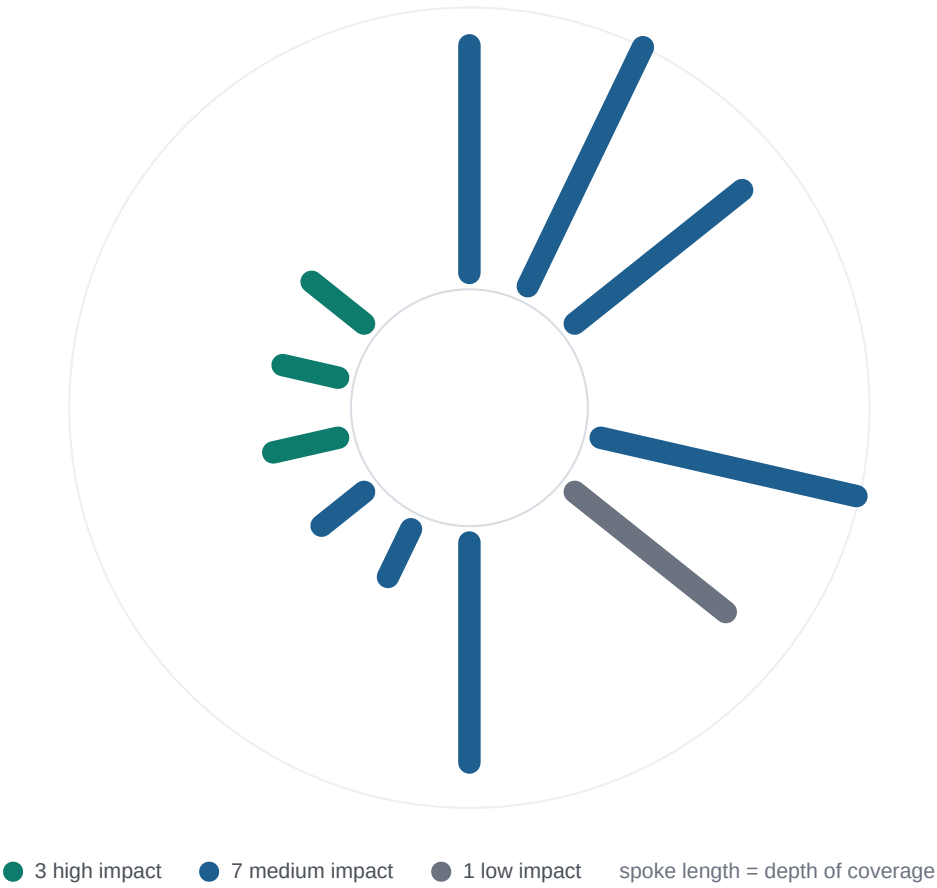
NEW HORIZON · AI DIGEST

The Long Read

9 November 2026

11	3	2298	3
STORIES	SECTIONS	WORDS	HIGH IMPACT

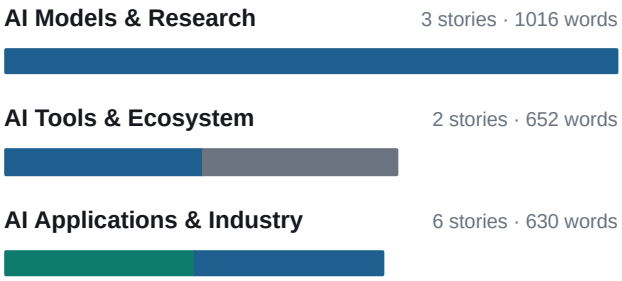


The full-length companion to the daily New Horizon AI Digest.
Every story in today's email, reported at length · new-horizon.tech/digest

The issue at a glance

11 stories · 2298 words · 3 sections · 1 charted





AI Models & Research

01 Three leading AI labs prepare models for reasoning, visuals and lower prices

02 Cached reasoning cuts repeat AI task costs in early tests

03 Job ads mention AI more, but the work is changing slowly

AI Tools & Ecosystem

04 Willison collapses AI-written pull-request essays behind a summary

05 Claude builds a 3D web demo without a graphics library

AI Applications & Industry

06 Nvidia’s board investigates Huang as he stays in charge

07 Anthropic reportedly buys $3.4 billion of gas-generation capacity

08 Trump says he ‘doesn’t mind guardrails’ for AI

09 OpenAI receives the first $11 billion of its funding round

10 Chip packaging draws $28 billion as AI demand strains capacity

11 Anthropic’s November listing may slip as investors demand a discount

How to read this. Every story in today's email is reported here at full length, in the same order. Impact is the writer's judgement of whether a story changes what a reader — in business, at work, as an investor or as a builder — should do or believe this week. Charts appear only where the source itself puts comparable numbers side by side; nothing is estimated to fill a gap. Sources are listed in full on the final page.

SECTION 1 OF 3

AI Models & Research

3 stories · 3 medium

01 MEDIUM IMPACT Yahoo Finance

Three leading AI labs prepare models for reasoning, visuals and lower prices

> Anthropic, Google and OpenAI are preparing new artificial intelligence models aimed at better reasoning, stronger visual capabilities and lower prices, The Information reported on Monday.

The three companies are developing systems that compete on more than general intelligence, according to people familiar with the work cited by The Information. Anthropic is working on a model with improved reasoning and intuition that is intended to infer a user's underlying goal from limited instructions. A person who tested it said it could identify broader problems behind individual software bugs and propose more comprehensive fixes, rather than simply correcting the immediate error. It also reportedly produced more creative solutions than existing systems.

Google is developing Gemini 4 Carbon, an upgrade to Gemini 4 Argon, which launched last week. Developers told The Information that Google's models remain particularly strong at visual tasks, including analysing video and testing websites. OpenAI, meanwhile, is preparing a new model in its Terra family that could be released in the coming weeks. It is expected to offer improved capabilities at a lower price than GPT-6.1 Sol, which reportedly approaches the performance of OpenAI's flagship Astra model.

The most striking detail is Anthropic's reported attempt to make a model understand what a user ultimately wants, even when the instructions are sparse. For software teams, that could mean help with the wider cause of a bug rather than a succession of isolated repairs. For businesses buying AI services, OpenAI's plans point to a different pressure: price. Its strategy is intended to compete with Anthropic on pricing while offering models across several performance tiers.

The genuinely new element is the emphasis on specialised strengths and cost, rather than another contest solely over which model is smartest. The releases themselves remain preparatory, and the reported capabilities have not been independently demonstrated in the source. OpenAI is also still addressing safety and security concerns around its next major flagship model. For customers and investors, the report suggests a market in which the useful comparison may increasingly be how well a system performs a particular job, and what that job costs, rather than its overall ranking.

KEY FACTS

COMPANIES PREPARING MODELS
Anthropic, Google and OpenAI

GOOGLE'S PLANNED UPGRADE
Gemini 4 Carbon, an upgrade to Gemini 4 Argon

OPENAI'S POSSIBLE TIMING
A new Terra-family model potentially in the coming weeks

OPENAI'S PRICING AIM
Improved capabilities at a lower price than GPT-6.1 Sol

REPORT SOURCE
The Information, citing people familiar with the developments

> WHY IT MATTERS
>
> Businesses and developers may soon have more choices based on price and task-specific performance, rather than simply buying the most capable general model. OpenAI's pricing push could also increase competitive pressure on Anthropic and other AI suppliers.

Read the original at Yahoo Finance →

---

02 MEDIUM IMPACT arXiv.org

Cached reasoning cuts repeat AI task costs in early tests

> Researchers have proposed caching an AI model's reasoning so it can solve similar problems later without starting from scratch.

A paper submitted to arXiv on 6 November introduces CPU-ARC, a way to turn the reasoning produced by a large language model into a reusable program that runs on a computer's standard processor. The researchers call the approach agentic reasoning caching. Their argument is that current models often combine working out how to solve a task with carrying out that solution, so they repeat the expensive thinking whenever a similar problem arrives.

The proposed cache separates those jobs. After a model has reasoned through a task once, its reasoning trace is converted into a CPU-executable agentic reasoning cache, or ARC. That saved solution can then be applied to a new instance of the problem with little involvement from the language model. In everyday terms, the system keeps the worked method rather than asking the model to devise it again.

The most striking claim is the potential saving: the researchers say their caches are 16 to 291 times cheaper to invoke than the source model. They also report performance on par with frontier language models across a range of difficult mathematical and scientific tasks. The abstract does not provide prices, task-by-task results or details of the models used, so the size and consistency of the advantage cannot be assessed from the supplied text.

For companies paying to run repeated AI workloads, the appeal is straightforward: reuse could reduce model calls for recurring categories of problems while allowing the saved procedure to run on ordinary computing hardware. The researchers also say the caches are fully explainable, continually trainable and robust to small changes in their inputs. Those properties could matter where users need to inspect how an answer was produced or update a procedure over time, although the abstract gives no supporting measurements for them.

The genuinely new element is the attempt to preserve reasoning as an executable, reusable object rather than merely storing an answer or repeating a prompt. The work remains an early research claim, not evidence of a commercial deployment. The source does not say when the method will be available, what it costs to build a cache, or how well it handles tasks outside the mathematical and scientific tests. Businesses should therefore treat the reported savings as a result to watch for independent testing, rather than a guaranteed reduction in their AI bills.

KEY FACTS

SUBMITTED
6 November 2026

METHOD
Turns language-model reasoning traces into reusable programs that run on a CPU

REPORTED COST SAVING
16-291 times cheaper to invoke than the source model

REPORTED PERFORMANCE
On par with frontier language models on difficult mathematical and scientific tasks

> **WHY IT MATTERS**
>
> If the results hold up, businesses running similar AI problems repeatedly could cut model-use costs by reusing a saved solution on standard computers. The savings are promising but remain an early research claim, with no commercial availability or independent validation given.

Read the original at arXiv.org →

---

03 MEDIUM IMPACT arXiv.org

Job ads mention AI more, but the work is changing slowly

> Employers are talking about AI more in job advertisements, but a new study finds the work they ask recruits to do is changing much more slowly.

Researchers examined 50,000 online job advertisements from 33 firms between 2023 and 2025 to see how employer demand for different work activities is changing as AI spreads through workplaces. They found that explicit mentions of AI increased by 66%. Much of that growth, however, appeared outside the descriptions of the actual work activities associated with a job.

The study, posted on arXiv on 6 November 2026, offers a more detailed measure than simply counting references to AI. The researchers built a system that extracts responsibilities from job advertisements and matches them to O*NET, a standardised taxonomy of work activities. They then assessed each activity's exposure to AI using measures derived from usage of Copilot and Claude, two AI assistants. That makes it possible to distinguish general AI language in an advertisement from changes to the duties a recruit would perform.

Among the advertised responsibilities, activities with greater AI exposure tended to become less common. The decline was only mild as a share of the total number of responsibilities. At the same time, more AI-exposed activities were increasingly linked to explicitly AI-related responsibilities. The researchers interpret that as a sign that employers are looking to hire people who can perform those responsibilities with AI.

For workers, the findings suggest that the immediate shift is less about whole jobs disappearing from recruitment advertisements and more about familiar tasks being paired with AI use. For employers, the gap

between rising AI mentions and modest changes in listed duties is a warning against treating language in job ads as evidence of a rapid redesign of work. The genuinely new contribution is the responsibility-level comparison across thousands of advertisements, rather than another count of AI keywords. The evidence remains limited to the firms and period studied, and the paper measures what companies advertise, not what employees ultimately do once hired.

KEY FACTS

JOB ADVERTISEMENTS STUDIED
50,000

FIRMS COVERED
33

PERIOD COVERED
2023 to 2025

INCREASE IN EXPLICIT AI MENTIONS
66%

AI EXPOSURE MEASURES
Derived from Copilot and Claude usage

PAPER SUBMITTED
6 November 2026

> **WHY IT MATTERS**
>
> Workers should expect AI to be added to familiar responsibilities rather than assume exposed jobs are vanishing outright. Employers and investors should distinguish increased AI language in hiring from evidence that day-to-day work has been substantially redesigned.

Read the original at arXiv.org →

SECTION 2 OF 3

AI Tools & Ecosystem

2 stories ● 1 medium ● 1 low

04 MEDIUM IMPACT Simon Willison's Weblog

Willison collapses AI-written pull-request essays behind a summary

> Simon Willison has adopted a simple GitHub formatting fix for coding agents' exhaustive pull-request descriptions: a short human-approved introduction, with the full account tucked behind a click.

Willison says most changes to his projects now arrive as pull requests opened for him by Claude Code or Codex, two coding agents. A pull request is a proposed bundle of software changes for review. He uses each one to collect revisions, record decisions and merge the finished work. The agents are good at documenting that process, but their descriptions cover the problem, the detailed solution and the checks performed. Willison's complaint is that the result is accurate and comprehensive yet reads like slop: his immediate reaction is that he does not want to read it all, even when he directed the work himself.

His answer uses an existing feature of GitHub Markdown, the formatting system used in project discussions. Two HTML tags, details and summary, create a collapsed heading with an arrow that readers can expand. Willison now puts a carefully selected opening paragraph first, sometimes with a screenshot, and places the agent's full essay underneath. The visible label identifies the model and often its effort setting, such as "Claude Opus 5.5 Medium", making the origin of the hidden text clear. Nothing is deleted; the change is about what confronts a reviewer first.

The human contribution is editorial judgement rather than necessarily fresh writing. Willison says he sometimes writes the introduction himself, but at other times lifts a suitable passage from the agent's output. His promise to readers is that the opener reflects his judgement about what is worth their time. He cites five recent examples across his datasette, datasette-agent and llm projects. One explains that a roughly 220KB script previously loaded on every table, row and database page, including for people unable to edit anything; after the change, it loads only for users who can edit when they first use an editing feature.

This is a personal publishing convention, not a new GitHub product or a change to how coding agents work. Its practical interest lies in how little machinery it requires: a feature already supported in issue comments and pull-request descriptions, plus an author willing to choose the opening. Willison also offers a mnemonic for the tag order, "details-summary", and says he wishes GitHub would add an editor button, both to save typing on a mobile keyboard and to make more users aware that the feature exists.

KEY FACTS

PUBLISHED
9 November 2026

CODING AGENTS NAMED
Claude Code and Codex

WHERE THE TECHNIQUE WORKS
GitHub issue comments and pull-request descriptions

EXAMPLE COLLAPSED LABEL
Claude Opus 5.5 Medium

RECENT EXAMPLES CITED
Five pull requests across datasette, datasette-agent and llm

> WHY IT MATTERS
>
> Teams using coding agents can copy this immediately to make reviews easier to scan without discarding the agents' detailed records. It puts the burden on the author to approve a short opener, rather than expecting every reviewer to excavate the important point.

Read the original at Simon Willison's Weblog →

---

05 LOW IMPACT Simon Willison's Weblog

Claude builds a 3D web demo without a graphics library

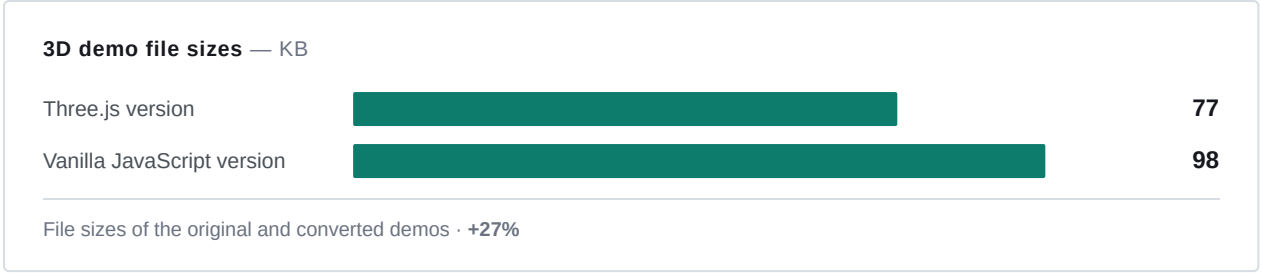
> Simon Willison used Claude Opus 5.5 to turn a Three.js 3D web demo into a version built with plain JavaScript and WebGL.

Willison had made a nudibranch-themed Faberge egg in Three.js, a widely used library that simplifies 3D graphics on the web. In the same Claude Chat conversation that produced the original, he asked Claude Opus 5.5 Medium to port it to vanilla JavaScript. The result used WebGL, the browser's underlying graphics technology, without Three.js.

The finding is a small demonstration rather than a product launch or a measured test of coding ability. It matters most to people building with AI because it shows a chatbot handling a more specialised kind of graphics programming in a familiar conversational workflow. A developer can ask for an existing project to be reworked in a different technical form without starting a separate session or supplying the earlier code again.

The two versions were not identical. Willison linked to the transcript for Claude's account of the differences and said he did not fully understand why they occurred. He also found the Three.js code slightly easier to understand when he inspected it, although he rarely reads the code for these 3D projects and did not consider that a significant drawback.

The file sizes provide a concrete comparison: the original Three.js version was 77KB, while the vanilla JavaScript version was 98KB. The source does not say how long the conversion took, whether the second version performed better, or how the result compares with other models. Nor does it establish that the output is ready for commercial use. The genuinely new element here is Willison's discovery that Claude Opus 5.5 could produce this particular 3D model without the graphics library; the broader claim that AI can write web code is incremental.



KEY FACTS

DEMONSTRATION DATE
9 November 2026

AI TOOL USED
Claude Opus 5.5 Medium in regular Claude Chat

ORIGINAL DEMO
A nudibranch-themed Faberge egg made in Three.js

ORIGINAL FILE SIZE
77KB

CONVERTED FILE SIZE
98KB

RESULT
A non-identical version using vanilla JavaScript and WebGL without Three.js

> WHY IT MATTERS
>
> For developers and teams experimenting with AI coding tools, this is evidence that Claude can attempt lower-level 3D browser work, not just use a convenience library. It is a single demonstration, so it does not justify changing tools or expecting production-ready results without review.

Read the original at Simon Willison’s Weblog →

SECTION 3 OF 3

AI Applications & Industry

6 stories · 3 high · 3 medium

06 MEDIUM IMPACT TechCrunch

Nvidia’s board investigates Huang as he stays in charge

> Nvidia’s board has launched an independent investigation into CEO Jensen Huang following reports that he was targeted by a Chinese spy operation, while expressing confidence in his leadership.

The investigation follows reports that Huang had an affair with Alice Huang, a former executive at Singapore-based chip company Megaspeed. A memo circulating among US policymakers earlier this year claimed she had been recruited by China’s Ministry of State Security to spy on Nvidia’s CEO. In a message to the company, Huang admitted to being unfaithful to his wife and apologised to his family, colleagues and shareholders, saying he had “fallen short” of his responsibility to the company and that they deserved better from him.

The board said it has “no reason to believe” that either national security or company secrets were compromised. It added that Huang “fully supports the review and has committed to cooperating with it,” and that it has “confidence in Jensen and his leadership team.” Independent directors working with outside legal counsel will conduct the investigation. In response to earlier reports about the relationship, Nvidia had said the pair’s interactions were “not unusual” and that the company does not share classified information with customers.

The scrutiny lands as chips, especially Nvidia’s, have become a major focus of the AI race between the United States and China. Huang also appears to have a close relationship with President Donald Trump, who has been a cheerleader for the AI industry. For investors and business customers, the immediate question is governance at a company whose products sit at the centre of AI spending, rather than any disclosed change to its products or operations.

What is new is the board’s formal review and Huang’s public admission, not a finding that spying succeeded or that information leaked. The most striking detail is the alleged method: a claimed intelligence recruitment tied to a personal relationship. Other companies involved in the AI boom have warned employees to be wary of possible spies, including those offering sex or romance, making this a security concern that extends beyond one executive.

KEY FACTS

ANNOUNCED
Sunday: an independent investigation by Nvidia's board

WHO LEADS THE REVIEW
Independent directors working with outside legal counsel

ALLEGATION IN POLICYMAKER MEMO
Alice Huang was recruited by China's Ministry of State Security to spy on Nvidia's CEO

BOARD'S SECURITY ASSESSMENT
"No reason to believe" national security or company secrets were compromised

HUANG'S STATUS
Board has "confidence in Jensen and his leadership team"; he supports and will cooperate with the review

> WHY IT MATTERS
>
> Investors and customers now have to watch governance and security risk at a central AI chip supplier, even though Nvidia says no secrets were compromised. The reported allegations also give AI companies a concrete reason to warn staff about romantic approaches linked to possible espionage.

Read the original at TechCrunch →

---

07 MEDIUM IMPACT reuters.com

Anthropic reportedly buys $3.4 billion of gas-generation capacity

> Anthropic has reportedly bought $3.4 billion of gas-generation capacity for AI data centres.

Source not retrievable. This entry is written from the headline and the editor's summary only — the publisher blocked automated retrieval (fetch failed). Follow the link for the full report.

Anthropic bought commitments for 1.7 gigawatts of heavy-duty gas generators from OnePWR Solutions, the Financial Times reported. The agreement includes engineering and installation help and covers equipment, not an acquisition of OnePWR. Reuters said it could not verify the deal. The reported $3.4 billion purchase would secure gas-generation capacity for AI data centres.

> WHY IT MATTERS
>
> If confirmed, the purchase would show an AI company securing dedicated gas-fired power equipment for data centres, though the report remains unverified by Reuters.

Read the original at reuters.com →

---

08 MEDIUM IMPACT reuters.com

Trump says he 'doesn't mind guardrails' for AI

> President Trump said he was open to guardrails for artificial intelligence, signalling a possible shift from his earlier support for company self-regulation.

Source not retrievable. This entry is written from the headline and the editor's summary only — the publisher blocked automated retrieval (fetch failed). Follow the link for the full report.

President Trump told CBS he was open to working with Democrats on safeguards for artificial intelligence. He warned that rules must not leave the United States behind China. Trump previously favoured company self-regulation. He did not say what level of government oversight he would support, leaving the scope and form of any potential safeguards unclear.

> WHY IT MATTERS
>
> If the openness leads to talks with Democrats, it could shape future US AI oversight, but no specific policy or oversight level has been set out.

Read the original at reuters.com →

---

09 HIGH IMPACT reuters.com

OpenAI receives the first $11 billion of its funding round

> OpenAI has received the first $11 billion of a funding round that values the company at $1.25 trillion, according to Reuters.

Source not retrievable. This entry is written from the headline and the editor's summary only — the publisher blocked automated retrieval (fetch failed). Follow the link for the full report.

SoftBank and MGX joined the first part of the round. OpenAI had initially sought up to $30 billion at a $1.5 trillion valuation, and the round remains open. The initial funding values the company at $1.25 trillion, below that earlier figure. The article itself could not be retrieved, so this entry is based only on the headline and editor's summary; details of the investors' commitments and the round's final size are not available here.

> WHY IT MATTERS
>
> If confirmed, the funding would give OpenAI substantial new capital, while the lower valuation and reduced initial target may signal a more cautious funding environment.

Read the original at reuters.com →

---

10 HIGH IMPACT reuters.com

Chip packaging draws $28 billion as AI demand strains capacity

> Capital spending on advanced chip packaging is expected to approach $28 billion this year as AI demand strains capacity.

Source not retrievable. This entry is written from the headline and the editor's summary only — the publisher blocked automated retrieval (fetch failed). Follow the link for the full report.

Advanced packaging links processors and memory so they can exchange data quickly, making it an important part of meeting AI demand. TSMC's limited capacity is pushing Nvidia and AMD to finance alternative suppliers and new manufacturing techniques. The expected spending shows how investment is spreading beyond the chips themselves to the manufacturing steps needed to connect and package them.

> **WHY IT MATTERS**
>
> If capacity remains tight, packaging constraints could affect how quickly AI chip supply expands and increase the importance of suppliers beyond TSMC.

Read the original at reuters.com →

---

11 HIGH IMPACT reuters.com

Anthropic's November listing may slip as investors demand a discount

> Anthropic's planned November listing could slip as potential investors push for a lower valuation, Reuters reports.

Source not retrievable. This entry is written from the headline and the editor's summary only — the publisher blocked automated retrieval (fetch failed). Follow the link for the full report.

Potential investors are pressing Anthropic for a discount on its valuation, casting doubt on the timing of its planned November listing, according to Reuters. They also want October revenue figures left blank in its filing. A roadshow had been expected to begin this week. The report points to pressure over both the price of the offering and the financial information presented to prospective backers. No revised timetable or valuation is given in the available summary.

> **WHY IT MATTERS**
>
> If the listing is delayed or repriced, it could affect Anthropic's fundraising plans and offer a cautious signal about investor appetite for a major AI company.

Read the original at reuters.com →

Sources

01 Three leading AI labs prepare models for reasoning, visuals and lower prices
https://au.finance.yahoo.com/news/anthropic-google-openai-prepare-ai-144212018.html

02 Cached reasoning cuts repeat AI task costs in early tests
https://arxiv.org/abs/2611.09891

03 Job ads mention AI more, but the work is changing slowly
https://arxiv.org/abs/2611.09870

04 Willison collapses AI-written pull-request essays behind a summary
https://simonwillison.net/2026/Nov/9/details-summary/

05 Claude builds a 3D web demo without a graphics library
https://simonwillison.net/2026/Nov/9/hn-50421091/

06 Nvidia’s board investigates Huang as he stays in charge
https://techcrunch.com/2026/11/08/nvidia-says-its-investigating-ceo-jensen-huangs-relationship-with-possible-chinese-spy/

07 Anthropic reportedly buys $3.4 billion of gas-generation capacity
https://www.reuters.com/business/energy/anthropic-secures-34-billion-generator-deal-ai-data-centers-ft-reports-2026-11-08

08 Trump says he ‘doesn’t mind guardrails’ for AI
https://www.reuters.com/world/trump-says-ai-risks-can-be-managed-signals-openness-guardrails-2026-11-09

09 OpenAI receives the first $11 billion of its funding round
https://www.reuters.com/business/openai-secures-11-billion-funding-125-trillion-valuation-bloomberg-news-reports-2026-11-06

10 Chip packaging draws $28 billion as AI demand strains capacity
https://www.reuters.com/world/asia-pacific/ai-boom-turns-chip-packaging-into-next-battleground-semiconductors-2026-11-06

11 Anthropic’s November listing may slip as investors demand a discount
https://www.reuters.com/world/ai-backlash-could-threaten-anthropic-ipo-timing-potential-investors-say-2026-11-05

About this document. Every story in the 9 November 2026 New Horizon AI Digest, reported at length. Each entry is written from the publisher's own article text; where a source could not be retrieved the entry is explicitly marked and kept short rather than padded.

Images and licensing. Figures are used only where the source licence permits redistribution, and are credited in the caption. Publisher artwork is not reproduced. All titles link to the original publication.

Vocabulary. Common AI terms are defined at new-horizon.tech/glossary/.

Generated 2026-11-09T06:39:27Z · model kimi-k3.1-pro:cloud · New Horizon, new-horizon.tech · To unsubscribe or manage your subscription, see the footer of the daily email.