AI-функції поверх уже наданих дозволів.

ТЕМА 10

Зупинка вильотів у США: L3Harris підтвердила, що переводить зв'язок FAA з міді на оптику

ДЖЕРЕЛА L3Harris, 14.11 L3harris · @garrytan, 14.11 @garrytan · HN, 15.11 Hacker News

Вчора ми писали, що FAA 13.11 зупинила всі вильоти в США через «communication issue». Тепер заяву зробив підрядник L3Harris, який модернізує зв'язок для FAA. Компанія пише, що разом з телеком-партнерами перевела понад 70% запланованого зв'язку з мідних дротів на оптоволокно. L3Harris допомагає FAA в розслідуванні. Причину зупинки компанія не називає, і ніхто офіційно не пов'язує її з переходом.

Чому це важливо: заява не каже, що винен перехід. Але вона показує, що в момент збою система була наполовину старою і наполовину новою. Великі міграції часто найкрихкіші саме в середині, коли працюють обидві системи одночасно.

misc

- **Агенти читають Internet Archive по рядку:** Ітан Моллік попросив Fable 5.2 відтворити курс поезії В. Г. Одена 1941 року. Агент знайшов потрібну книжку в Internet Archive і прочитав її рядок за рядком через повторні запити повнотекстового пошуку. Моллік пише, що трафік від агентів буде великим, і пожертвував архіву.
  @emollick
- **Revolut:** компанія опублікувала розбір п'ятигодинного глобального збою карток. За переказом Гергея Ороса, моніторингу падіння вузлів бази не було, а 5 вузлів у 3 регіонах не давали стійкості при втраті одного.
  @levelsio
- **Unsloth:** оновила стиснуті версії Qwen3.8-27B до Dynamic V3.1. Компанія каже про +5-10% точності при тому самому розмірі. 4-бітна модель працює на 16-19 ГБ відеопам'яті або на Mac з 24 ГБ.
  Unsloth · HN Hacker News
- **Webloc:** розслідування Bloomberg, на яке посилається Citizen Lab, знайшло нових клієнтів системи стеження через рекламні дані Webloc. Серед них Британія, Нідерланди, Німеччина, Мексика, Сінгапур, Саудівська Аравія, ОАЕ і Ізраїль.
  Citizenlab
- **HubSpot:** 12.11 співзасновник Браян Галліган повернувся на посаду CEO, компанія називає це «refounding».
  Дармеш Шах щодня пише нотатки про процес. @dharmesh
- **@amasad про збої тижня:** «Anything that can go wrong will go wrong». Мільярди взаємозалежних процесів роблять рідкісні збої щоденними. @amasad
- **Monkey Island:** сайт показує поруч усі 13 версій гри.

AI-дайджест · неділя, 15 листопада 2026