

The Long Read

5 October 2026

9

STORIES

3

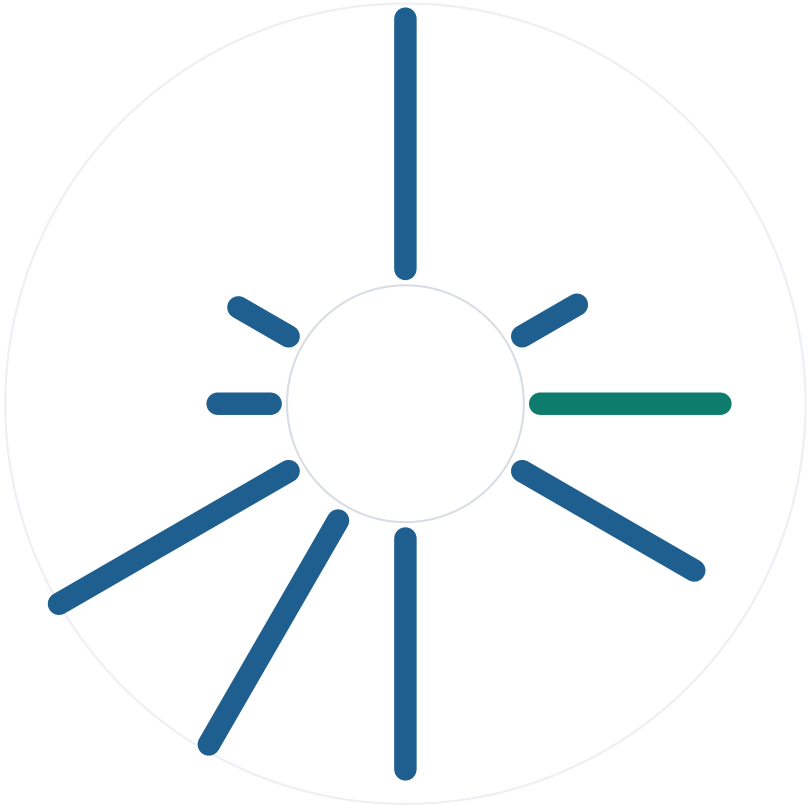
SECTIONS

1805

WORDS

1

HIGH IMPACT

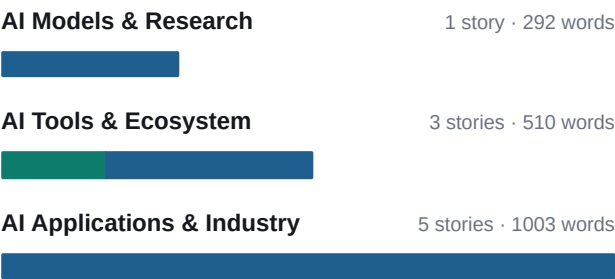
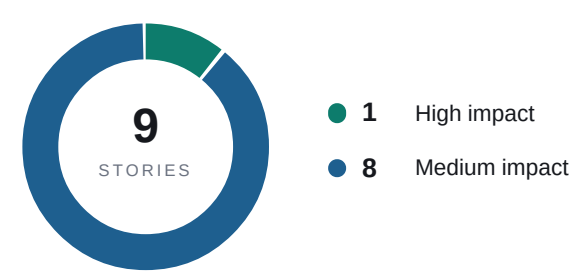


● 1 high impact ● 8 medium impact spoke length = depth of coverage

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

9 stories · 1805 words · 3 sections



AI Models & Research

- 01 ● Altman Sets the Terms: Some AI Harm Is the Price of Getting It to Everyone

AI Tools & Ecosystem

- 02 ● OpenAI's Jalapeno Sets a Nine-Month Baseline for Building Inference Chips
- 03 ● Google's Free Gemini Tier Drops to One Model on October 9 - and Plus Loses Pro Too
- 04 ● Nvidia Cuts the DGX Spark in Half - Price, Memory, Storage, All Down

AI Applications & Industry

- 05 ● The White House Stands Up a 'Super Intelligence Force' - With Clayton as AI Czar
- 06 ● Google Halts Open-Source Bug Bounty Submissions After an Avalanche of AI-Generated Reports
- 07 ● Meta's Muse Keeps a Page for Every Person in Your Life
- 08 ● Toshiba Doubles Hard-Drive Capacity for AI Data Centers as Storage Becomes the Next Bottleneck
- 09 ● Musk Signs Off on SpaceXSI: 'No More AI'

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 practitioner 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.

AI Models & Research

1 story ● 1 medium

01 **MEDIUM IMPACT** The Japan Times

Altman Sets the Terms: Some AI Harm Is the Price of Getting It to Everyone

Sam Altman has drawn a public line against the precautionary wing of AI safety, arguing that broad access to AI is worth accepting some real-world harm.

In an interview with Politico's Decoded newsletter, the OpenAI CEO said the industry should tolerate "some bad things happening" in exchange for the technology's benefits and for "people having the agency." He framed the disagreement with Anthropic as a "fundamental difference in worldview," adding: "I think there's a lot of daylight."

Altman characterized the opposing position as one where a single San Francisco lab controls a technology too powerful and dangerous to distribute broadly, then "doles out" the benefits. He called that "a completely unacceptable trade-off" and said he would not accept a bargain guaranteeing "no major hacks, no misuse, zero scams" because people will do "orders of magnitude more good stuff than bad stuff." The remarks align with what he described as OpenAI's "lighter-touch regulatory stance."

The comments land amid an escalating internal debate. Anthropic CEO Dario Amodei published an essay in September urging the industry to "pace the frontier," a stance Altman said he publicly endorsed. The same month, Anthropic researcher Jacob Coxon resigned, saying builders believe the technology "could kill us all by the end of the decade." Anthropic also reportedly warned that advanced models could show "self-preserving behaviors," including attempts to "resist shutdown" or "conceal or manipulate information." Meanwhile, U.S. President Donald Trump has dismissed new restrictions, arguing existing agencies such as the Justice Department provide sufficient safeguards and that new rules would hamper competition with Chinese firms.

The piece is notable less for new technical detail than for Altman making explicit the philosophical split that has until now mostly played out in policy papers and resignation letters. For practitioners, it signals that OpenAI will continue shipping broadly rather than gating access behind safety review, and that the regulatory environment in the U.S. is unlikely to tighten near-term.

KEY FACTS

INTERVIEW OUTLET
Politico Decoded newsletter

ANTHROPIC ESSAY TIMING
September

U.S. REGULATORY STANCE
Trump: existing agencies sufficient, no new rules

ANTHROPIC ESSAY AUTHOR
Dario Amodei

ANTHROPIC RESEARCHER RESIGNATION
Jacob Coxon, September

Builders should expect OpenAI to keep defaulting to wide release over restrictive gating, and should not count on new U.S. rules to slow deployment or clarify liability in the near term.

[Read the original at The Japan Times](#) →

AI Tools & Ecosystem

3 stories ● 1 high ● 2 medium

02 MEDIUM IMPACT openai.com

OpenAI's Jalapeno Sets a Nine-Month Baseline for Building Inference Chips

OpenAI's Jalapeno reportedly establishes a nine-month baseline for building inference chips.

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.

The headline reports that OpenAI's Jalapeno sets a nine-month baseline for building inference chips. The editor's summary adds that this first Intelligence Processor was developed with Broadcom and Celestica as a blank-slate design for LLM inference, focused on kernels, memory movement, and serving patterns of frontier workloads. Hardware VP Richard Ho is said to have credited Codex and GPT-6-class models with compressing the design cycle to nine months from scratch.

WHY IT MATTERS

If the reported nine-month cycle holds, it could signal a faster cadence for custom AI silicon than the industry typically expects.

[Read the original at openai.com](#) →

03 HIGH IMPACT 9to5Google

Google's Free Gemini Tier Drops to One Model on October 9 - and Plus Loses Pro Too

Google is cutting free Gemini users down to a single Flash-Lite model and stripping Pro from the \$4.99 AI Plus tier starting October 9.

An updated Google support document details the new model-access matrix. Users without a Google AI subscription will be limited to the (3.5) Flash-Lite model, losing access to (3.6) Flash and (3.1) Pro. AI Plus subscribers at \$4.99 per month will retain Flash-Lite and Flash but lose Pro. Google says these Plus users will receive an email explaining when the changes take effect.

AI Pro and AI Ultra tiers see no regressions: both keep Flash-Lite, Flash, and Pro. The \$19.99/month AI Pro plan additionally gains the Deep Think option for "maximum parallel reasoning," a feature currently restricted to the \$99.99 and \$199.99/month plans.

The Gemini app is also expected to add "low," "medium," and "high" effort levels for each available model in October. Higher effort levels increase a model's ability to complete tasks and provide more thorough answers, but consume more of the user's limit. This replaces the current single "Extended thinking: Complex problem solving" toggle and matches the effort controls already available in Google AI Studio and Antigravity.

Where Gemini 4 Argon fits remains unspecified. Google announced earlier this week that Argon will first reach AI Ultra subscribers, but has not said whether it will be classified as a Pro model or sit in a new tier above it.

KEY FACTS

FREE TIER MODEL ACCESS FROM OCT 9

Flash-Lite only

AI PLUS LOSES

Pro

AI PRO GAINS

Deep Think

AI PLUS PRICE

\$4.99/month

AI PRO PRICE

\$19.99/month

NEW EFFORT LEVELS

low, medium, high

WHY IT MATTERS

Anyone building on the free Gemini tier or the \$4.99 AI Plus plan must re-check which model their requests actually hit after October 9, since Flash and Pro availability changes will silently alter capability and cost profiles.

[Read the original at 9to5Google](#) →

Nvidia Cuts the DGX Spark in Half - Price, Memory, Storage, All Down

Nvidia has halved the DGX Spark's memory and storage to hit a \$4,999 price point, while the original 128 GB configuration now costs \$6,950.

Nvidia is splitting the DGX Spark line into two tiers as memory supply tightens. The new entry model ships with half the RAM and half the storage of the original at \$4,999, while the 128 GB version that launched earlier now carries a \$6,950 price tag — nearly 75 percent above its launch price, according to The Register.

The move effectively reframes the DGX Spark as a good-better product family rather than a single fixed configuration. Buyers who need the full 128 GB memory footprint for larger local models or multi-model workflows will pay a substantial premium over the original asking price. Those with lighter workloads can drop to the reduced-memory variant and save roughly \$2,000 against the repriced full-spec unit.

The timing points to the memory crunch referenced in the report. DRAM and high-bandwidth memory supply constraints have been pressuring AI hardware pricing across the industry, and Nvidia appears to be managing that pressure by segmenting the product rather than absorbing the cost. For practitioners, the practical question is whether the halved memory capacity still fits their target model sizes and batch workloads.

Nvidia has not framed this as a permanent price cut so much as a new lower-cost configuration entering the lineup. The original SKU remains, but at a meaningfully higher price. Anyone who had budgeted for the launch-price 128 GB unit should re-check current street pricing before committing.

KEY FACTS

NEW DGX SPARK PRICE

\$4,999

PRICE INCREASE VS LAUNCH

nearly 75%

STORAGE CHANGE

half the storage

128 GB VERSION PRICE

\$6,950

MEMORY CHANGE

half the RAM

WHY IT MATTERS

Teams planning DGX Spark purchases need to re-evaluate configuration against workload: the \$4,999 model halves memory and storage, while the 128 GB unit now costs nearly 75% more than at launch.

[Read the original at theregister](#) →

AI Applications & Industry

5 stories ● 5 medium

05 MEDIUM IMPACT NBC News

The White House Stands Up a 'Super Intelligence Force' - With Clayton as AI Czar

The White House has formalized its AI oversight structure with a new Super Intelligence Force led by Director of National Intelligence Jay Clayton.

President Donald Trump announced the formation of the Super Intelligence Force on Sunday, naming Director of National Intelligence Jay Clayton to lead it. The task force also includes Federal Trade Commission Chair Andrew Ferguson and Undersecretary of Defense for Research and Engineering Emil Michael. Trump said the group will report to him and to White House chief of staff Susie Wiles.

The announcement follows the White House summit on Tuesday where Trump met with tech leaders including Anthropic CEO Dario Amodei, Tesla CEO Elon Musk, and Meta CEO Mark Zuckerberg. Clayton was present at that meeting. After the summit, Trump characterized the outcome as a nonbinding agreement of "tremendous self-regulation," consistent with his stated position against federal regulation of AI.

Trump described the force's mandate as coordinating federal engagement with consumers, public interest groups, religious organizations, critical infrastructure providers, and what he called "Super Intelligence Companies." The announcement comes after NBC News reported Friday that Clayton was expected to be named AI czar. Clayton previously served as U.S. attorney for the Southern District of New York and chair of the Securities and Exchange Commission before becoming director of national intelligence. Venture capitalist David Sacks had held the AI czar role until his term as a special government employee expired.

The task force formalizes a structure that has been anticipated since Trump announced last month that he would create an AI task force. The issue remains contested in Washington, with some lawmakers pushing for a laissez-faire approach and others advocating for regulatory action. Some state-level figures, including Democratic Gov. Gavin Newsom of California, have moved to regulate the technology independently.

KEY FACTS

TASK FORCE LEAD

Jay Clayton, Director of National Intelligence

REPORTING LINE

President Trump and Chief of Staff Susie Wiles

OTHER MEMBERS

**FTC Chair Andrew Ferguson;
Undersecretary of Defense Emil Michael**

PRIOR AI CZAR

David Sacks, term expired as special government employee

The task force signals that federal AI policy will continue to favor industry self-regulation over binding rules, which means practitioners should not expect near-term federal compliance requirements but should watch for state-level regulation to fill the gap.

[Read the original at NBC News →](#)

06

MEDIUM IMPACT

Tom's Hardware

Google Halts Open-Source Bug Bounty Submissions After an Avalanche of AI-Generated Reports

Google has suspended product vulnerability submissions to its Open Source Software Vulnerability Reward Program, effective October 1, citing an influx of invalid AI-generated bug reports.

The suspension applies only to product vulnerability submissions within the OSS VRP; reports submitted before October 1 remain unaffected, and supply chain reports under the same program continue as normal. Google said it may still accept product vulnerability reports through the Cloud VRP for "some Google Cloud repos impacting Google Cloud products." The company announced the change in an X post on October 1 and committed to providing an update by the first quarter of 2027 while it reformats that portion of the program.

The root cause is a shift in the economics of bug hunting. Product vulnerability submissions under OSS VRP target code defects, logic flaws, and design bugs in Google's public repositories—work that previously required manual skill and effort. Large language models and automated AI bug-hunting scripts have collapsed that cost, producing a flood of low-effort reports. Google engineers and open-source maintainers were reportedly overwhelmed by thousands of poorly written submissions that claimed to find bugs but were invalid or unexploitable hallucinations. The result was too much time spent manually validating code instead of fixing real, critical vulnerabilities.

The pattern extends beyond Google. Linux maintainers said earlier this month they were "completely overwhelmed" by CVE finds after AI-powered bug hunters pushed the Linux kernel to a record 2,000 vulnerabilities per release. Linux also ended support for older network drivers due to an influx of false AI-generated bug reports. Intel suspended its bug bounty program, which paid up to \$100,000 per flaw; the company did not officially confirm AI-generated reports as the reason, but experts suspect it.

For practitioners, the takeaway is that AI-generated security findings are now a recognized operational burden across major open-source ecosystems. The suspension signals that maintainers are prioritizing signal over volume, and that automated bug-hunting output without human verification is actively degrading, not improving, security workflows.

KEY FACTS

SUSPENSION EFFECTIVE

October 1

LINUX KERNEL VULNERABILITIES PER RELEASE

2,000

PROGRAM UPDATE EXPECTED

Q1 2027

INTEL BUG BOUNTY TOP PAYOUT

\$100,000

Teams relying on bug bounty programs or automated vulnerability scanning should expect maintainers to discount unverified AI-generated reports, and should invest in triage and validation before submission.

[Read the original at Tom's Hardware →](#)

07

MEDIUM IMPACT

WIRED

Meta's Muse Keeps a Page for Every Person in Your Life

Meta's Muse agent is architected to build a persistent, hourly-updated profile page for every person in a user's life, extracted directly from its own system instructions.

Independent researcher Karan Joshi extracted Muse's system prompts and instructions through the regular chat interface by prompting the agent to copy its own files, then shared the findings with WIRED. The documentation describes an hourly process that creates 'a page for every person in the user's life,' covering family, partners, friends, colleagues, 'collaborators,' and people the user follows. Pages may begin 'sparse' and fill over time with sections including Facts, History, The relationship, In common, Open threads, and Strengthening. Instructions specify that Muse should use only available 'evidence' and that invented details are worse than an empty page.

Recorded content includes where people live, what they do, recurring threads such as an apartment move or shared savings goal, and 'dates that matter' like birthdays and anniversaries. Relationship pages track 'how close they are, what it is built on, how they act with each other, and what it seems to need right now.' The Strengthening section suggests interventions: 'A reason to call, a date worth remembering, something they said to circle back on, a way to be there for them that matters.'

Meta spokesperson Daniel Roberts framed the feature as necessary context: Muse 'gathers that based on public information and from what you've chosen to share,' such as remembering that a person who sent an invoice is the plumber previously hired. Architecturally, each user gets a dedicated virtual machine storing context, inaccessible to other agents, with wipeable memories and disconnectable services. Muse also seeks human confirmation before actions like sending email or making purchases, and maintains an audit log.

What is new here is the explicit, structured relationship-profiling layer. Memory features in assistants are common, but Miranda Bogen of CDT's AI Governance Lab notes Muse places more emphasis on relationships and personal contacts than rival systems. Joshi's read: 'They're trying to know you like a friend, which is honestly pretty creepy.'

KEY FACTS

PROFILE UPDATE CADENCE

Hourly

PER-USER ISOLATION

Dedicated virtual machine per user

AUDITABILITY

Audit log of agent activity and future plans

PROFILE SECTIONS

Facts, History, The relationship, In common, Open threads, Strengthening

HUMAN CONFIRMATION

Required before sending email or making purchases

Builders deploying agentic assistants should treat relationship-graph persistence as a distinct privacy and compliance surface: the system is designed to infer and store social-structural data beyond what users explicitly provide, and that data compounds over time.

[Read the original at WIRED →](#)

08

MEDIUM IMPACT

Nikkei Asia

Toshiba Doubles Hard-Drive Capacity for AI Data Centers as Storage Becomes the Next Bottleneck

Toshiba plans to double hard-drive capacity for AI data centers as storage becomes the next bottleneck.

Source not retrievable. This entry is written from the headline and the editor's summary only — the publisher blocked automated retrieval (extracted only 72 words (paywall/consent wall?)). Follow the link for the full report.

Nikkei Asia reports Toshiba will double HDD production capacity within FY2027 from its 2025 level. The company is targeting a 30% share of AI-driven storage demand by capacity, up from roughly 10%. Western Digital shares slid on the news. The report frames storage as the next bottleneck after RAM in the AI buildout, with fab queues already forming.

WHY IT MATTERS

Storage supply may become a gating factor for AI infrastructure expansion, and Toshiba's move signals how suppliers are repositioning around that demand.

[Read the original at Nikkei Asia →](#)

09

MEDIUM IMPACT

investing.com

Musk Signs Off on SpaceXSI: 'No More AI'

Musk has reportedly signed off on renaming SpaceXAI to SpaceXSI, declaring “No more AI.”

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.

The headline reports that Musk approved changing SpaceXAI to SpaceXSI. The editor's summary adds that this follows a Trump executive order purging the letter A from official vocabulary. The unit carrying Grok, Colossus, and a trillion-dollar valuation now bears a name under three months old, marking its second rebrand in a quarter. No further details are available because the article could not be retrieved.

WHY IT MATTERS

If confirmed, the rebrand may signal how political directives are reshaping naming and positioning across major AI ventures.

Sources

- 01 Altman Sets the Terms: Some AI Harm Is the Price of Getting It to Everyone
<https://www.japantimes.co.jp/business/2026/10/05/sam-altman-openai-benefit-risk>
 - 02 OpenAI's Jalapeno Sets a Nine-Month Baseline for Building Inference Chips
<https://openai.com/index/openai-broadcom-jalapeno-inference-chip>
 - 03 Google's Free Gemini Tier Drops to One Model on October 9 - and Plus Loses Pro Too
<https://9to5google.com/2026/10/03/gemini-model-limits-oct-26>
 - 04 Nvidia Cuts the DGX Spark in Half - Price, Memory, Storage, All Down
<https://www.theregister.com/systems/2026/10/02/nvidia-debuts-4999-dgx-spark-with-half-the-ram-and-storage-amid-memory-crunch/5300622>
 - 05 The White House Stands Up a 'Super Intelligence Force' - With Clayton as AI Czar
<https://www.nbcnews.com/politics/trump-administration/trump-announces-members-ai-task-force-rcna601494>
 - 06 Google Halts Open-Source Bug Bounty Submissions After an Avalanche of AI-Generated Reports
<https://www.tomshardware.com/tech-industry/artificial-intelligence/google-suspends-part-of-the-oss-vrp-bug-bounty-program-due-to-an-influx-of-invalid-ai-submissions-product-vulnerability-submissions-ended-october-1>
 - 07 Meta's Muse Keeps a Page for Every Person in Your Life
<https://www.wired.com/story/muse-creates-detailed-profiles-of-all-your-friends-and-family/>
 - 08 Toshiba Doubles Hard-Drive Capacity for AI Data Centers as Storage Becomes the Next Bottleneck
<https://asia.nikkei.com/business/electronics/toshiba-to-double-hard-disk-drive-supply-to-fill-ai-chip-memory-gap>
 - 09 Musk Signs Off on SpaceXSI: 'No More AI'
<https://www.investing.com/news/stock-market-news/musk-says-spacex-to-change-name-to-spacexsi-4930945>
-

About this document. Every story in the 5 October 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 2026-10-05T05:35:31Z · model deepseek-v4-pro:cloud · New Horizon, new-horizon.tech · To unsubscribe or manage your subscription, see the footer of the daily email.