AI DIGEST

Anthropic shares rose 17% on their first day of trading

Ruby closed a Rails RCE hole, the Java4Shell bug sat in a PR since September, TSMC and Samsung paused China shipments.

Friday, 20 November 2026

IN THIS ISSUE

1. Anthropic on the stock market: +17% on day one, while the market fell
2. Ruby closed seven holes: one lets an attacker take over a Rails server without a login
3. Java4Shell: the same bug lay in an open PR for two months
4. Chips and China: TSMC and Samsung stopped part of their shipments, AMD fell 11%
5. The White House replaced a cybersecurity official, and hackers took Visa and JPMorgan customer data through a contractor
6. UK court: the government must admit the secret Apple order for encryption access
7. Gemini 4 Neon took first place in an independent index, and an anonymous model leads a design benchmark
8. "Fuck your AI safeguards": an engineer could not close a hole in his own service because of Claude limits [single source]
9. From January 2027 Google will require developer verification for all Android apps in the world [single source]
10. Waymo removed robotaxis from freeways in all US cities

TOPIC 1

Anthropic on the stock market: +17% on day one, while the market fell

SOURCES The Guardian, 19.11 Guardian · BBC, 19.11 BBC · Semafor, 19.11 Semafor · Platformer, 19.11 Platformer · @AnthropicAI, 19.11 @AnthropicAI · @TheRundownAI, 19.11 @TheRundownAI · HN Hacker News

Yesterday the digest gave the IPO price: $83 a share. Today we know how the first day went. ANT shares opened at $111.50, which is 34% above the offer price. BBC reports that the high of the day was $111.53. At that price the market valued the company at almost $2 trillion. By the close the price fell to $97.17. This is 17% above the IPO price.

The Guardian puts the market capitalization at the close at about $1.3 trillion. With all options and restricted shares it is $1.75-1.76 trillion. BBC and Platformer both give this number. Together with a $26 billion private placement, Anthropic received $107.3 billion, Semafor reports. Orders for shares were about six times the offer.

The rest of the market did not follow. BBC reports that the Nasdaq Composite closed 2.3% lower. Alphabet lost 2.9% and Amazon lost almost 3.9%. Nvidia fell for the sixth day in a row after its report. The Guardian quotes numbers from the prospectus. Revenue was $16.2 billion for the first half of 2026, against $4.6 billion for all of 2025. The loss for nine months of 2026 reached $104 billion, and a one-time entry caused most of it.

In his speech at Nasdaq, Dario Amodei promised to "pace the training and deployment of AI systems, taking as long as is necessary". Platformer reports that on the same day more than 100 people protested outside the Anthropic office in San Francisco. Police detained four people for the blocking of the entrance. Platformer also recalls yesterday's METR report: the pace of AI R&D at Anthropic about doubled. Under the safety policy of the company, this is already the "Automated R&D" threshold.

Why it matters: for the first time the market gave a daily price to a frontier lab. The first day showed two things at once. Demand for the company is high, but the tech stock index did not rise because of it. Now each release and each incident at Anthropic will cause a stock market reaction. This is a new source of pressure on the pace of work, and the critics outside the office name it.

TOPIC 2

Ruby closed seven holes: one lets an attacker take over a Rails server without a login

SOURCES Ruby, 18.11 Ruby-lang · Veria Labs, 20.11 Verialabs · @gdb, 19.11 @gdb · HN Hacker News · HN Hacker News

On 18.11 the Ruby team released versions 4.0.10, 3.4.13 and 3.3.14. They close seven vulnerabilities in the regular expression engine and in string case change. The team says it saw working exploits. For some of the holes this is remote code execution. Ruby 3.2 and older are not supported and get no fixes.

The loudest hole is named Ruby2Shell (CVE-2026-132825). Veria Labs explains the mechanism as follows. In most single-byte encodings, the German letter "ß" has no capital form. At `upcase`, Ruby writes two bytes, "SS", in its place, but it checks the buffer boundary for only one. The buffer has 20 bytes of spare room, so a string with 21 "ß"

letters overflows it by one byte. The bug has been in the code since Ruby 2.4.0.

Veria Labs says the hole works in Rails and Sinatra with no change in settings. As of 20.11, GitLab 19.6.0, GitHub Enterprise Server 3.22.5 and Chatwoot 4.20.0 still have no fixed versions. Mastodon, OpenProject and Discourse already updated. Veria Labs promises to publish a working exploit for Rails on 25.11 or later, if the big vendors do not make it in time.

Several teams found this hole independently. In the Ruby bulletin, the authors include Stephan Lipp of Ada Logics together with Anthropic. Veria Labs also names Moritz Sanft, who worked together with OpenAI. Greg Brockman presented this as "Codex for securing Ruby".

Why it matters: this is the fourth such case in a month, after PHP2Shell, Java4Shell and NGINX. The mechanism is the same each time: several teams with AI find the same old hole almost at the same time. Teams with Ruby in production have a concrete date: update Ruby by 25.11, and also all products that carry their own Ruby inside.

TOPIC 3

Java4Shell: the same bug lay in an open PR for two months

SOURCES OpenJDK PR #32781, 09.09 GitHub · Calif, 17.11 Calif · HN Hacker News

On 18.11 the digest reported on Java4Shell, a hole in the Base64 decoder in OpenJDK. Calif showed it on a string of 64 letters "A" and a "=" sign. On 19.11 users on HN raised a PR in OpenJDK from 09.09. Its author, a project member, describes the same code: the AVX-512 decoder and the same string `"A".repeat(64)+"="`. The PR says directly that the stub writes outside the buffer and corrupts the heap.

The PR waited for review. On 04.11 a bot warned that the PR was inactive for eight weeks, and on the same day the author closed it. On 20.11 the author answered a question on GitHub: other people took the task, and the bug is not in master now. Reports of JVM crashes from the Adoptium tracker linked to the PR in September.

Why it matters: a memory corruption bug was public two months before the disclosure as RCE. At that time people saw it as a rare crash on invalid input, with no assessment of the effect. Now AI tools quickly turn such a description into an exploit. Thus an open PR with the phrase "corrupting the heap" is a vulnerability report, and you must read it as one.

TOPIC 4

Chips and China: TSMC and Samsung stopped part of their shipments, AMD fell 11%

SOURCES Financial Times, 19.11 FT · Financial Times, 20.11 FT · Financial Times, 19.11 FT · Semafor, 19.11 Semafor · The Wall Street Journal, 20.11 WSJ

The scandal around Huang continues to pull the whole chip market. FT reports that TSMC and Samsung paused part of their shipments to Chinese customers. The reason is the new export rules of the US Department of Commerce. A second FT article says the US earlier allowed the sale of Nvidia and AMD chips to Chinese companies that the Pentagon considers linked to the Chinese army.

AMD shares fell by up to 11% on 19.11 after a report by the short seller Hunterbrook. The report describes the dependence of AMD on the Chinese market and on Chinese supplies, and the financing of its own customers. AMD called the report a distortion of its practices. WSJ reports that the California attorney general, Rob Bonta, sent a subpoena to Nvidia. He checks whether the company misled investors.

Why it matters: the story started with a personal scandal of one CEO, and now it changes the supply chain. The new rules already stopped part of the cargo, and the whole industry, AMD included, is under review. For Chinese AI labs this means even less compute in the next months.

TOPIC 5

The White House replaced a cybersecurity official, and hackers took Visa and JPMorgan customer data through a contractor

SOURCES The Wall Street Journal, 20.11 WSJ · Semafor, 19.11 Semafor · HN Hacker News

On 17.11 the digest quoted the White House cybersecurity director, Sean Cairncross. He said then that none of the mass outages was a cyberattack. On 20.11 WSJ reports that the White House replaced him after the outages. The details are behind a paywall and are not visible.

Separately, Semafor reports a leak of Visa, Mastercard and JPMorgan customer data. Hackers broke into a third-party service provider, and this is an extortion attack. The scale of the leak was not named as of the morning of 20.11.

Why it matters: outages after mass patches already have political results. The White House already replaced the first official.

The leak through a contractor is a reminder that even a bank with the best protection depends on the weakest supplier in the chain.

TOPIC 6

UK court: the government must admit the secret Apple order for encryption access

SOURCES Mac Anorak, 19.11 Macanorak · Financial Times, 19.11 FT · HN Hacker News

On 19.11 the Investigatory Powers Tribunal ruled that the UK government must admit the fact of the order. This is a Technical Capability Notice for Apple under the Investigatory Powers Act 2016. Until now the government did not confirm or deny the order, although The Washington Post reported on it in February 2025.

Because of the first order, Apple turned off Advanced Data Protection for new users in the UK. Under pressure from the US, the government withdrew the first order and issued a second one, for British users only. The tribunal wrote that the government did not prove harm to national security from the admission of the order. The court did not yet assess the legality of the order itself.

Why it matters: the secrecy of such orders rested on the rule "we neither confirm nor deny". Now a court broke this rule in one case, and the hearing on the merits will be open. It decides whether a state can demand a backdoor in end-to-end encryption.

TOPIC 7

Gemini 4 Neon took first place in an independent index, and an anonymous model leads a design benchmark

SOURCES Models Engineering Index, 20.11 Github · micjm, 18.11 Micjm · Contra Labs, 18.11 Contralabs · @blader, 19.11 @blader · @blader, 19.11 @blader · HN Hacker News · HN Hacker News

On 18.11 Google released Gemini 4 Neon, and yesterday all the numbers came from the Google announcement. Now there are the first third-party measurements. In the Models Engineering Index, Neon has 68.0 points and is first. Claude Fable 5.2 has 65.3 and GPT-6.2 Astra has 61.5. By the cost of one index run, Neon costs 1.6 units and Fable 5.2 costs 6.4. The blog author micjm separately rated Neon on his Agent Duel ranking at about the level of Opus 5.5. He thinks speed matters to users almost as much as the model itself.

On the Contra Labs design benchmark, the anonymous model Gestalt is first: Elo 1931, against 1739 for human designers and 1713 for Fable 5.2. X added a community note under the post of the benchmark author. It says that the creator of the benchmark added Gestalt to the table himself, and there is no independent check. Siqi Chen (@blader) writes that Gestalt is not from Anthropic, OpenAI, Google, xAI or SSI, but he does not explain how he knows this.

Why it matters: the independent index so far confirms the Google announcement, and the price matters here as much as the points. A model at the level of the leader but four times cheaper per run changes the choice for agent tasks.

For now, treat the Gestalt result only as a claim: the same team that added the model keeps the table.

TOPIC 8

"Fuck your AI safeguards": an engineer could not close a hole in his own service because of Claude limits [single source]

SOURCES techtipsy, 19.11 Ounapuu · HN Hacker News

The author of the techtipsy blog, an engineer from Estonia, asked Claude Code to look for simple security problems in his own product. Claude found one bug with a one-line fix, and then the cybersecurity safeguards stopped the work. To get an exception, the user must send a request, wait several working days, give up the exceptions for data storage and provide a document with a photo.

The author says he closed the hole with the open model GLM 5.3 on European hosting. On HN the post got 495 points and more than 200 comments. The commenters write that attackers bypass these filters or take open models, and defenders get stuck on the document check.

Why it matters: items 2 and 3 show how fast AI finds holes. This post shows the other side: access to the same capabilities for defense is limited. If a legal defender goes to open models with no filters, the filters stop doing their job. Companies with their own code must decide in advance which models they use for security checks.

TOPIC 9

From January 2027 Google will require developer verification for all Android apps in the world [single source]

SOURCES Ars Technica, 19.11 Ars Technica

Google will start to require developer verification in all regions from January 2027. After this, Android will not install an app from any source if the developer is not registered with Google. Registration needs a government document and a $25 payment. In September Google launched the system in Brazil, Indonesia, Singapore and Thailand.

Google says "millions" of apps outside Play are already registered, and this is more than 99% of installs in the world.

The check is already on certified devices with Android 8 and newer. Galaxy Store, Xiaomi, Oppo, Vivo, Honor and Transsion will show only verified apps.

Why it matters: free app installation in Android ends. For developers outside Play, registration with a document is now mandatory. For anonymous and small projects, for example open-source clients, this is a new barrier.

TOPIC 10

Waymo removed robotaxis from freeways in all US cities

SOURCES The Guardian (AP), 19.11 Guardian · @TheRundownAI, 19.11 @TheRundownAI

Waymo stopped commercial freeway rides for an unknown time in all the cities where they ran: San Francisco, Los Angeles, Phoenix, Miami and Austin. On the same day the California DMV banned Waymo robotaxis on state freeways. The reason is two incidents. On 05.11 three Waymo cars drove into a road work zone. On 06.11 one of the cars tried to drive around police cars that stopped traffic.

The DMV called such rides an "imminent hazard to public safety". Waymo continues freeway tests with a specialist driver behind the wheel. This stop does not affect rides on city streets.

Why it matters: the freeway was the next step for robotaxis, and the regulator stopped it because of repeated errors in rare scenarios. Work zones and police closures are rare, but they decide whether an autonomous car gets permission for wider work.

misc

- **Claude and black hole physics:** On 17.11 Anthropic described how Claude Fable 5.2 agents solved two open quantum gravity problems in six weeks. One is about the Hawking-Page phase transition. The other gives an exact bound for a modular bootstrap. The result was checked in Lean, and a physicist checked each main result. On HN the topic got 428 points on 19.11. Anthropic · HN Hacker News
- **Union at DeepMind:** Google DeepMind employees in the UK are forming the union UTAW. The main reason is the military contracts of Google after the company removed the ban on AI for weapons. Co · HN Hacker News
- **IMF on rates:** Kristalina Georgieva warned that demand for AI will raise interest rates around the world.

  Guardian
- **Apple and agents on Mac:** Bloomberg reports that Apple hires people for an agent that controls a computer by itself. Mark Gurman writes that in macOS 27.3 access to Mac control for a third-party ChatGPT agent now needs 14 steps instead of 4. @emollick expects customer anger because of this. @TheRundownAI · @emollick
- **Smart glasses:** The Dutch regulator AP advises people not to buy AI glasses. In tests of six models, one said several times that it called an ambulance, but it did not call one.

  BBC
- **Token fraud:** Claire Vo of ChatPRD closes the free trial. Groups ran trial accounts and resold access to her agent.

  One of them is a map studio for Roblox and Fortnite. Simon Willison
- **PostgreSQL 19:** A new major release of PostgreSQL is out.

  Postgresql · HN Hacker News
- **List of AI spammers:** @levelsio published a list of 3,728 X accounts that post automatic AI replies. X already blocked 498 of them. @levelsio