

Super Intelligence Force Трампа очолив Клейтон

Strata запускає 125-мільярдну Qwen на ігровому ПК, суд скасував вирок через AI-відео, протест проти дата-центру Claude.

понеділок, 5 жовтня 2026

У ЦЬОМУ ВИПУСКУ

1. «Super Intelligence Force»: Білий дім зібрав групу з AI, і вона знову поступається компаніям
2. Strata: 125-мільярдна Qwen на ігровому ПК з картою на 12 ГБ
3. Аризона: суд скасував вирок, на який вплинуло AI-відео з загиблим
4. Дата-центр для Claude у Квінсленді: 2,16 ГВт і 21 500 підписів проти
5. RemoveMacAI: як вимкнути Apple Intelligence у macOS 27 повністю
6. AI і робота: Леві рахує нові професії, Моллік сперечається з Кубаном

ТЕМА 1

«Super Intelligence Force»: Білий дім зібрав групу з AI, і вона знову поступається компаніям

ДЖЕРЕЛА BBC, 05.10 **BBC** · The Guardian, 04.10 **Guardian** · Semafor, 04.10 **Semafor**

У неділю Дональд Трамп оголосив, що директор національної розвідки Джей Клейтон стає новим відповідальним за AI у Білому домі і очолює групу, яку президент назвав «Super Intelligence Force».

За даними BBC, до неї входять голова Федеральної торгової комісії Ендрю Фергюсон, заступник міністра оборони з досліджень Еміль Майкл і керівник кадрового відомства Скотт Купор. Semafor додає віцепрезидента, міністрів оборони і фінансів і Девіда Сакса, який був відповідальним за AI до березня. Група звітує напряду президенту і главі апарату Сьєзі Вайлз. Завдання з поста Трампа: координувати роботу уряду зі споживачами, громадськими і релігійними організаціями, операторами критичної інфраструктури і «компаніями суперінтелекту».

Учора тут писали, що WSJ дав групі 120 днів на звіт, а призначення Клейтона було лише чуткою. Тепер воно офіційне, і Semafor уточнює зміст: звіт має оцінити ризики й можливості AI і визначити відповідальність федерального уряду. Сам Клейтон сказав WSJ, що очікує «механізмів, вироблених у діалозі» з компаніями, але, за даними Semafor з посиланням на інших посадовців, управління ризиками група залишить

самим компаніям. Міністр фінансів у суботу повторив головне: «Ми не можемо віддати лідерство китайцям».

Реакція показова. Сенатор Марк Ворнер, віцеполова розвідувального комітету, обережно привітав призначення, але нагадав, що адміністрація «опиралась змістовним запобіжникам, не пропонуючи серйозної альтернативи». Ілон Маск того ж дня оголосив, що перейменує AI-підрозділ SpaceX зі SpaceXAI на SpaceXSI. А Сем Альтман в інтерв'ю Politico, опублікованому 04.10, сказав, що «світ має прийняти, що стануться деякі погані речі, заради переваг цієї технології». Термін «super intelligence» замість AI адміністрація закріпила окремим указом, про перейменування тут писали 23.09.

Чому це важливо. Група складається з розвідки, оборони, фінансів і регулятора конкуренції, тобто у Вашингтоні AI тепер насамперед питання національної безпеки. Для компаній, які продають у США, практичний сигнал такий: жорстких правил найближчі 120 днів не буде, а зміст звіту визначатимуть ті самі люди, з якими вони вже домовились про «морально обов'язковий» пакт 29.09.

ТЕМА 2

Strata: 125-мільярдна Qwen на ігровому ПК з картою на 12 ГБ

ДЖЕРЕЛА · репозиторій Strata на GitHub [GitHub](#) · модель Qwen3.8-Flash-Next на Hugging Face [Hugging Face](#) · обговорення на HN, 04.10 [Hacker News](#)

Strata - безкоштовний установник і рушій під ліцензією MIT, який запускає Qwen3.8-Flash-Next на звичайному комп'ютері. Сама модель вийшла в серпні: 125 млрд параметрів, з них на кожен токен працюють 6 млрд, плюс 51 млрд параметрів n-gram-ембедингів. Усередині 512 експертів на шар, з яких на токен вмикаються 10 і один спільний. Strata використовує саме цю архітектуру: найчастіше вживані експерти тримає у відеопам'яті, усі інші в оперативній пам'яті, процесор рахує решту паралельно, а велика таблиця ембедингів лежить на SSD.

Цифри з README автора: на RTX 5070 з 12 ГБ і 64 ГБ оперативки найстисліша версія (Q2_0) пише 94 токени на секунду, найякісніша з доступних (IQ3_S) - 53. На AMD RX 9070 XT 60 токенів на секунду.

Коментатор на HN з RTX 4090 і 128 ГБ пам'яті повідомив про 124 токени на секунду. Вимоги: від 12 ГБ відеопам'яті, від 32 ГБ оперативки і близько 80 ГБ на диску, завантаження моделі близько 70 ГБ.

Є окрема версія Coder з половиною експертів: за вимірами її авторів вона набирає 91% результату повної моделі на SWE-bench Verified і вміщується в 32 ГБ.

Репозиторій створений 24.09, за 11 днів зібрав понад 11 тис. зірок і 845 комітів, багато з них написані разом з Claude. Скепсис на HN стосується якості: моделі стиснуті до 2-3 біт на параметр, і ніхто не опублікував, скільки розуму вони при цьому втрачають. Ще одне зауваження: README пропонує вставити в агента для коду фразу «встанови Strata

за інструкцією з репозиторію», і один з коментаторів порівняв це з `curl | bash`, тільки гіршим.

Чому це важливо. Ще рік тому модель такого розміру потребувала серверних GPU. Тепер локальний запуск з десятками токенів на секунду коштує як ігровий ПК, і для задач, де дані не можна віддавати в хмару, це вже реальний варіант. Перед тим як на нього покладатись, варто перевірити стиснуту версію на власних задачах: швидкість генерації виміряна, якість на 2 бітах ні.

ТЕМА 3

Аризона: суд скасував вирок, на який вплинуло AI-відео з загиблим

ДЖЕРЕЛА The New York Times, 04.10 [NYT](#) [одне джерело]

Апеляційний суд Аризони залишив у силі обвинувальний вирок Габріелю Оркаситасу, засудженому за вбивство через необережність, але скасував його покарання: 10,5 року в'язниці. У 2021 році в передмісті Фенікса він під час дорожнього конфлікту застрелив 37-річного Крістофера Пелкі. На оголошенні вироку сестра загиблого показала відео, згенероване AI: Пелкі в ньому каже вбивці «я вірю в прощення». Голос узяли з YouTube-ролика, обличчя з плаката з похорону, а текст написала сама сестра. Суддя сказав «Мені сподобався цей AI», назвав відео щирим і призначив максимальне покарання, більше за дев'ять років, яких просили прокурори.

Колегія з трьох суддів назвала це «фундаментальною помилкою». За їхніми словами, відео не відображає реальних подій, але подає слова як такі, що йдуть від самого загиблого, і «стирає дистанцію між тим, що, на думку родини, сказав би загиблий, і його власним голосом». Право потерпілих висловитись, пишуть судді, не може переважити право підсудного на вирок, ухвалений на точній і надійній інформації. Оркаситаса чекає новий розгляд щодо покарання.

Чому це важливо. Це один з перших випадків, коли вищий суд прямо оцінив синтетичне відео в залі суду, і висновок стосується впливу: відео було відкрито позначене як AI, і все одно вплинуло на вирок. Для юристів і суддів аргумент практичний. Реконструкція голосу людини вже має вагу показань, хоча показанням не є.

ТЕМА 4

Дата-центр для Claude у Квінсленді: 2,16 ГВт і 21 500 підписів проти

ДЖЕРЕЛА The Guardian, 04.10 [Guardian](#) [одне джерело]

Біля Далбі, за 180 км на захід від Брисбена, планують найбільший дата-центр Австралії вартістю \$31 млрд для моделей Anthropic. Забудовник - сингапурська Zerra DC, ділянка

в 725 га, де зараз відгодівельний майданчик на 24 тис. голів худоби посеред газових свердловин. Пікова потужність 2,16 ГВт, це приблизно чверть усього споживання штату. Забудовник обіцяє охолодження, яке споживає води як середній торговий центр, і близько 1 400 постійних робочих місць.

Місцеві пам'ятають бум і спад вугільного газу в тому ж регіоні і бачать ту саму історію. Петицію проти за кілька тижнів підписали понад 21 500 людей. Організація Environmental Defenders Office каже, що з січня запитів про дата-центри на її юридичну лінію стало на 250% більше, і закликає до мораторію. У п'ятницю уряд штату забрав рішення про дата-центри в місцевих рад і передав їх на рівень штату, хоча мер Далбі ще місяць тому розраховував, що проект лишиться за радою.

Учора тут писали, що у вівторок керівники OpenAI, Anthropic, Microsoft і Google виступлять перед парламентським комітетом Австралії з питань AI. Дата-центр під Далбі, найімовірніше, теж стане однією з тем.

Чому це важливо. Опір дата-центрам перестає бути американською історією: за даними Guardian, у США їм опираються три чверті населення, а понад 500 округів запровадили заборони.

Чим більший проект, тим більше рішення переїжджає з місцевого рівня на державний, і тим довшими стають строки будівництва, на які розраховують лабораторії.

ТЕМА 5

RemoveMacAI: як вимкнути Apple Intelligence у macOS 27 повністю

ДЖЕРЕЛА · репозиторій RemoveMacAI на GitHub [GitHub](#) · обговорення на HN, 04.10 [Hacker News](#)

У macOS 27, за описом автора утиліти, більше нема одного перемикача для Apple Intelligence, а моделі лишаються на диску після вимкнення функцій. RemoveMacAI робить три речі: вимикає функції (Siri, Writing Tools, Genmoji, Image Playground, розширення ChatGPT, підсумки в Mail, Messages, Safari і сповіщеннях, передбачення тексту, автодоповнення коду в Xcode), видаляє моделі через системний сервіс Apple і не дає системі завантажити їх знову. Для останнього профіль конфігурації перенаправляє завантаження кожної моделі на закритий локальний порт. System Integrity Protection лишається увімкненою, усе можна повернути однією командою.

Тема збрала 402 бали і 265 коментарів на HN. Окремо автор пояснює дрібниці, які заплутують користувачів: у налаштуваннях сховища Apple Intelligence ще деякий час рахується після видалення, бо система стирає файли за власним розкладом, а процес «Siri» лишається запущеним, бо в macOS 27 так називається вікно Spotlight.

Чому це важливо. Вбудований AI у системі тепер займає місце і вимикається складніше, ніж вмикається, і частина користувачів хоче це контролювати. Для компаній

з керованими маками важливіше інше: те саме робиться штатними ключами обмежень Apple у профілі, без сторонніх скриптів.

ТЕМА 6

AI і робота: Леві рахує нові професії, Моллік сперечається з Кубаном

ДЖЕРЕЛА @levie у X, 05.10 @levie · @emollick у X, 05.10 @emollick [одне джерело на цифри]

Аарон Леві, CEO Vox, процитував оцінку LinkedIn, яку переказав акаунт The Kobeissi Letter: з 2023 року в США з'явилося понад 750 тис. робочих місць, пов'язаних з AI, і найбільше серед них, 282 тис., становлять розмітники даних. Леві додає, що статистика недораховує тих, кого всередині великих компаній просто перевели на роботу з AI: банки, фарма, виробники і навіть юридичні фірми наймають інженерів для впровадження агентів.

Марк Кубан відповів класичним «AI не забере твою роботу, її забере той, хто краще за тебе вміє користуватись AI». Етан Моллік з цим не погодився, і в трьох пунктах. Незрозуміло, яка навичка «кращого користування» лишиться цінною, коли AI стає простішим. Дедалі частіше вигоду від AI гальмують організаційні системи, і окремий працівник тут уже не головне. А вплив на робочі місця невизначений і може зачепити цілі категорії професій. Висновок Молліка: робочі системи варто будувати так, щоб AI посилював людей, і автоматизація тут лише один з варіантів.

Чому це важливо. Обидві сторони кажуть про реальне, і цифра LinkedIn це підтверджує:

найбільше нових місць з'явилося в обслуговуванні моделей, розмітці і дата-центрах. Для людини, яка планує кар'єру, корисніше дивитись, як змінюються процеси в її компанії, ніж просто вчитись «промптити краще».

misc

- **Юридичні ризики OpenAI ростуть.** Semafor і FT пишуть про правові наслідки інцидентів з агентами компанії: сповіщення вже отримали понад 100 організацій. Продовження вчорашньої історії про перевірку на 50 ПБ записів. Текст FT за пейволом.
- **WSJ: бюджет на AI майже неможливо спланувати.** За дослідженням, яке цитує газета, лише 11% з майже 400 компаній змогли точно спрогнозувати витрати на AI. Продовження вчорашнього пункту Віллсона про ліміти витрат. Доступний лише перший абзац [одне джерело].

WSJ

- **Tarbell Center виділяє \$10 млн на журналістику про AI** до кінця 2027 року. Перший грант, \$480 тис., отримав новий відеоподкаст NPR Machine Gods від колишніх ведучих Hard Fork Кевіна Руза і Кейсі Ньютона. [Semafor](#)
- **@lennysan** випустив розмову з Тібо Сотто, який керує ChatGPT і Codex в OpenAI: чому основна ставка тепер на персонального агента Dots і чому цикли й графи в агентних конвеєрах, на його думку, тимчасова фаза. Розшифровка платна. [@lennysan](#)
- **@emollick** одним промптом отримав від Fable 5.1 браузерний містобудівний симулятор у стилі бруталізму і порівняв з тим, що рік тому вмiла GPT-5. Код на GitHub.
[@emollick](#)
- **@emollick** про безпеку агентів: йому потрібна операційна система, зроблена для спільної роботи з AI, з детальними правами, видимістю дій агента і окремими агентами-аудиторами. Відповідь на пост Шрірама Крішнана, що давати агенту повний доступ до диска - ознака «незручної проміжної ери». [@emollick](#)
- **@garrytan** про те, що всі будують однакове (цикл агента, пам'ять, пісочниці, конектори): це ознака потреби, і ринок зрештою зійдеться на одній «правильній ОС». [@garrytan](#)
- **Невдале затирання в звіті показало воду і електрику дата-центру Google у Лінкольні, Небраска:** 52,65 МВт пікового споживання і 13,3 млн галонів води за рік. Текст під чорним прямокутником копіювався звичайним виділенням. Новина від 30.09, на HN вистрілила вчора.
[1011now](#)
- **Тайлер Коуен написав блерб до нової книжки Молліка «Coexistence» і адресував його самим моделям.** [Marginal Revolution](#)
- **headstart** запускає компіляцію залежних Rust-крейтів раніше, ніж залежності закінчать перевірку типів, і за словами автора прискорює збірку до двох разів.
[GitHub](#)