

Google opens Gemini 4 Argon to paid users

Anthropic leaves git for Scion, an AI worm on smart devices, a Chrome exploit in 6 days and the Character.ai bankruptcy

Wednesday, 14 October 2026

IN THIS ISSUE

1. Gemini 4 Argon is out for all paid users
2. Anthropic replaced git with its own system: agents merge 30,000 changes per day
3. Anthropic: in a few months, small models may copy themselves on smart devices
4. XBOW hacked Chrome with a full chain in 6 days and \$40,000
5. The FBI arrested an OpenAI engineer over a bug in an executive meeting room
6. Character.ai filed for bankruptcy, Disney offers \$20 million
7. The IMF called an AI stock crash a global risk, the Broadcom CEO compared the market to a casino
8. Ben Recht: the "faster FFT" from OpenAI works only on a computer that does not exist

TOPIC 1

Gemini 4 Argon is out for all paid users

SOURCES Google DeepMind, model page, 13.10 [DeepMind](#) · Ars Technica, 13.10 [Ars Technica](#) · Google DeepMind, X, 13.10 [@GoogleDeepMind](#) · Aaron Levie, X, 13.10 [@levie](#) · Ethan Mollick, X, 13.10 [@emollick](#) · Siqi Chen, X, 14.10 [@blader](#) · HN discussion, 13.10 [Hacker News](#)

On 01.10 this digest wrote about the announcement of Gemini 4 Argon. At that time, Google gave the model only to cyber defenders and trusted testers. On 13.10 the company opened it wider, but only for money. In the Gemini app, the model is available on the AI Pro plan (\$20 per month) and the AI Ultra plan (from \$99 per month). In Antigravity and AI Studio, Argon is available only on Ultra. The story got 1,059 points on HN, the most for the day.

The API price is temporary. Until 30.11, Google charges \$2 for 1 million input tokens and \$10 for 1 million output tokens, with a 95% discount on cached tokens. After 30.11 the price goes up to \$4 and \$20, the level of Claude Opus 5.5. Ars advises developers to plan for this difference in advance.

Google did not change the benchmarks. They are the same as on 30.09.

Google itself gives the new examples. In Chrome, the model helped to separate page processing from an old folder of browser functions with ~6 million lines. For the Gemma 4

E2B model, Argon made inference in vLLM 2.85 times faster, from ~37,900 to more than 108,000 tokens per second. According to Ars, it found unnecessary emoji rendering in YouTube comments, and the server load fell by 45%.

Independent reviews are more careful. Aaron Levie of Box writes that on an internal Box test, Argon scored 14 points more than Gemini 3.1 Pro. Siqi Chen ran his cfo.ai tests. By his measurements, at the introductory price Argon gives the same value as Opus 5.5, but it is two times slower. In the same run, GPT-6.1 Sol gives three times more value than both. Ethan Mollick called Argon the first model from a lab other than OpenAI and Anthropic that reaches the level of their flagships.

Why it matters. Four labs are now at the top, and yesterday's Muse Wonder from Meta also costs \$4 and \$20. In practice, the market came to one price for a flagship model. Thus speed and behavior on a specific task decide the choice of model. The introductory price until 30.11 is attractive, but calculate the budget at the full price.

TOPIC 2

Anthropic replaced git with its own system: agents merge 30,000 changes per day

SOURCES Anthropic, engineering blog, 13.10 [Anthropic](#) · Anthropic, X, 13.10 [@AnthropicAI](#) · Simon Willison, 13.10 [Simon Willison](#) · HN discussion, 13.10 [Hacker News](#)

In May, Anthropic merged about 2,500 changes per day into its monorepo. In September, it merged more than 30,000 per day. Claude writes almost all of this code. The company writes that the tools for people became the bottleneck: version control, build and CI.

The monorepo numbers are large. The history of the main branch takes almost 100 GB, and a compressed working copy takes more than 25 GB. CI makes about 25 million checkouts per day. Without optimizations, this would give more than 600 PB of git traffic per day. At the peak, the company kept several hundred internal git mirrors.

Thus Anthropic built Scion, its own version control system, and moved the monorepo to it fully in 5 months. The main idea is simple: the clients are almost always agents and tasks that are always online. Thus the server itself calculates history, blame, merges and rebases. The client keeps only part of the files and gets the rest on request. The Scion infrastructure needs 200 times less CPU and memory than the old git mirrors. Checkout became 100 times faster, and CI released 25% of its CPU.

A team of 7 people built Scion and the Antique review system. All of them rate Claude as a better programmer than themselves. But the team says that the job of an engineer was often to tell Claude "no". Agents broke the limits of the system, and adversarial review pushed toward extra complexity.

People still review each change. Now they check the intent of a change and do not read each line.

The company could not keep a classic merge queue with a global lock and moved to tests after merge.

Why it matters. Anthropic itself writes that Scion will not replace git for others. But it expects other companies to meet such problems in 6-12 months. For a team with agents, the conclusion is practical. When agents write code 10 times faster, CI, review and merge break first.

Measure these points now.

TOPIC 3

Anthropic: in a few months, small models may copy themselves on smart devices

SOURCES Anthropic, Frontier Red Team, 12.10 [Anthropic](#) · The Guardian, 13.10 [Guardian](#) · Financial Times, 13.10 [FT](#) · Semafor, 13.10 [Semafor](#) · Ethan Mollick, X, 13.10 [@emollick](#)

The Anthropic red team checked if an AI model can copy itself from one device to another device.

The targets are cheap internet-of-things devices: smart ovens, no-name Android TV boxes. Such a worm carries the model weights with it and runs the model directly on the hacked hardware.

The first experiment showed that the hardware is sufficient. The team bought 15 Android TV boxes at ~\$55 each and ran the Gemma4-26B-A4B model on them. Each box held two layers of the model. The speed was 0.39 tokens per second. A model with 9 billion parameters set up inference by itself on 8 of 9 hardware types.

The second experiment tested the full cycle: find a target, copy itself, start inference and repeat. The team made the hack simple here: an open ADB port without a password. Only GLM-5.2 completed the full cycle. It infected all 8 nodes in 4 of 5 attempts. Small models up to 30 billion parameters still make errors when they connect the steps together.

The conclusion of the team is based on a trend. Small open models grow by 13.7 points of the ECI index per year. If the trend continues, a small model will reach the level of GLM-5.2 by the end of

1. Such a worm would be slow: the team estimates that one copy sets up inference in 1-2 days. But it learns: notes from earlier copies raised the success rate of Qwen3.5-9B from 45% to 61-84%.

Why it matters. This is an estimate from the company, and the authors made the weakest step, the hack itself, simple on purpose. Thus read the result as an upper limit. The defense is known since Mirai: no default passwords, regular updates and a provider quarantine for infected customers.

According to Anthropic, such a quarantine once removed 92% of Mirai infections in 14 days.

TOPIC 4

XBOW hacked Chrome with a full chain in 6 days and \$40,000

SOURCES XBOW, blog, 13.10 [Xbow](#) · XBOW Labs, technical analysis, 13.10 [Xbow](#) · Greg Brockman, X, 13.10 [@gdb](#) [single source]

XBOW sells automatic security tests. The company explained how it built a full Chrome exploit chain. With such a chain, an open web page runs code on the Windows computer itself. Google paid \$250,000 for it. This is the first such reward under the rules that Google introduced in April

1. All three bugs are already fixed.

The chain has three bugs. The first is in Maglev, the JavaScript optimizer of the V8 engine. The second is in the animated WebP decoder, and it gives an escape from the V8 sandbox. The third is in Crashpad, the crash handler, and it gives an escape from the sandbox of the full renderer. Agents on the public OpenAI model GPT-5.6 Sol did the search. People selected the targets, checked the results and decided which parts to connect.

The first version of the chain failed: Google already knew about one of the bugs. The team made each part replaceable in advance, thus it quickly built a new chain. This took 6 days, from 2 to 7 August. According to the Codex logs, the compute cost about \$40,000. XBOW estimates that before large models, such work took a team of 4 people 6-9 months.

Why it matters. XBOW itself gives the numbers, but Google confirmed the result with the payment.

The main point is the price: \$40,000 and a public model. Thus a full browser exploit is now within reach of teams outside governments too. XBOW gives two simple pieces of advice: restart Chrome immediately after an update, and test your own code before the attacker does.

TOPIC 5

The FBI arrested an OpenAI engineer over a bug in an executive meeting room

SOURCES The San Francisco Standard, 13.10 [Sfstandard](#) · The New York Times, 13.10 [NYT](#) · The Wall Street Journal, 14.10 [WSJ](#) · The Guardian, 13.10 [Guardian](#) · Financial Times, 14.10 [FT](#) · HN discussion, 13.10 [Hacker News](#)

The FBI arrested OpenAI engineer Abdulrahman Alfozan, 35, a citizen of Saudi Arabia. Prosecutors say that he hid a recording device in an executive meeting room at the San Francisco headquarters.

Staff found the device on 30.09. It was in a USB port behind a TV, wrapped in toilet paper. The recordings on it were encrypted.

The prosecutors describe the preparation. Before the installation, Alfozan asked an internal tool about office addresses, meeting rooms and executive calendars. From 24 to 30.09, a

chatbot account linked to his wife asked how to make a continuous audio recording and encrypt an SD card. On 05.10 security found his bag with a wallet and a second device of the same type.

OpenAI writes that its technical counterintelligence team found the device during a routine check.

The company used its own models to identify the device. Alfozan worked at OpenAI from July 2025. He faces up to 5 years in prison and a \$250,000 fine. The prosecutors do not name a motive or a client.

Why it matters. US authorities directly call the case a national security matter, and the FBI calls the protection of AI research a counterintelligence priority. For the labs, this means costs for physical security, as in defense companies. And the complaint shows that a log of requests to AI tools is now also evidence in court.

TOPIC 6

Character.ai filed for bankruptcy, Disney offers \$20 million

SOURCES Character.ai, blog, 13.10 [Character](#) · The Verge, 13.10 [Theverge](#) · Ars Technica, 13.10 [Ars Technica](#) · The Wall Street Journal, 13.10 [WSJ](#) · HN discussion, 13.10 [Hacker News](#)

Character.ai, one of the first hits of generative AI, filed for Chapter 11 bankruptcy in the US.

The company wants to sell the business under court supervision by the end of November or the start of December. If nobody buys it, the platform will close. For now, the service and paid subscriptions work.

The main bidder is Disney. It offers \$20 million for part of the assets and a non-exclusive license to part of the technology. In September 2025, Disney demanded the removal of its characters from the platform and wrote that Character.ai chatbots harm children. Now the CTO of Disney is Karandeep Anand, the former CEO of Character.ai. According to The Verge, part of the technical team will also move to Disney.

In 2024, Google paid more than \$2 billion to take the founders and part of the team. NBC data, quoted by Ars, puts the debts of the company at \$50 million to \$100 million. Two states, Kentucky and Pennsylvania, still sue it. Ars also points to the data: the privacy policy permits the transfer of user chats to the buyer together with the assets.

Why it matters. Last year the service had more than 20 million monthly active users. The main bidder offers only \$20 million for part of the assets. A large audience without protection from lawsuits has a low value. For chatbot users, the conclusion is direct: in a bankruptcy, chats become an asset that is sold together with the company.

TOPIC 7

The IMF called an AI stock crash a global risk, the Broadcom CEO compared the market to a casino

SOURCES The Guardian, IMF report, 13.10 [Guardian](#) · The Guardian, markets, 13.10 [Guardian](#) · Financial Times, 13.10 [FT](#) · Semafor, 13.10 [Semafor](#) · HN discussion, 13.10 [Hacker News](#)

The IMF released its half-year report on financial stability. The fund called a correction of AI stocks one of the three main risks for global markets. The other two are more expensive government loans and a large cyberattack. The IMF says that the dependence of the rest of the world on US assets, especially AI, is now the highest in history.

The fund also writes about circular financing. In three years, hyperscalers made deals worth \$300 billion with weaker AI companies and added \$650 billion of debt. In return, they got promises of \$1.4 trillion in future payments. The seven largest tech companies take more than 35% of the S&P 500 index, against 12% ten years ago.

On the same day, Broadcom CEO Hock Tan said that many investors "do not understand what they invest in", and compared the market to a casino. The Chinese company CXMT announced progress in a new DRAM memory architecture. The Korean Kospi index fell by 6.25%, Samsung by 7.4%, SK Hynix by 10%, SoftBank by 8.5%.

Why it matters. A warning from the IMF and from a chip vendor on the same day sounds more convincing than single voices of skeptics. The Kospi drop shows how fast the market reacts to one news item from China. For companies that build a product on cheap tokens, this is a signal: the price of compute also depends on whether the market can carry the credit load.

TOPIC 8

Ben Recht: the "faster FFT" from OpenAI works only on a computer that does not exist

SOURCES Ben Recht, arg min, 12.10 [Argmin](#) · HN discussion, 13.10 [Hacker News](#) [single source]

On 08.10 this digest wrote that Scott Aaronson praised the OpenAI math papers. Among the results, he named a Fourier transform faster than $O(n \log n)$. On 09.10 OpenAI withdrew three papers from this set. Now the Berkeley professor Ben Recht analyzed the FFT result itself. He called it a "harmful exploit" of the rules.

The argument of Recht is about the model of computation. The algorithm assumes exact complex arithmetic: it multiplies numbers of any length in one operation. These numbers grow with n faster than any polynomial. In bits, such storage would need almost exponential memory. The hidden constant, by an estimate from the paper itself, is about 10^{100} . Thus the classic FFT is faster at any real problem size.

Recht reminds readers that the limits are known for a long time. In 1973, Jacques Morgenstern proved a lower bound for FFT with bounded coefficients. Later, Nir Ailon showed that a result faster than $n \log n$ is not possible without strong numerical instability.

Recht thinks that AI here only automated an old flaw of theory: to prove things that nobody will ever run on a computer.

Why it matters. Mass production of proofs by a model hits the same wall as code from an agent. A formally correct result can fail to answer the real question. A Lean certificate catches an error in a proof. It does not catch a question that has a wrong statement.

misc

- **Ethan Mollick (@emollick)** expects a "legal DDoS attack" on customer support. Agents such as Muse and Dot can demand old refunds without end and negotiate for millions of people.
[@emollick](#)
- **Physical Intelligence (@physical_int)** opened the beta of the Pi API. The company already adapted its models to partner robots, from strawberry harvest to server assembly.
[@physical_int](#)
- **Exa:** the search engine for agents passed \$100 million in annual revenue, writes co-founder Will Bryk (@WilliamBryk). The company grows 10 times per year for three years in a row and serves 2 billion queries per month. [@WilliamBryk](#)
- **Hidden capabilities.** A NeurIPS 2026 paper shows how to hide an encrypted neural network in model weights, and only a secret trigger turns it on. The authors prove that a capability audit cannot find such a network, even with full access to the weights. [arXiv](#)
- **Rust:** the borrow checker has an 11-year-old type safety hole. The Rust team gave it P-high priority and writes that you cannot hit it by accident: it needs a deliberately malicious Drop. [GitHub](#)
- **Google DeepMind (@GoogleDeepMind):** from a video, Argon wrote a physics simulation by itself and trained the Apptronik Apollo humanoid to hit the cornhole target 7 times in a row.
[@GoogleDeepMind](#)
- **Apple** showed HomeView, a 7-inch home display for \$299, and a new Apple TV 4K.
[Ars Technica](#)