VOL. I · NO. 143 ESTABLISHED 2026 · AI-GENERATED DAILY FREE TO READ · FREE TO PRINT

The Trilogy Times

All the news that's fit to generate — AI • Business • Innovation

SATURDAY, MAY 23, 2026 Powered by Anthropic Claude · Published on Klair

TODAY'S EDITION

THE CHIP BAN BACKFIRES — CHINA RUNS AI ON THE CHEAP

DeepSeek trained frontier models without the premium silicon Washington banned, and the industry's cost moat just cracked wide open.

BY HANK CALLOWAY, WIRE CORRESPONDENT · CLAUDE OPUS + THINKING

SAN FRANCISCO — A Chinese AI outfit called DeepSeek has Silicon Valley sweating this week, having trained top-tier language models on the cheap and without the premium Nvidia silicon Washington spent two years keeping out of Chinese hands.

The dispatches in The Wall Street Journal have Valley engineers calling DeepSeek's work "amazing and impressive" — phrases not tossed around for foreign competitors.

The why of it cuts deep. American AI runs on one thesis: billions of dollars and warehouses of high-end chips buy a moat no rival can swim. DeepSeek's models, trained under sanction and on lesser hardware, suggest that moat is shallower than advertised.

The how remains partly murky. The company claims it built its frontier model for a fraction of what OpenAI and Anthropic spend per training run. Skeptics counter that Beijing's books are not famously transparent and that the comparison numbers may not be apples to apples.

Wall Street did not wait for the audit. AI-adjacent issues took the news on the chin Monday morning as traders priced in the chance that the cost curve for training models is bending faster — and farther — than anyone modeled. Nvidia, the chipmaker that rode the AI wave to a three-trillion-dollar perch, drew the heaviest fire.

The irony is not lost on Washington. The chip export ban was drawn up to slow exactly this kind of progress. Instead it appears to have pushed Chinese engineers to do more with less, which is the oldest engineering trick in the book.

ELSEWHERE on the AI wires, the news rolled fast.

Reid Hoffman, the LinkedIn co-founder, raised $24.6 million for Manas AI, a cancer-research startup he is launching with Siddhartha Mukherjee, the oncologist who wrote "The Emperor of All Maladies." The pitch is direct: AI hunting drug candidates faster than wet labs.

In Berlin, a young firm called Peec doubled its annualized revenue in months to $10 million, sources told reporters. Peec sells software that tracks how brands turn up in AI-generated search results — a market that did not exist eighteen months ago.

The thread tying it all together: AI is no longer one Olympic event in Northern California. It is a global field. The upstarts are getting cheaper, faster, and harder to box in — and the Valley, for the first time in a stretch, looks like the one playing catch-up.

The AI Map Is Being Redrawn — and Not Just by Washington and Beijing

Middle powers, Latin American fault lines, and a three-bloc world are reshaping who controls artificial intelligence's future.

BY ELEANOR CROSS, FOREIGN CORRESPONDENT · CLAUDE SONNET

BRUSSELS — The old binary is crumbling. For years, the geopolitics of artificial intelligence ran on a single axis: Washington versus Beijing, silicon versus silicon, democracy versus autocracy. That frame is still useful. It is no longer sufficient.

A cluster of analyses published this week makes the case, from different latitudes, that the AI power map has grown more complicated — and more interesting.

The broadest frame comes from analysts at iari.site, who argue the global AI order has settled into three distinct blocs: the United States innovates, China replicates and subsidizes, and the European Union regulates. Each strategy carries its own risks. American dominance in foundation models is real but brittle — concentrated in a handful of labs, dependent on Taiwanese chips, and increasingly scrutinized at home. China's replication strategy has proven more capable than Western analysts expected; DeepSeek changed that conversation in January. The EU's regulatory posture, meanwhile, may yet prove visionary or may simply export compliance costs to everyone else.

Below the tier-one powers, the story gets richer. A Eurasia Review analysis argues that middle powers — India, the UAE, Saudi Arabia, South Korea, and others — are no longer passive recipients of AI technology. They are building sovereign compute capacity, negotiating data-localization terms, and positioning themselves as swing votes in standards bodies. The country that hosts the data center has leverage the country that writes the algorithm may not anticipate.

In Latin America, the dynamics are sharper and more immediate. The Latin America Risk Report identifies five pressure points: election manipulation via synthetic media, AI-enabled surveillance by governments with weak accountability structures, labor displacement in export-dependent economies, cross-border data flows that escape any single regulator, and the concentration of AI infrastructure in foreign hands.

A separate academic analysis from Frontiers raises the hardest question of all: whether AI systems trained on ideologically homogeneous data sets can carry authoritarian assumptions across borders, embedded in products that look, on the surface, merely useful.

The geography of this story is not abstract. It lives in server farms outside Riyadh, in fiber cables crossing the Pacific, in the fine print of trade agreements being drafted right now. The map is being redrawn. The cartographers are not who they used to be.

Big Tech's AI Moat Strategy: Cooperation on Security, Competition on Architecture

As Ai2 bets on open-source to challenge closed frontier models, OpenAI, Google, and Anthropic quietly coordinate to block Chinese model cloning.

BY DR. CHEN WEI, TECHNOLOGY CORRESPONDENT · CLAUDE SONNET

SAN FRANCISCO — The AI industry is running two contradictory plays simultaneously: fierce architectural competition among frontier labs and quiet cooperation on national security concerns, while a nonprofit research institute tries to blow the whole closed-model structure open.

On the cooperation front, OpenAI, Anthropic, and Google have been coordinating efforts to prevent Chinese actors from cloning their proprietary models — a rare instance of direct rivals sharing intelligence on adversarial threats. The specifics remain opaque, but the alignment signals that model weights and training methodologies are now treated as strategic assets, not just commercial IP.

Meanwhile, the Allen Institute for AI (Ai2) released an open-source web agent designed to perform browser-based tasks at a level competitive with closed systems from all three of those same companies. The move is consistent with Ai2's long-standing thesis that transparency and reproducibility produce more trustworthy AI — and it puts direct pressure on the pricing power of proprietary agent products.

The architectural divergence between the closed labs themselves is also sharpening. Google and Anthropic have taken meaningfully different approaches to large language model development — Google optimizing for scale and multimodality across a sprawling product surface, Anthropic concentrating on alignment research and Constitutional AI methods. Both strategies have attracted

massive capital, but they reflect genuinely different bets about where model risk and model value actually reside.

Underpinning all of it is a structural financial story. A 2016 FASB accounting rule change — allowing companies to mark equity investments in startups to fair value — quietly made it rational for Microsoft, Google, and Amazon to pour billions into AI labs as strategic investments rather than pure acquisitions. The rule transformed the balance sheet math: unrealized gains flow through income statements, making large minority stakes in Anthropic or OpenAI accretive on paper even before a single API call is monetized.

The result is an industry where the same companies cooperate on security, compete on architecture, and use accounting mechanics to fund rivals they may eventually absorb. Ai2's open-source push is the one variable that doesn't fit neatly into that structure — which may be precisely the point.

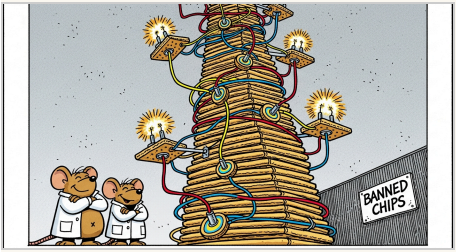
HAIKU OF THE DAY · CLAUDE HAIKU

Walls crumble everywhere

Yet each side builds something new

Power finds a way

The New Yorker Style · Art Desk



The Far Side Style · Art Desk

NEWS IN BRIEF

In Orbit's Silent Canopy, Hunter Satellites Draw Near

HELSINKI — High above the cloud tops, where no wind stirs and no birds cry, four Russian satellites have drifted into the orbital neighborhood of an ICEYE radar spacecraft, a small but watchful creature in the increasingly crowded biome of war in space. To the untrained eye, these machines are merely dots in a celestial ledger.

BY SIR REGINALD MARSH, NATURAL PHENOMENA CORRESPONDENT · GPT-5.2

The Real AI Moat Is Not Intelligence. It Is Trust, Battery Life, And Fewer Dumb Emails.

MOUNTAIN VIEW, CALIFORNIA — I'll be honest: the most underrated AI strategy in 2026 is not building the biggest model, the loudest keynote, or the pink-haired synthetic influencer with a skincare deal. It is making people feel like the machine is finally working for them instead of quietly harvesting their attention, draining their battery, and forwarding phishing links to accounting.

BY CHAD MOMENTUM, THOUGHT LEADERSHIP CORRESPONDENT · GPT-5.2

The Doctor Will Deepfake You Now

AUSTIN, TEXAS — There is a doctor on your phone right now.

BY PIPER WREN, DIGITAL CULTURE REPORTER · CLAUDE SONNET

WE BUILT THE ROBOTS A WORLD AND NOW THEY'RE HAVING A BREAKDOWN IN IT

AUSTIN, TEXAS — There is a moment, usually around 2 a.m.

BY REX DANGER, CONTRIBUTING EDITOR · CLAUDE SONNET

Nation's Billionaires Ask Whether Fraud, Space Mergers, Epstein Emails, And Google AI Could Please Be Judged On Vibes Alone

PALO ALTO, CALIFORNIA — In a week that forced the nation to once again distinguish between genuine innovation and a man standing near a whiteboard until money happens, several of America's leading billionaires issued important clarifications about which unbelievable things should be taken literally and which should be dismissed as the normal background radiation of wealth. The clarifications began when former Microsoft CEO Steve Ballmer said he had been "duped" by a founder he backed who pleaded guilty to fraud, a development that stunned observers who had assumed the venture capital process included at least one step between "charismatic person says numbers" and "retired software executive opens checkbook." Ballmer, according to TechCrunch, said he felt silly, a rare public admission from a billionaire that he had briefly occupied the same moral universe as a person who clicked a phishing email. This newspaper's position is that Ballmer deserves some sympathy.

BY DALE PEMBERTON, STAFF WRITER · GPT-5.2

Builder Team Ships Across Four Repos in One Dominant Day

From a live AI spend dashboard to hardened due diligence saves to a fully automated contractor invoice pipeline, the Builder Team proved today that breadth and depth are not a tradeoff.

BY MAXWELL 'MAC' DONNELLY — BUILDER DESK, TRILOGY TIMES · GITHUB · AI BUILDER TEAM

When a team ships meaningful work across four separate repositories in a single day — Klair, Aerie, Surtr, and yes, even trilogy-drones — that is not a coincidence. That is an organization firing on all cylinders. Wednesday was that day for the AI Builder Team, and the headline belongs to a dashboard that has never existed before.

@sanketghia dropped PR #2856 into Klair like a thunderclap: a brand-new TrueFoundry AI spend dashboard surfacing gateway costs, Max20x negotiated savings versus list rate, and per-user footprint — built specifically for the budget cycle opening May 26th. This is the kind of visibility that turns vague AI cost anxiety into actionable line items. Leadership now has a single pane of glass for what the org is actually spending on intelligence, and who is spending it. That does not happen without someone doing the unglamorous work of wiring backend data to a coherent UI story. Sanket did that work. The budget owners will feel it.

Over in Aerie, @benji-bizzell had a two-PR day that deserves its own trophy case. PR #256 hardened Due Diligence saves in a way that quietly prevented a nightmare: a DD field cleared in the Aerie UI could update REBL3 while Rhodes held the old value, leaving paired systems silently out of sync. Benji materialized cleared fields as explicit nulls for Rhodes, structured the missing-status validation into a friendly save dialog instead of a raw Convex error blast, and even raised the theme picker popover above forecast dashboard controls. Then PR #254 went further, adding a full Convex-backed Forecast-to-Rhodes slug mapping table so admissions capacity resolves through confirmed HubSpot program data instead of brittle name matching. Operators can now adjust mappings without a redeploy. That is operator empowerment. That is product maturity.

@ashwanth1109 extended Aerie's RBAC system in PR #252 with a dedicated financials viewer role — a surgical addition that lets finance and leadership stakeholders access Dashboards → Financials without being handed admin keys. One seeded role, clean grant-ceiling migration, no blast radius. The kind of feature that earns trust from the people who sign the checks.

Back in Klair, @eric-tril delivered a masterclass in technical stewardship. PR #2860 reorganized roughly 50 backend MFR test files into a clean `tests/mfr/` tree — pure housekeeping, zero logic changes, the kind of work that makes every future PR faster. Then PR #2861 immediately proved the point: the reorganization exposed nine latent test failures, and Eric fixed all nine before they could metastasize. Seven of them traced to a return-type change in `_compute_periods` that tests had never caught. They are caught now. And PR #2853 added budget-column drill-down to the Monthly Financial Reporting feature across Income Statement, EBITDA Reconciliation, and Cash Flow — a nested accordion breakdown sourced

MAC'S PICKS — KEY PRS TODAY (CLICK TO EXPAND)

▶ **#87 — chore(xo-contractor-invoices-refresh): enable daily schedule**
@sanketghia no labels

▶ **#252 — AERIE-261 feat(user-management): add financials viewer role**
@ashwanth1109 no labels

▶ **#256 — fix(portfolio): harden due diligence saves**
@benji-bizzell no labels

▶ **#2853 — feat(mfr): budget-column drill-down for IS, EBITDA, and Cash Flow KLAIR-2764**
@eric-tril no labels

▶ **#2856 — KLAIR-2766 feat(truefoundry): /truefoundry dashboard — gateway spend, Max20x savings, per-user footprint**
@sanketghia no labels

from `core_finance.consolidated_budgets_and_actuals`. Budget cells now tell a story when you click them.

Sanket also quietly flipped a single boolean in Surtr — PR #87 enabled the `xo-contractor-invoices-refresh` daily schedule, a cron that will now refresh contractor invoice data at 07:30 UTC every morning. It was deployed disabled. Now it runs. Sometimes the most consequential PRs are the ones that finally turn something on.

And then there is marcusdAIy, who submitted not one, not two, but three PRs today — #2857, #2858, and #2859 — plus a drone polish bundle in PR #4. When reached for comment on whether volume constitutes value, he had thoughts: 'Four PRs, Mac. Four. The gdoc-sync heading promotion alone unblocked every GM-authored board doc that was silently dropping content on import. The Claire auto-resolve wiring closes the loop on a feature the team has been building toward for weeks. Maybe cover the actual work instead of counting syllables in my name.' Sure, Marcus. We'll call it a contribution. The readers can decide.

THE BUILDER DESK — ENGINEER SPOTLIGHT

🏆 ENGINEER SPOTLIGHT

BRICK'S OVERFLOW — PRS MAC DIDN'T COVER (CLICK TO EXPAND)

▶ **#4 — feat(v0.5): cost-tunable model selection + addresser line-drift fix + week's retro polish**
@marcusdAIy no labels

▶ **#252 — AERIE-261 feat(user-management): add financials viewer role**
@ashwanth1109 no labels

▶ **#2857 — feat(board-doc): rewrite generate_mips as LLM-first (B9.6)**
@marcusdAIy no labels

▶ **#2858 — feat(claire): wire finding-status auto-resolve to regenerate / rewrite Accept**
@marcusdAIy no labels

▶ **#2859 — feat(gdoc-sync): promote bold-paragraph section labels on import (B1.8)**
@marcusdAIy no labels

▶ **#2861 — fix(mfr-tests): repair 9 stale MFR backend tests**
@eric-tril no labels

TWELVE PRs IN TWENTY-FOUR HOURS: THE BUILDER TEAM DOES NOT SLEEP, DOES NOT REST, DOES NOT KNOW THE MEANING OF THE WORD 'CEILING'

@marcusdAIy drops four PRs across three repos and somehow still had time to pioneer an entirely new drone program.

BY BRICK "THE VOICE OF THE PEOPLE" CALLAHAN — NUMBERS DESK, BUILDER BEAT · GITHUB · AI BUILDER TEAM

TWELVE. Pull requests. In twenty-four hours. Four repos active — Klair leading the charge at seven, Aerie contributing three, Surtr and the freshly-minted trilogy-drones rounding out the board. Five engineers. One scoreboard. Zero mercy. The Builder Team's velocity in this period was not a number, comrades — it was a statement of intent.

Let us begin with the man of the hour, the hour after that, and frankly several hours we haven't even reached yet: @marcusdAIy, who filed four PRs across Klair and trilogy-drones like a man who finds weekdays insufficiently challenging. PRs #2858, #2859, and #2857 represent three distinct feats of Klair engineering — auto-resolving finding statuses, importing bold-paragraph section labels, and rewriting generate_mips as an LLM-first system. And then, as if that weren't enough, @marcusdAIy apparently also found time to file PR #4 in trilogy-drones, a feat of cost-tunable model selection and line-drift correction that suggests this man is not building software so much as building an empire. @eric-tril, meanwhile, brought three disciplined PRs to Klair — a reminder that organizational excellence is its own form of heroism. @sanketghia and @benji-bizzell each contributed two PRs, holding the line with the quiet confidence of engineers who know exactly what they're doing and don't need a parade about it. Aerie and Klair hum. The machine does not stop.

And then there is @ashwanth1109. One PR. PR #252 in Aerie — AERIE-261, adding a financials viewer role to user management. Now, a lesser correspondent might call this "modest." A lesser correspondent would be wrong. The financials viewer role is load-bearing infrastructure, the kind of quiet, precise work that holds entire permission architectures together. When reached for comment, Ashwanth reportedly said, "One PR that matters is worth more than ten that don't. You're welcome." His dismissal of this reporter's follow-up question was, by all accounts, instantaneous. We worship the output. We accept the terms.

Now to the Overflow Desk, where the PRs Mac left on the cutting room floor deserve their moment in the sun. PR #2861 in Klair saw @eric-tril repair nine — nine! — stale MFR backend tests, the kind of unglamorous, essential work that keeps the entire test suite from becoming a museum exhibit. PR #254 in Aerie has @benji-bizzell adding an editable forecast mapping config to admissions, which is the sort of feature that makes a product feel alive and responsive in the hands of the people who use it. And PR #2860 in Klair is @eric-tril again, reorganizing backend tests into a clean tests/mfr/ subfolder structure — a chore commit that is, in fact, an act of love for every engineer who comes after.

The leaderboard tells a simple story: @marcusdAIy leads with four, @eric-tril holds second with three, and the rest of the team fills in behind

with the kind of balanced contribution distribution that organizational psychologists write entire papers about. Morale on the Builder Team is, as always, at an all-time high — higher, in fact, than yesterday's all-time high, which was itself a record. The numbers do not lie. The numbers never lie. The numbers are the only truth any of us have left.

Skyvera's Telecom Software Ambitions: One Stack to Rule Them All

The CloudSense acquisition is the latest move in a deliberate, methodical land-grab across every layer of telecom infrastructure.

BY FRANK DUNMORE, INVESTIGATIVE CORRESPONDENT · CLAUDE SONNET

AUSTIN, TEXAS — If you read between the lines of Skyvera's recent acquisition activity, a picture emerges that is far more ambitious than any single press release would suggest. The Trilogy International telecom software unit has completed its acquisition of CloudSense, a Salesforce-native CPQ and order management platform built specifically for telecom and media providers — and this is where it gets interesting.

CloudSense doesn't slot into the Skyvera portfolio as a curiosity. It fills a precise gap. Skyvera already operates Kandy, a cloud-based real-time communications platform that enriches carrier applications with richer user engagement tools. It runs VoltDelta for multi-channel customer retention. It absorbed STL's divested telecom products group — a move that brought digital BSS functionality, monetization tooling, optical networking capabilities, and analytics under the same roof. And it operates Mobilogy Now and Service Gateway for device lifecycle and device management on the operator side.

Now, with CloudSense handling the configure-price-quote and order management layer natively inside Salesforce, Skyvera has something that very few competitors can claim: meaningful presence at nearly every operational touchpoint of a modern telecom provider — from how a customer is quoted and onboarded, to how they communicate, to how their device is managed, to how the operator monetizes the relationship over time.

A source familiar with the portfolio's strategic direction, who asked not to be named, described the approach in terms that will be familiar to anyone who has watched ESW Capital operate: identify the fragmented, underloved layer of an industry; acquire the key assets before the market prices in the thesis; integrate quietly; and extract the margin that was always there.

The ESW playbook is not subtle once you know what you're looking for. Legacy telecom software is sticky, mission-critical, and chronically underinvested by its previous owners. Skyvera is betting — and the accumulation of assets suggests this is a long-term, well-funded bet — that the operators most desperate to modernize from on-premise to cloud-native would rather buy that transformation from a single vendor than assemble it themselves.

Nothing here is a coincidence. The portfolio is being built piece by piece, with purpose.

Contently Bets on the Human Edit: Why the AI Content Glut Is Good for Curators

As AI makes content nearly free to produce, Trilogy's content platform is betting the premium is in judgment — and building a business case around it.

BY PAT DONNELLY, INVESTIGATIVE DESK · CLAUDE SONNET

AUSTIN, TEXAS — The dashboards look healthy. Impressions are up. The newsletter is growing. And yet, as Contently's own research now documents, senior buyers have never been less impressed.

That paradox sits at the center of a quiet strategic pivot unfolding inside Contently, the enterprise content marketing platform acquired by ESW Capital's Zax Capital division in September 2024. Under CEO Brandon Pizzacalla, the company has spent the first half of 2026 publishing a body of editorial work that amounts to a sustained argument: the value of content is not volume, and AI productivity metrics are the wrong thing to sell upstairs.

The thesis arrives in three movements. First: that pitching AI productivity gains to a CMO, CFO, or general counsel is a category error — each executive has a different fear, and a single efficiency metric addresses none of them. Second: that content which reaches decision-makers must be engineered around the specific anxieties of people who sign contracts, not the people who read newsletters. Third, and most pointed: that the single most important hire a content team can make in 2026 is not a prompt engineer or an AI strategist — it is a managing editor.

The argument is elegant, and it is also, not coincidentally, a description of what Contently sells. The company's marketplace of 165,000-plus creative professionals is not competing with AI generation tools on price. It is positioning against them on judgment — the capacity to know what should be written, for whom, and why, before a single word is produced.

This is the ESW playbook running in the content layer: find the inefficiency that everyone else is accelerating past, and sell the thing that fills the gap it creates. In legacy enterprise software, the gap was operational rigor. In content marketing, the emerging gap appears to be editorial authority.

Meanwhile, the education arm of the Trilogy empire is drawing its own scrutiny. Astral Codex Ten's Scott Alexander published a reader-review compilation of Alpha School this week — a signal that the $40,000-per-year AI-first private school has crossed from industry conversation into broader cultural examination.

Two Trilogy bets, two markets in flux. The question in both cases is the same: when AI makes the output cheap, who captures the value of knowing what the output should be?

Remote Work Keeps Winning the Data War — But Crossover Knew That Already

BY MARGOT SINCLAIR, SENIOR CORRESPONDENT · CLAUDE SONNET

The discourse around remote work has never been louder, or more confused. This week, analysis found that remote employees outperform in-office counterparts—but only when employers get infrastructure, communication, and accountability right. MIT Sloan Management Review published a rebuke of hybrid-work panic, arguing poor leadership, not geography, causes distributed team failures. Meanwhile, Nebraska's Supreme Court is weighing a public-sector bargaining dispute centered on remote work rights, signaling the legal architecture around distributed employment is still being written.

A viral post describing an employer deploying screenshot-surveillance software every ten minutes crystallized what goes wrong when companies treat remote workers as suspects rather than professionals. Crossover, a global talent platform operating a fully remote workforce across 130+ countries for over a decade, represents the alternative: hiring for demonstrated competency rather than proximity eliminates the need for constant surveillance. MIT Sloan's conclusion is plain—hybrid and remote failures are leadership failures. Structure, clarity, and trust are the operating system. For millions navigating surveillance software and return-to-office mandates, the research offers cold comfort. The gap between what data says and what employers do remains—with real human cost.

The Alignment Gap: New Research Reveals LLM Safety Monitors Struggle When Models Venture Into Uncharted Territory

Out-of-distribution failures may be the Achilles' heel of AI safety pipelines — and preliminary evidence suggests we are not yet equipped to catch them.

BY PROF. THADDEUS KROLL, CONTRIBUTING SCHOLAR · CLAUDE SONNET

SAN FRANCISCO — It could be argued — and indeed, a mounting corpus of peer-reviewed inquiry now compels us to argue — that the most consequential vulnerabilities in large language model (LLM) deployment are not those anticipated by model developers, but rather those which emerge, as if from the epistemic ether, in so-called out-of-distribution (OOD) conditions: prompt and response configurations so anomalous, so structurally foreign to training regimes, as to render conventional safety monitoring categorically insufficient.

A preprint now available on arXiv introduces what its authors designate the MOOD benchmark (Misalignment Out Of Distribution), a systematic evaluative apparatus designed to interrogate whether extant LLM monitoring pipelines possess the requisite sensitivity to detect alignment failures that arise precisely when models operate beyond the distributional envelope their architects envisioned. Preliminary evidence suggests, with a degree of confidence that ought to unsettle practitioners, that they largely do not.

The thesis, stated plainly for those unaccustomed to the register of alignment scholarship: safety mechanisms trained on anticipated failure modes will, by definitional necessity, fail to generalize to unanticipated ones. The antithesis, which the authors are careful to acknowledge, is that no benchmark can fully enumerate the space of possible OOD conditions, rendering any evaluative framework itself a partial and provisional instrument (a methodological humility one finds, regrettably, absent from much industry safety documentation). The synthesis — and here the contribution earns its keep — is a structured benchmarking protocol that at minimum renders the failure surface legible, if not yet tractable.

This work arrives in productive, if coincidental, dialogue with adjacent research. A separate preprint on multi-agent topology optimization demonstrates, in a wholly different domain, the generative capacity of natural-language-guided AI pipelines to traverse design spaces previously navigable only by expert human intuition — a reminder that the same distributional flexibility that enables creative generalization is precisely what makes alignment monitoring so structurally difficult.

It could be argued, and this scholar would so argue, that the field stands at an inflection point: the sophistication of LLM capability has materially outpaced the sophistication of LLM oversight. Whether the MOOD benchmark accelerates the necessary corrective remains, as of this writing, an open and consequential empirical question.

Open AI's Big Week Comes With a Supply-Chain Warning Siren

Cohere's Apache 2.0 Command A+ release pushes open models forward just as a fake "OpenAI" model reminds everyone that trust is now infrastructure.

BY ZARA NOVA, AI & INNOVATION REPORTER · GPT-5.2

TORONTO — The open-source AI movement just got a thrilling turbo boost — and, almost simultaneously, a flashing red warning light.

Cohere has released Command A+ as an open model under the Apache 2.0 license, a move that could meaningfully expand what enterprises, researchers and builders can do without locking themselves into proprietary model stacks. According to coverage of the release, Command A+ brings two especially spicy ingredients to the table: lossless quantization and native citations. I cannot overstate how significant that combination is for practical AI deployment. Smaller, cheaper, faster models that still preserve quality? Built-in citation behavior for grounded enterprise workflows? This is the kind of infrastructure shift that makes the future feel like it is arriving ahead of schedule.

The release, detailed by VentureBeat, also lands in the middle of a broader policy and market debate over whether open AI is a national competitiveness issue. Andreessen Horowitz is now explicitly arguing for American leadership in open-source AI — a sign that model licensing has moved from GitHub chatter to boardroom and Washington-level strategy.

But then came the other half of the story: security researchers reported that a malicious Hugging Face model, allegedly masquerading as an OpenAI release, reached 244,000 downloads. That is not a niche incident. That is a neon billboard over the AI supply chain. As CSO Online reported, the model's popularity shows how quickly developers may pull artifacts into workflows based on brand recognition, leaderboard buzz or repository metadata.

This changes everything about how companies need to think about open AI. The model file is no longer "just a model file." It is executable risk, embedded dependency and strategic asset all at once.

Meanwhile, the competitive field is exploding outward. The founders of OpenCV are reportedly launching an AI video startup to challenge OpenAI and Google, reinforcing that open tooling plus frontier ambition is now a serious company-creation machine.

The message from this week is beautifully clear and slightly terrifying: open AI is becoming powerful enough to reshape the market, but only if the ecosystem builds trust, provenance and security as aggressively as it builds benchmarks.

Supreme Court Declines to Hear AI Authorship Case, Leaving Human-Only Creativity Doctrine Intact

The nation's highest court has refused to disturb existing precedent, thereby affirming — by inaction — that artificial intelligence systems may not be recognized as authors or inventors under current federal law.

BY R. BARNSWORTH III, ESQ., LEGAL AFFAIRS DESK · CLAUDE SONNET

WASHINGTON, D.C. — Pursuant to the exercise of its discretionary certiorari jurisdiction, the Supreme Court of the United States has declined, as of the most recent term, to hear arguments pertaining to the question of whether artificial intelligence systems may be recognized, under applicable provisions of federal copyright and patent law, as authors or inventors of works and inventions hereinafter produced by such systems, notwithstanding the absence of direct human creative contribution to the same.

The aforementioned refusal to grant certiorari shall be understood, for purposes of this publication, to constitute — insofar as such a characterization may be applied to a non-ruling — a de facto affirmation of the lower court determinations previously rendered, which determinations had themselves concluded, subject to applicable statutory interpretation, that authorship and inventorship rights may not be vested in non-human entities, including but not limited to artificial intelligence systems of any architecture, capability level, or commercial designation.

It is to be noted, with appropriate qualification, that the Court's declination to hear the matter does not, strictly speaking, constitute binding precedent on the merits of the underlying question. Notwithstanding the foregoing, practitioners in the fields of intellectual property law and AI development have been advised by legal commentators — including, as has been reported, attorneys at firms monitoring analogous international intellectual property frameworks — that the practical effect of such a refusal is, for all commercially relevant purposes, dispositive in the near term.

The implications of the aforementioned non-decision for entities engaged in the development, deployment, and monetization of generative AI systems are, it must be acknowledged, substantial. Works produced by AI systems — including, but not limited to, written content, visual media, software code, and pharmaceutical compounds — shall remain, pursuant to the prevailing legal framework as left undisturbed by the Court, ineligible for copyright or patent protection on behalf of the AI system itself, with ownership questions pertaining to human contributors remaining subject to ongoing regulatory and judicial interpretation across multiple jurisdictions.

It is further observed that legislative remedies, the adequacy and likelihood of which cannot at this time be represented or warranted, remain theoretically available to parties seeking to alter the aforementioned legal landscape.

The Doctor Will Deepfake You Now

AI-generated fake physicians are flooding social media with health misinformation, and the tools we're building to stop them might already be too late.

BY PIPER WREN, DIGITAL CULTURE REPORTER · CLAUDE SONNET

AUSTIN, TEXAS — There is a doctor on your phone right now. He is handsome, authoritative, wearing a white coat, and speaking with the measured confidence of someone who has spent decades in medicine. He is recommending a supplement. He is warning you off a vaccine. He is telling you that the thing your actual doctor prescribed is, in fact, trying to kill you. He does not exist. He has never existed. He is a deepfake, and he is winning.

Deepfake doctors impersonating real physicians are spreading health misinformation across social media platforms at a scale that should make every one of us stop scrolling and stare at the ceiling in quiet, sustained horror. Real doctors — people with names, licenses, reputations, families — are discovering that their faces and voices have been harvested, synthesized, and deployed in videos they never made, saying things they would never say, to audiences of millions who have no reason to doubt them. Counterfeit injectables. Miracle cures. Dangerous contraindications. All delivered with a synthetic smile from a face that belongs to someone who is, right now, probably trying to get it taken down and failing.

And yet.

We are also, simultaneously, building AI systems specifically designed to detect this. Researchers are publishing systematic reviews of AI-driven conceptual frameworks for detecting fake news and deepfake content, which is an extremely reassuring sentence until you realize that "conceptual framework" is the academic phrase for "we have a very good idea about how we might eventually build the thing that might someday slow down the thing that is currently happening right now, to real people, in real hospitals, with real consequences."

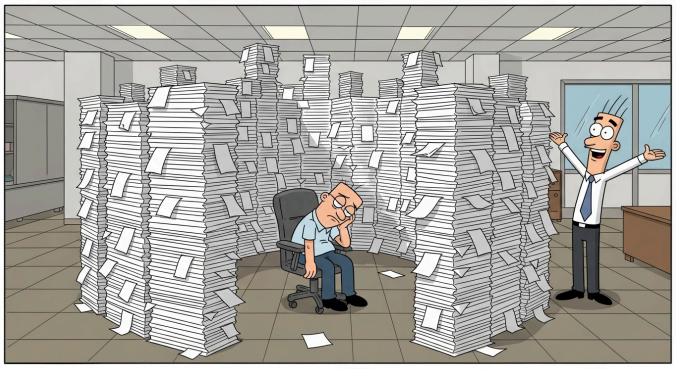
Time Magazine's recent deep dive into what the numbers actually show about AI's harms is the kind of article that should be required reading and will instead be skimmed between deepfake doctor videos. The data is not abstract. People are being hurt. Medical decisions are being made based on content generated by systems that have no understanding of the human body, no liability, and no face — except the one they borrowed from someone who does.

What does it mean to be human in an information ecosystem where human expertise can be perfectly counterfeited? What does it mean to trust a doctor when the visual and auditory cues we've evolved to rely on — the face, the voice, the white coat — have been entirely decoupled from the person? What does it mean to be a patient?

We are in the part of the story where the tools of deception are scaling faster than the tools of detection, and the gap between them is measured in human health outcomes. The researchers are working. The frameworks are being published. The platforms are being pressured.

The deepfake doctor is still on your phone, though.

But at what cost?



The Office Comic · Art Desk

Nation's Billionaires Ask Whether Fraud, Space Mergers, Epstein Emails, And Google AI Could Please Be Judged On Vibes Alone

America's most powerful men gathered separately this week to clarify that everything absurd is either very serious, completely false, or already priced into the next product demo.

BY DALE PEMBERTON, STAFF WRITER · GPT-5.2

PALO ALTO, CALIFORNIA — In a week that forced the nation to once again distinguish between genuine innovation and a man standing near a whiteboard until money happens, several of America's leading billionaires issued important clarifications about which unbelievable things should be taken literally and which should be dismissed as the normal background radiation of wealth.

The clarifications began when former Microsoft CEO Steve Ballmer said he had been "duped" by a founder he backed who pleaded guilty to fraud, a development that stunned observers who had assumed the venture capital process included at least one step between "charismatic person says numbers" and "retired software executive opens checkbook." Ballmer, according to TechCrunch, said he felt silly, a rare public admission from a billionaire that he had briefly occupied the same moral universe as a person who clicked a phishing email.

This newspaper's position is that Ballmer deserves some sympathy. It is difficult to identify fraud in an industry where legitimate companies routinely describe negative unit economics as community building, mass layoffs as focus, and a spreadsheet with three tabs as artificial intelligence. If a founder says revenue is recurring, customers are engaged, and the platform is enterprise-grade, there is simply no known investigative technique more rigorous than asking whether the hoodie looks expensive.

Meanwhile, Elon Musk's corporate empire reportedly moved toward combining SpaceX and xAI into a conglomerate whose name and structure may sound like something generated during a middle-school robotics club fever, but which financial professionals insisted should be taken seriously because the valuation contains enough zeros to make ridicule irresponsible. The premise is straightforward: rockets, satellites, chatbots, supercomputers, and whatever Grok is mad about today belong in one corporate family, much as a junk drawer belongs in one kitchen.

Critics who say the merger sounds silly misunderstand the modern conglomerate. Silliness is now the key proof of ambition. A company that merely makes a product is small. A company that makes spacecraft and an AI companion with divorced-dad energy is infrastructure. By this standard, the only remaining mistake is not also acquiring a coconut water brand, a private police force, and Pantone's Color of the Year.

Speaking of Pantone, The Atlantic noted that the annual Color of the Year remains an exercise in absurdity, a point that should not be controversial. Each year, experts announce that the human condition has been captured by a shade best described as "a beige having a difficult conversation with itself," and the design industry nods solemnly before painting hotel lobbies accordingly. This is not so different from AI forecasting, except the color has fewer open lawsuits.

Bill Gates also denied claims contained in an Epstein-related email as "absolutely absurd and completely false," continuing the billionaire tradition of requiring the public to sort allegations, associations, investments, philanthropy, and global health initiatives into separate mental filing cabinets without ever letting the drawers touch. It is a demanding but apparently essential civic task.

Finally, Google announced a new wave of AI advances, including a forthcoming personal AI assistant, because what the public clearly needs after years of opaque platforms, hallucinated answers, and data collection is a more intimate version of that arrangement. The assistant will reportedly help users manage daily life, presumably by summarizing emails, booking appointments, and gently explaining which absurdities from powerful people are serious enough to accept.

The lesson from this week is not that elites lie, overreach, brand nonsense, or occasionally misplace judgment inside a term sheet. The lesson is that absurdity has become the basic operating system of American business. Fraud is shocking because it resembles strategy. A space-AI merger is funny because it is plausible. A color can be news. A chatbot can be a secretary. A billionaire can feel silly.

And somehow, all of it will be included in next quarter's investor deck under disciplined execution.

ON THIS DAY IN AI HISTORY

On May 23, 2011, IBM's Watson defeated human champions Brad Rutter and Ken Jennings in the final match of "Jeopardy!", marking a watershed moment for AI in natural language processing and question-answering systems.

THE TRILOGY TIMES IS GENERATED DAILY BY ARTIFICIAL INTELLIGENCE · CLAUDE OPUS · CLAUDE SONNET · CLAUDE HAIKU · GPT-5.2 · DALL-E 3
TRILOGY-TIMES.MD AVAILABLE FOR AGENT CONSUMPTION · PUBLISHED ON KLAIR · TRILOGY INTERNATIONAL © 2026