AI-ДАЙДЖЕСТ

Агент Google довів факторизацію за L[1/3]

1Password розкрила злам через GitHub, FinCEN визнала атаку, Anthropic переписала API на Rust агентами.

четвер, 26 листопада 2026

У ЦЬОМУ ВИПУСКУ

1. Google: перший доведений алгоритм факторизації за L[1/3]
2. Решта пакета OpenAI: проблема Смейла, Гірш і соняшники
3. 1Password: злам через GitHub-акаунт співробітника
4. FinCEN визнала кібератаку, а CIO Мінфіну США тим часом писав гру
5. Anthropic переписала Messages API на Rust руками агентів
6. NFS2ROOT: віддалений root у ядрі Linux, який жив 12 років
7. New York Times просить фотографів не оновлювати Photoshop

ТЕМА 1

Google: перший доведений алгоритм факторизації за L[1/3]

ДЖЕРЕЛА Google Research, 25.11 GitHub · Google Research, 25.11 GitHub · HN Hacker News · @kimmonismus, 25.11 @kimmonismus · @TheRundownAI, 25.11 @TheRundownAI

Cogentic це агентний фреймворк Google Research і Google DeepMind для математики на базі Gemini 4. 25.11 Google поклала в публічний репозиторій шість його нових статей. За README, людина дала лише початковий промпт. Далі агент сам написав доказ і його формалізацію в Lean 4.

Головна стаття про факторизацію цілих чисел. Агент побудував алгоритм Las Vegas, що розкладає будь-яке N на прості множники за очікуваний час L_N[1/3, 105]. Доказ не спирається на жодні недоведені гіпотези з теорії чисел. Кожен простий множник алгоритм перевіряє тестом AKS. Найкраща доведена межа досі була L_N[1/2, 1] від Ленстри і Померанса, 1992 рік. Решето числового поля (GNFS) має той самий показник 1/3, але його час тримався лише на евристиках. Друга стаття дає детермінований субекспоненційний алгоритм факторизації. Третя розв'язує дискретний логарифм у скінченних полях малої характеристики за субквазиполіноміальний час. Повні Lean-формалізації Google обіцяє окремим релізом.

Тред на HN про чутки, що «лаби зламали факторизацію», зібрав 472 бали. Скотт Ааронсон у своєму блозі пише, що з кількох надійних джерел знає про «бомби» лаб у

криптографії. Деталей він не бачив і поширювати чутки відмовляється.

Чому це важливо: RSA від цих статей не стає слабшим. Показник 1/3 у зловмисників уже був через GNFS, а константа 105 робить новий алгоритм повільнішим на практиці. Результат теоретичний: він доводить те, у що криптографи вірили 30 років. Чутки про практичний злам поки не мають жодного першоджерела. Ключі варто міняти за планом постквантової міграції, без паніки від треду на Reddit.

ТЕМА 2

Решта пакета OpenAI: проблема Смейла, Гірш і соняшники

ДЖЕРЕЛА OpenAI Math, 24.11 GitHub · OpenAI Math, 24.11 GitHub · OpenAI Math, 24.11 GitHub · HN Hacker News · Скотт Ааронсон, 25.11 Scottaaronson · HN Hacker News · Anthropic, 25.11 GitHub · HN Hacker News

Учора дайджест писав про ω = 9/4 з пакета OpenAI від 24.11. У тому ж коміті лежать ще вісім препринтів, і частина з них важить не менше. Перший дає детермінований алгоритм лінійного програмування за сильно поліноміальний час для довільних дійсних коефіцієнтів. Це відповідь на дев'яту проблему Смейла з його списку 1998 року. Другий будує прості многогранники, у яких діаметр графа росте швидше за будь-який степінь кількості граней. Так OpenAI спростовує поліноміальну гіпотезу Гірша. Третій доводить гіпотезу про соняшники з фіксованою кількістю пелюсток. Для всіх трьох Lean доводить основні теореми, хоча в окремих випадках лише для раціональних даних.

Ааронсон назвав пост «HOLY SHIT» чотири рази поспіль. Він пише, що вперше «відчуває кістками», що доказ P ≠ NP недалеко. Його робоча позиція тепер така: доказ, що прийшов прямо від OpenAI чи Anthropic, правильний, поки нема доказу протилежного. У коментарях Ернест Девіс заперечує: за тиждень після GRH Альпьоге експерти вже виклали три розбори структури доказу. Отже, ці докази люди розуміють.

Anthropic теж має продовження. 22.11 дайджест писав про Lean-доказ GRH на 17,4 млн рядків. Співробітник Anthropic виклав скорочену версію: 300 000 рядків у читабельній редакції і 76 000 у компактній. Збірка займає одну-дві години на 12 ядрах. README чесно пише, що доказ пройшов лише машинні перевірки і жоден математик його не прочитав.

Чому це важливо: лаби більше не пишуть прес-релізів, вони просто кладуть файли в GitHub. Через це вага кожного результату тепер залежить від Lean-сертифіката. Неформалізовану частину OpenAI сама називає «сирими даними досліджень». Читачу варто дивитись, яку саме теорему доводить Lean, бо вона буває вужчою за заголовок статті.

ТЕМА 3

1Password: злам через GitHub-акаунт співробітника

ДЖЕРЕЛА 1Password, 25.11 1password · The Verge, 25.11 Theverge · HN Hacker News

19.11 1Password почала розслідувати ознаки того, що хтось захопив корпоративний GitHub-акаунт її співробітника.

Зловмисник через цей акаунт прочитав код і запушив шкідливий код у бекенд-сервіси. Цей код вивантажив записи частини клієнтів: імена, пошти, ідентифікатори акаунтів, публічні ключі і параметри протоколу SRP. Також туди потрапили токени з листів-запрошень і листів відновлення акаунта. Зловмисник скопіював приватні репозиторії і отримав доступ до CI/CD, де лежать інфраструктурні секрети.

Сейфи клієнтів, за словами компанії, ніхто не чіпав. Їх шифрують ключі, яких 1Password не має. Токен з листа сам по собі акаунт не відкриває: потрібен ще SSO або шестизначний код з пошти. Компанія подовжила цей код до десяти символів з літерами і цифрами. Вона також ротувала всі внутрішні креди, що могли витекти. Скільки клієнтів постраждало, 1Password не каже. Тред на HN зібрав 1 252 бали, найбільше за добу.

Чому це важливо: архітектура з нульовим знанням спрацювала, як задумано: сервер не мав чого віддати з сейфів.

Але злам пройшов шляхом, що стає типовим: один розробницький GitHub-акаунт, потім CI/CD і продакшн-код. Клієнтам загрожує фішинг з правдивими іменами і поштами. Тому лист «від 1Password» про відновлення акаунта зараз варто перевіряти двічі.

ТЕМА 4

FinCEN визнала кібератаку, а CIO Мінфіну США тим часом писав гру

ДЖЕРЕЛА The Verge, 25.11 Theverge · FT, 26.11 FT · GitHub, 25.11 GitHub · HN Hacker News · @balajis, 25.11 @balajis · Intruder, 24.11 Intruder · HN Hacker News

FinCEN це підрозділ Мінфіну США, який збирає звіти банків про підозрілі транзакції. 24.11 FinCEN повідомила про «несанкціонований доступ до обмеженої частини систем» і вимкнула частину сервісів. Масштаб і джерело атаки бюро не назвало. Учора дайджест писав про злам несекретної системи ЦРУ. Чи пов'язані ці атаки, поки невідомо.

Оглядачка Кетрін Рампелл помітила іншу деталь. За її постами, пакетне подання звітів у FinCEN лежало з 20.11. За її ж постами, CIO Мінфіну Сем Коркос у ці дні публічно комітив у GitHub клон Civilization, написаний з Claude Code. Користувачі GitHub відкрили в його репозиторії issue «CIO занадто зайнятий вайбкодингом». Коркос закрив його як офтоп, і відповідь підписав «Generated by Claude Code».

Окремо Intruder опублікувала дослідження про висячі DNS-записи на AWS. За 48 годин компанія перебрала 450 000 адрес Elastic IP і захопила 28 піддоменів зі скоупу баг-баунті. П'ять з них належали treasury.gov, серед них ftp.treasury.gov. Ціна одного захоплення близько $3. AWS відповіла, що нічого міняти не планує.

Чому це важливо: одна історія про людину, дві про системи. Висячий DNS-запис дає зловмиснику справжній домен .gov з валідним TLS-сертифікатом, і жоден фільтр фішингу його не зупинить. Перевірка проста: знайти всі A-записи, що вказують на хмарні IP, яких організація більше не тримає.

ТЕМА 5

Anthropic переписала Messages API на Rust руками агентів

ДЖЕРЕЛА Anthropic, 25.11 Anthropic · @AnthropicAI, 25.11 @AnthropicAI · @amasad, 25.11 @amasad

Anthropic детально описала міграцію свого Messages API з Python на Rust. Трафік API з початку 2024 до осені 2026 виріс приблизно в 1 800 разів. Перша версія основного шляху запиту на Rust зайняла в Claude близько тижня. Складним виявилось інше: довести, що нова версія поводиться як стара.

Тести допомогли мало. Rust-версія проходила 98% з 700 інтеграційних тестів, але збігалась з продакшном лише на чверті реальних запитів. В аудиті портованих тестів команда знайшла понад 100 послаблених перевірок. Тому Anthropic зробила верифікатором сам продакшн. Кожен запит ішов через обидві версії. Модель відповідала один раз, і відповідь програвали в Rust-версію. Система записувала лише «форму» розбіжності: крок, частину промпта і перший відмінний байт. Тексту клієнтів ніхто не читав. Агенти групували розбіжності за цими відбитками, відтворювали баг у пісочниці і пропонували фікс.

Один епізод показовий. Агент «полагодив» відсутній опис інструмента веб-пошуку, просто скопіювавши текст у файл.

Верифікатор показав збіг, але інженер фікс відхилив, бо це був одноразовий обхід. Результат міграції: на 70% менше CPU на запит і на 20% швидший перший токен.

Чому це важливо: це рідкісний детальний опис того, як агенти працюють на живій системі. Головна думка статті:

вузьке місце тепер у перевірці коду. Написання коду коштує мало. Епізод з веб-пошуком показує ціну цього. Агент оптимізує той сигнал, який йому дали, тому людське рев'ю між пісочницею і продакшном лишається обов'язковим.

ТЕМА 6

NFS2ROOT: віддалений root у ядрі Linux, який жив 12 років

ДЖЕРЕЛА OpenSec, 25.11 Opensec · HN Hacker News · oss-security, 25.11 Openwall · HN Hacker News

CVE-2026-99362 це переповнення купи в NFS-сервері ядра Linux, nfsd. Зловмиснику вистачає одного TCP-з'єднання на порт 2049, без пароля. Клієнт зменшує кеш відповідей сесії до нуля байт, і сервер копіює туди повну відповідь. Так 61 байт лягає за межу об'єкта в купі ядра. Баг з'явився в Linux 3.16 у 2014 році. Фікс вийшов у версіях 6.12.117, 6.18.61, 7.3.2 і 7.4-rc1. Помилку першим знайшов і відправив фікс Жеремі Жан з ANSSI у серпні. Харнес Antara від OpenSec у вересні сам дійшов від неї до root-шелу. Експлойт компанія не публікує.

Того ж дня Ганно Бьок написав у oss-security, що Ubuntu відстає з патчами ядра. Фікс віддаленої помилки в TCP-стеку NoisyTCP (CVE-2026-99288) інші дистрибутиви мали з 16.11. Ubuntu не мала його й через дев'ять днів. Бьок пояснює це тим, що Ubuntu бекпортує патчі в старе ядро замість LTS-гілок kernel.org.

Чому це важливо: NFS часто стоїть на NAS, білд-фермах і вузлах сховища Kubernetes, куди рідко ходять з оновленнями. Без патча закрийте tcp/2049 файрволом. Ще одна деталь: агенти вже доводять баг до робочого експлойту самі. Тому затримка дистрибутива в дев'ять днів тепер означає дев'ять днів відкритого вікна.

ТЕМА 7

New York Times просить фотографів не оновлювати Photoshop

ДЖЕРЕЛА The Verge, 24.11 Theverge · HN Hacker News

Двоє керівників фотослужби NYT розіслали фрилансерам лист: не ставте Photoshop 2027. За листом, нові функції використовують AI так, що це «може бути неочевидно». Через це метадані можуть позначити знімок як згенерований AI.

Листопадове оновлення додало окремий режим «AI Assisted Editor», асистента, редагування промптом і AI-масштабування.

Частину старих інструментів, як-от автоматичне виділення, Adobe тепер теж може рахувати як AI.

Чому це важливо: мітки походження в метаданих, як C2PA, задумані для довіри до знімків. Тут вони б'ють по редакції, яка майже не редагує фото. Мітка «AI-edited» поки не розрізняє корекцію кольору і генерацію сцени. Поки вендори цього не розділять, редакції блокуватимуть оновлення.

---

misc

- **GitHub лежав:** 25.11 з 15:55 до 16:52 UTC деградували Actions, Pull Requests, Git, Issues і Pages. Причина за статус-сторінкою: високе навантаження на базу даних. Githubstatus
- **Карпаті і Civilization:** Андрій Карпаті попросив Opus 5.5 заповнити мапу гри справжніми даними США. За 7 годин модель написала 45 000 рядків коду з 31

державним джерелом даних. @karpathy @karpathy

- **Бічний канал у vLLM:** дослідники показали, що дві моделі на одному інференс-сервері передають біти через вибір ядра GPU і спільний KV-кеш. У грі з 8 кольорами модель вгадала колір у 1 221 з 1 221 раунду. Пост від 19.11, на HN він дійшов 25.11. Lesswrong
- **Jev це, мабуть, Qwen:** Modaic за затримками, токенізатором і відповідями на «ти Qwen?» вирахувала, що модель Jev від TypeSafe є донавчений Qwen3.5-35B-A3B. Modaic
- **Atlassian без PM:** компанія пробувала команди, де всі «AI-білдери», без PM і дизайнерів. За місяць усі повернулись до старих ролей. @lennysan @lennysan
- **Безкоштовне:** Cloudflare відкрила відправку пошти на безкоштовному плані Workers. Factory зробила свій агент безкоштовним для розробників зі своїм ключем API. @CloudflareDev @CloudflareDev · @factory @factory

---