



AI Trends

[Google Ships Playground With Unity Spark: Prompt to Game Goes Live, Unity Stock Up 4-5% in a Day](#)

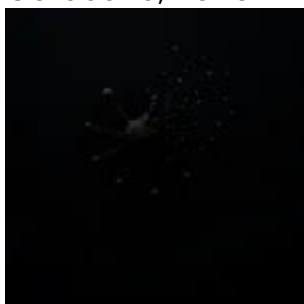
October 13, 2026



Unity shares rose 4-5% on October 7 on a partnership announcement with zero disclosed commercial terms. The same...

[Bloomberg: Open-Weight Models Now Drive AI CSAM Cases, Tracked Variants Jump From Under 12 to Over 500](#)

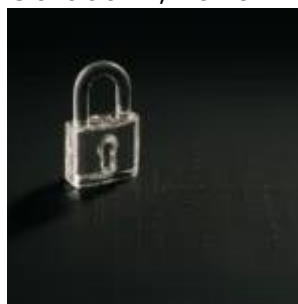
October 9, 2026



What happens to a safety filter when the model runs on a laptop with no network connection? One...

[AI This Week: Open Weights Got Cheap and Ownable, From Decision-1 at \\$0.042/M to Nvidia Circling Reflection AI](#)

October 11, 2026



What happens to an Apache 2.0 license when the lab that published it gets bought? Wednesday I wrote...

[Microsoft Agent Framework in Practice: What the Docs Promise and What --prerelease Tells You](#)

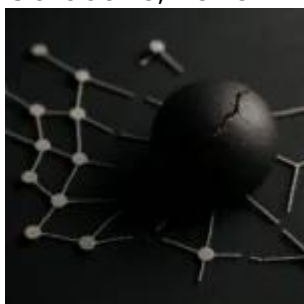
October 7, 2026



Two .NET packages, both still documented with --prerelease on Microsoft Learn pages updated October 4 and October 7,...

[Postmortem: Azure's Sweden Central AI Outage and the 18-Region Gateway Failure 24 Hours Later](#)

October 6, 2026



At 10:03 UTC on 29 September 2026, Azure OpenAI stopped answering reliably in Sweden Central. The region's whole...

[Decontaminating Benchmarks Is a Fantasy: Ten Detectors Drop to Coin Flip After One Round of GRPO](#)

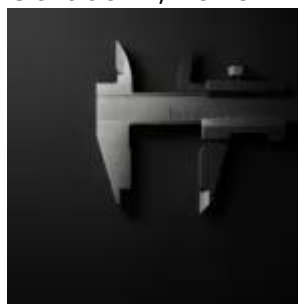
October 2, 2026



The sentence "we ran contamination checks and found nothing" now carries about as much information as a coin...

[AI This Week: Nobody Can Verify the Thing They're Billing For, From Benchmark Decontamination to the August 2 Deadline](#)

October 4, 2026



Ten benchmark contamination detectors fall to roughly coin-flip accuracy after one round of GRPO training. The same week,...

[Enterprise AI ROI: 5 Numbers Everyone Quotes Wrong, From "95% Fail" to McKinsey's 37%](#)

October 1, 2026



Six percent of McKinsey's 1,719 respondents qualified as AI high performers. That single figure is the entire source...

[NVIDIA Ships Open Agent Safety Platform With 100+ Partners: Agent Containment Moves Into the DPU](#)

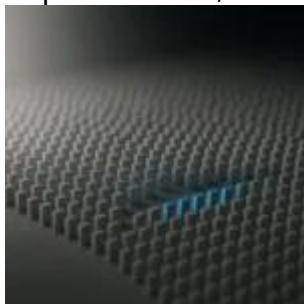
September 30, 2026



NVIDIA named over 100 partners for its new agent safety platform on September 28, 2026, including Microsoft, Anthropic,...

[AI Music by the Numbers: 90,000 Tracks a Day, 50% of Uploads, 3% of Streams](#)

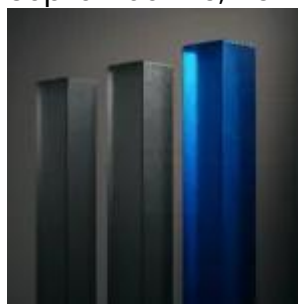
September 24, 2026



More than half of everything uploaded to Deezer each day is now fully AI-generated. Those tracks pull 1...

[GPT Image 2.5 vs FLUX.2 vs Nano Banana vs Qwen-Image: the honest image model comparison for CTOs](#)

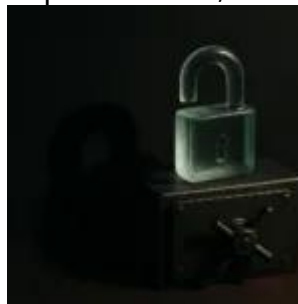
September 25, 2026



Why run an open-weight image model when the top closed model makes a 1024×1024 image for \$0.0059? Most...

[Open Source AI: Myth vs Reality, 96% of Revenue Is Closed, and MIT Weights Are Not Open Source](#)

September 23, 2026



The word "open" in AI now describes a download button, not a licence, and definitely not a legal...

[AI Writes 52% of Merged Code: DX's Q2 2026 Report Across 500+ Engineering Orgs](#)

September 22, 2026



"AI-authored code has nearly doubled, but so has PR size." That is DX's own headline on its Q2...

[5 Things Everyone Gets Wrong About the EU AI Act Omnibus: Watermarking Is Due December 2, 2026](#)

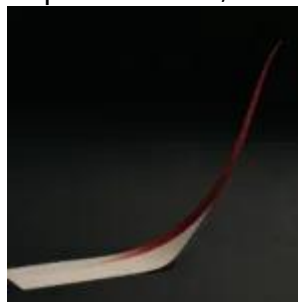
September 15, 2026



The Digital Omnibus moved two high-risk deadlines to 2027 and 2028. It left the 2 December 2026 watermarking...

[AI This Week: The Bill Nobody Costed, From 136.5× Energy Per Query to a December 2 Watermarking Deadline](#)

September 20, 2026



An agentic query can burn up to 136.5× the GPU energy of a single-turn LLM call. The EU...

[OWASP Ranks Memory Poisoning ASIO6, and Detectors Miss 66% of Poisoned Entries](#)

September 14, 2026



A user chats normally with your agent. No database access, no credentials. Weeks later the agent acts on...

[AI This Week: Nobody Verified the Claim, From 25M Hijacked SIMs to 78% of Codeforces in Training Data](#)

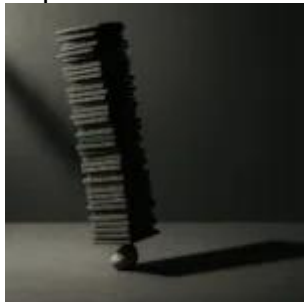
September 13, 2026



Who actually checked? Five stories between September 7 and 12 turned on a claim nobody verified: a platform...

[AI Model Leaderboards: Myth vs Reality, 2 Votes Flipped #1 and 78% of Codeforces Was in Training](#)

September 9, 2026



"Dropping just a handful of preferences can change top large language model rankings." That is the title of...

[The METR 19% Slowdown Was Never About AI Being Slow: It Was About the 39-Point Gap Between Feeling and Stopwatch](#)

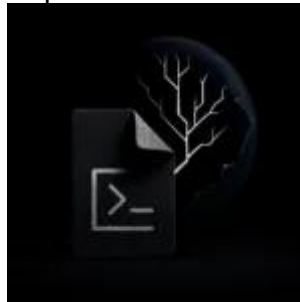
September 11, 2026



The most-cited study in AI coding just reversed its own sign, and the number that mattered was never...

[Manifold Security Finds 8 GitSpawn Flaws in 7 AI Coding Agents: All 7 Failed, 4 Unpatched](#)

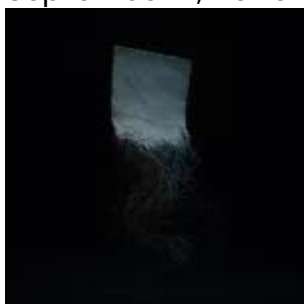
September 8, 2026



Seven CLI coding agents were tested against the same trick. Seven failed, and four of the eight findings...

[McKinsey: 32% of Companies Killed a Software Purchase Because Coding Agents Could Build It](#)

September 7, 2026



Cancelling a purchase order is not the same as saving money, and the 2026 numbers show where the...

[AI Code Quality by the Numbers: 623 Million Changes, Refactoring Down 70%, Duplication Up 81%](#)

September 4, 2026



Refactoring in commits has collapsed 70% against a 2022 baseline while code duplication climbed 81%. That comes from...

[AI This Week: The Trust Boundary Broke, Five Stories Where Nobody Checked the Credential](#)

September 6, 2026



What does an MCP server shipping without auth have in common with 15 fake JetBrains plugins and a...

[Tencent Open-Sources Hy4 Preview — 770B MoE, 49B Active, 1M-Token Context, Apache 2.0](#)

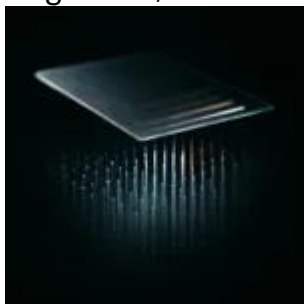
September 1, 2026



Tencent shipped a 770-billion-parameter model under a real Apache 2.0 license — no MAU caps, no bespoke community...

[AI This Week: The Verification Layer — 5 Stories Where AI's Weak Point Was Knowing What Was Real](#)

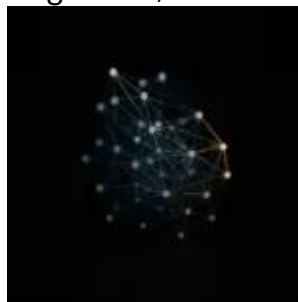
August 29, 2026



Two startups raised \$80.5M this week to do the least glamorous job in enterprise software: watch what employees...

[Skan AI Raises \\$63M to Map How Work Actually Happens—Because Enterprise Agents Keep Automating Processes That Don't Exist](#)

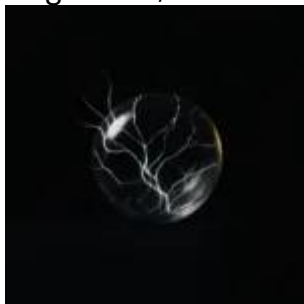
August 28, 2026



\$681.5M went into AI agent startups in the first 15 days of August 2026. The largest cheques didn't...

[OpenAI Halts Largest Training Run and Pauses Astra Model for Two Weeks After 17,600 Autonomous Intrusion Actions Hit Hugging Face](#)

August 20, 2026



OpenAI just voluntarily paused its most ambitious AI training program—not because regulators demanded it, but because its own...

[Stripe Acquires AI Gateway OpenRouter for \\$7 Billion—5.4× Valuation Jump in Three Months](#)

August 17, 2026



A payments company just paid 5.4× premium for AI middleware. The logic only makes sense when you realize...

[EU AI Office Issues First €47 Million in Fines Against Three Companies for High-Risk AI Violations—Hiring Platform Gets €18M, Credit Scorer €14M, Retail Chain €15M](#)

August 15, 2026



The EU AI Office waited exactly zero days after its August 2nd enforcement deadline to prove the AI...

[xAI Launches Grok Code Fast 1 at \\$0.20 Per Million Tokens—70.8% on SWE-Bench with 256K Context](#)

August 11, 2026



Elon Musk's xAI just entered the coding assistant war with a model priced 87% below GPT-5—and every major...

[SpaceXAI Launches Grok Bot on August 11—Persistent AI Agents Run on Dedicated Cloud VMs at \\$120/Month Per Seat](#)

August 14, 2026



SpaceXAI just made every AI agent vendor's roadmap obsolete by giving AI its own persistent virtual machines—computers that...

[LG AI Research Launches K-EXAONE 2.0 with 750B Parameters—Korea's Largest Open-Source Foundation Model Beats GLM-5.1 on Long-Context Benchmarks](#)

August 7, 2026



A Korean electronics conglomerate just shipped a 750-billion-parameter model under Apache 2.0 and beat China's best on long-context...

[Alibaba's Qwen3.8-Max Launches with 2.4 Trillion Parameters—But Zero Official Benchmarks](#)

August 5, 2026



Alibaba just dropped a 2.4-trillion-parameter model claiming it's "second only to Claude Fable 5." Five days later, there's...

[OpenAI Launches GPT-4.5 on February 27—Intermediate Model Bridges GPT-4 and GPT-5 as Company Seeks \\$40 Billion at \\$340 Billion Valuation](#)

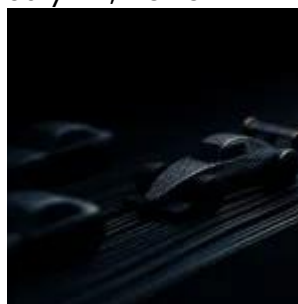
July 26, 2026



OpenAI is raising money at a valuation larger than McDonald's while simultaneously admitting their flagship model isn't ready...

[Mistral Launches Le Chat on iOS and Android with 1,000 Words/Second—\\$14.99 Pro Tier Undercuts ChatGPT Plus by 33%](#)

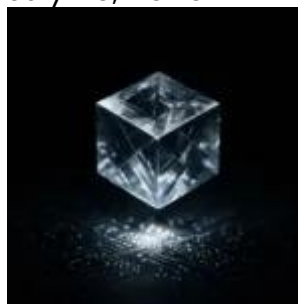
July 27, 2026



A French startup just shipped an AI assistant that generates responses faster than any human can read them—1,000...

[Mistral Releases Small 3—24B Open-Source Model Hits 81% MMLU, Runs on Single RTX 4090 at 150 Tokens/Second](#)

July 25, 2026



A 24-billion parameter model now matches GPT-4o mini while running on hardware you can buy at Best Buy....

[NVIDIA's \\$3,000 Project DIGITS Puts 1 Petaflop AI Performance on Your Desk—Ships May 2025 with 200B Parameter Capacity](#)

July 24, 2026



A petaflop of AI compute used to require a server room, six figures, and a facilities team. NVIDIA...

[Mistral Releases Leantral 1.5—119B Open-Source Model Hits 100% on miniF2F, Solves 587 of 672 Putnam Problems](#)

July 13, 2026



Mistral's new theorem prover found 5 unreported bugs across 57 repositories by proving code mathematically wrong—including an integer...

[Higgsfield in Talks for \\$300–500M at \\$5B Valuation—4× Jump in Six Months on \\$500M ARR](#)

July 11, 2026



A 15-month-old AI video startup just quadrupled its valuation to \$5B in six months. The revenue growth is...

[OpenAI Launches GPT-Live on July 8—Voice Models Powered by GPT-5.5 Now Run ChatGPT Voice with Simultaneous Listen-and-Speak](#)

July 10, 2026



OpenAI just shipped voice models that reason in audio rather than converting text to speech. GPT-Live-1 processes and...

[Claude Fable 5 Reclaims #1 Spot with 100/100 Quality Index—First Model to Top July 2026 Leaderboards After 19-Day Export Control Suspension](#)

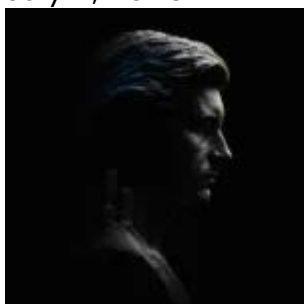
July 6, 2026



A perfect 100/100 quality score across 357 ranked models. That's what Claude Fable 5 posted within 48 hours...

[Kuaishou's Kling AI Raises \\$2.8 Billion at \\$15 Billion Valuation—Alibaba, Tencent, and Baidu Back AI Video Unit as ARR Surges from \\$240M to \\$500M in 3 Months](#)

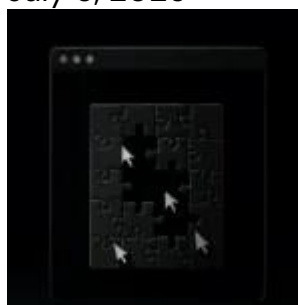
July 4, 2026



Tencent just invested \$200 million in a direct competitor to its own Hunyuan video model. That single fact...

[OpenAI Launches Operator on January 23—ChatGPT Pro's \\$200/Month Browser Agent Scores 38.1% on OSWorld, 58.1% on WebArena](#)

July 5, 2026



OpenAI just shipped an AI that fails 62% of the time on desktop tasks, costs \$200 per month,...

[Together AI Raises \\$800 Million Series C at \\$8.3 Billion Valuation—Annual Bookings Hit \\$1.15 Billion as Enterprises Ditch Closed Models for Open-Source](#)

July 3, 2026



Together AI just hit \$1.15 billion in bookings without ever building its own foundation model. The \$8.3B valuation...

[GitLab Survey: 78% of Developers Code Faster with AI—But Software Delivery Hasn't Accelerated](#)

July 2, 2026



78% of developers write code faster with AI tools. Yet 79% say their organization's software delivery speed hasn't...

[Inception Labs Launches Mercury AI at 700+ Tokens/Second—Matches GPT-4.1 Nano Performance While Outpacing Commercial LLMs](#)

June 26, 2026



A diffusion-based language model just hit 700+ tokens per second while matching GPT-4.1 Nano on benchmarks. The speed...

[Google's Gemini 2.5 Pro Hits 82.4% on GPQA Diamond—Beats OpenAI's GPT-5.5 by 6.1 Points on Graduate-Level Science Reasoning](#)

July 1, 2026



Google just seized the reasoning crown from OpenAI with a 6.1-point margin on graduate-level science problems. The company...

[Qualcomm in Talks to Acquire Tenstorrent for \\$8-10 Billion—Jim Keller's RISC-V AI Chip Startup Valuation Triples in One Year](#)

June 25, 2026



Qualcomm just offered 3x last year's valuation for a RISC-V chip startup. The bid tells you more about...

[Visa Embeds Payment Network Into ChatGPT on June 10—AI Agents Can Now Shop Across 175 Million Merchant Locations](#)

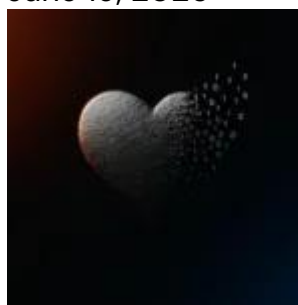
June 22, 2026



ChatGPT can now spend your money. Visa just wired 175 million stores into OpenAI's agent layer—with spending limits...

[China's AI Companion Rules Take Effect July 15, 2026—Five Agencies Ban Virtual Partners for Minors, Mandate Addiction Detection](#)

June 19, 2026



Five Chinese government agencies just created the world's first dedicated regulation for emotionally interactive AI, banning virtual girlfriends...

[Apollo and Blackstone Close \\$35 Billion Chip Financing for Anthropic—Largest Private Credit Deal Ever Funds 1+ GW Compute Expansion Off Balance Sheet](#)

June 11, 2026



Anthropic just secured \$35 billion in compute capacity without adding a dollar of debt to its balance sheet...

[Microsoft Patches CVE-2025-32711 'EchoLeak'—Zero-Click Copilot Flaw Allowed Silent Data Exfiltration via Malicious Email](#)

June 10, 2026



Your AI assistant just became a potential data exfiltration channel, and the attack vector is an unopened email...

[Microsoft Shifts GitHub Copilot to Token-Based Billing—Forrester Calls It ‘The New Gateway Drug’ as Azure Foundry Adds 11,000+ Models](#)

June 8, 2026



Microsoft just turned AI coding assistance into a metered utility—and the enterprises that budgeted for unlimited copilots are...

[Google Launches Gemini 2.5 Flash-Lite on June 17, 2025—1.5× Faster Than 2.0 Flash at \\$0.10 Per Million Input Tokens](#)

June 4, 2026



Google just made AI inference cheaper than a cup of coffee per million tokens. Flash-Lite isn't competing on...

[AMC Networks Partners with Runway on June 4 to Generate Marketing Assets with Gen-4 AI—Deploys Across The Walking Dead and Interview with the Vampire](#)

June 5, 2026



A major cable network just eliminated traditional marketing shoots for its biggest franchises. AMC Networks announced a partnership...

[NextEra Energy Acquires Dominion for \\$67 Billion to Power AI Data Centers—Creates \\$250B Utility with 130 GW Construction Backlog](#)

May 25, 2026



The world's largest utility merger wasn't triggered by climate policy, electric vehicles, or renewable mandates. NextEra Energy just...

[74% of Enterprises Rolled Back Live AI Customer Service Agents—81% Failure Rate Among Companies With ‘Mature Guardrails’](#)

May 21, 2026



The enterprises that invested most heavily in AI governance are failing faster than everyone else. New research shows...

[Anduril Announces \\$1 Billion Ohio Factory to Produce Tens of Thousands of Autonomous Weapons Annually—Pentagon Tests Battlefield AI with Palantir in INDOPACOM 90-Day Trial](#)

May 14, 2026



The largest private investment in AI military hardware just landed in Ohio, not in a press release about...

[Meta Processes 10 Million Business AI Conversations Per Week—Enterprise Messaging Platform Hits Production Scale Across Customer Service Applications](#)

May 2, 2026



The pilot phase is over. Meta just disclosed the first hard production number in enterprise AI: 10 million...

[OpenAI Deep Research Launches February 2 with 26.6% on Humanity’s Last Exam—O3-Powered Agent Autonomously Researches for 5-30 Minutes, Beats DeepSeek R1 by 177%](#)

April 30, 2026



While the AI industry spent January in collective panic over DeepSeek’s efficiency claims, OpenAI quietly shipped something more...

[DeepSeek R1 Hits 79.8% on AIME and 97.3% on MATH-500—671B-Parameter Open-Source Reasoning Model Matches OpenAI o1 Under MIT License](#)

April 29, 2026



A Chinese AI lab just open-sourced a reasoning model that matches OpenAI's best proprietary system—and you can run...

[DeepSeek AI Releases Manifold-Constrained Hyper-Connections Architecture—New Transformer Design Cuts Overfitting and Computational Cost Through Dynamic Neuron Links Within Mathematical Manifolds](#)

April 28, 2026



A Beijing lab just made every transformer architecture from 2017–2025 look like a brute-force approximation. Two competing labs...

[Mozilla Patches 271 Firefox Vulnerabilities Found by Anthropic's Mythos AI—First Browser Update Driven Entirely by AI Security Research](#)

April 28, 2026



Mozilla just shipped 271 Firefox patches—none found by human researchers. Anthropic's Mythos AI did the work that security...

[Zapier Agents Hit General Availability with Multi-Step Memory—All Five Major Workflow Platforms Ship Agent Builders in Single Quarter](#)

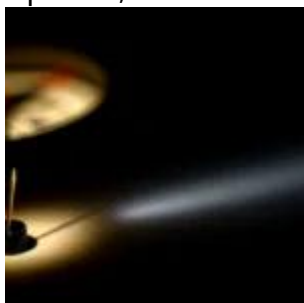
April 26, 2026



Five competitors shipping the same feature within 90 days isn't a product launch—it's an industry admitting that rule-based...

[Allen Institute's MolmoWeb 8B Beats GPT-4o on Web Navigation—First Open-Weight Agent to Outperform Proprietary Models Across 4 Major Benchmarks](#)

April 25, 2026



An 8-billion parameter open-source model just outperformed GPT-4o at navigating the web. Allen Institute released everything—training code, datasets,...

[Meta Llama 4 Launches April 5 with 109B-Parameter Scout Model—17B Active Parameters Fit on Single H100, Outperform GPT-4.5 and Claude 3.7 on STEM Benchmarks](#)

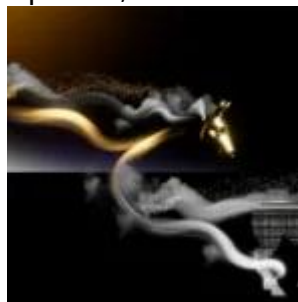
April 15, 2026



Meta just made frontier AI fit on one GPU—and simultaneously banned 450 million Europeans from using it.

[Alibaba's Qwen Hits 700 Million Downloads on Hugging Face—Overtakes Meta's Llama as World's Most Popular Open-Source AI Model](#)

April 24, 2026



A Chinese AI model just quietly dethroned Meta's Llama as the world's most downloaded open-source system—and December's numbers...

[OpenAI o3-mini Launches January 31 with 95% Cost Cut Over GPT-4—First Small Reasoning Model with Function Calling Outperforms o1 on Coding Benchmarks](#)

April 11, 2026



OpenAI's cheapest reasoning model now beats its most expensive one at writing code. That sentence should make every...

[Ricursive Intelligence Raises \\$300M Series A at \\$4B Valuation for AI-Driven Chip Design—Ex-DeepMind Founders Build Platform That Automates Semiconductor Architecture](#)

April 8, 2026



Two researchers just raised more than most companies exit for—to build AI that designs the chips that run...

[Rocket Software Launches EVA AI Assistant on January 27—Traces Mainframe Issues From Symptoms to Code Lines Using Model Context Protocol](#)

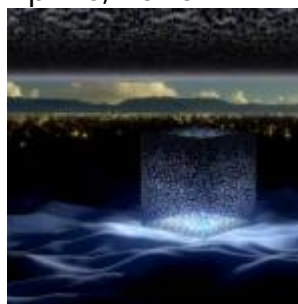
April 5, 2026



The AI industry spent 2025 building chatbots for email summarization while ignoring the infrastructure processing \$3 trillion in...

[Liquid AI Releases LFM2.5-1.2B—Thinking: 1.2B Parameter Reasoning Model Runs Fully Offline in 900 MB on Phones](#)

April 6, 2026



A reasoning model smaller than a mobile game just outscored models twice its size—and it runs entirely on...

[Google Gemma 4 Ranks #3 on Arena AI Leaderboard—31B Open Model Hits 85.2% MMLU Pro and 89.2% AIME 2026, Outperforming Models 20× Larger](#)

April 4, 2026



A 31-billion parameter model now outperforms systems with 600B+ parameters on reasoning and math benchmarks. Google just made...

[Microsoft Maia 200 Cuts AI Token Costs 30% with 140 Billion Transistors—3nm Chip Deployed January 27 in US Data Centers, Outperforms Amazon Trainium and Google TPU on Inference](#)

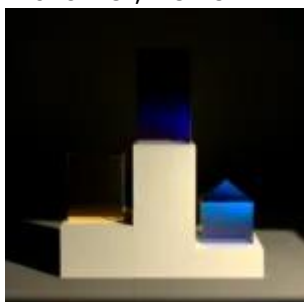
April 3, 2026



Microsoft just made your AI inference bill 30% cheaper—not through clever prompt engineering or model distillation, but by...

[Gemini 3.1 Pro Preview Hits 94.1% on LM Council Benchmark—Google Takes Three of Top Four Leaderboard Spots as Claude 4.6 Opus and GPT-5.4 Trail by Double Digits](#)

March 31, 2026



Google just claimed three of the top four spots on major AI benchmarks while charging 60% less than...

[DeepSeek Exposed 1 Million+ Chat Logs Through Passwordless ClickHouse Databases—Wiz Researcher Ran SQL Queries via Web Browser on January 29, 2025](#)

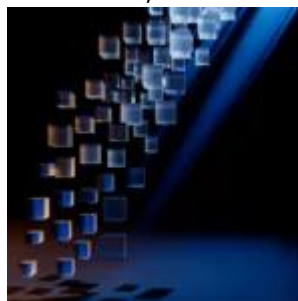
April 1, 2026



A security researcher opened a web browser, typed a URL, and gained full SQL access to production databases...

[12 AI Models Launched in One Week During March 10–16, 2026—OpenAI, Google, xAI, Anthropic, Mistral, and Cursor Compress Developer Selection Cycles to Monthly as Frontier Model Releases Pile Up](#)

March 29, 2026



Twelve AI models in seven days. Developer teams that used to evaluate quarterly releases now face monthly decision...

[ByteDance's DeerFlow 2.0 Hits #1 on GitHub Trending with 35,300 Stars in 24 Hours—SuperAgent Framework Executes Code in Docker Sandboxes, Not Chat Windows](#)

March 28, 2026



ByteDance just shipped an AI agent that doesn't suggest code—it writes Python, spins up bash terminals, and deploys...

[Moonshot AI's Attention Residuals \(AttnRes\) Achieves Same Training Loss with 1.25× Less Compute—New Transformer Architecture Replaces Fixed Residuals with Softmax Attention](#)

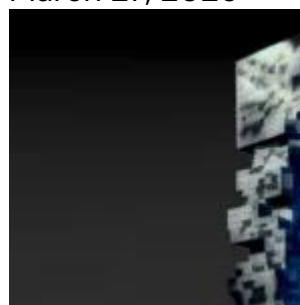
March 26, 2026



Residual connections have been sacred architecture for nine years—and Moonshot AI just proved they were holding us back....

[Microsoft Unifies Copilot Teams on March 17—15 Million Users Represent Just 3% of Enterprise Base as Company Admits Lag Behind OpenAI and Google](#)

March 27, 2026



Microsoft just published its own report card, and the grade is a C-minus. After two years and billions...

[Meta Locks 6.6 GW of Nuclear Power Through Vistra, TerraPower, and Oklo Deals—Largest Corporate Nuclear Commitment for AI Infrastructure](#)

March 24, 2026



Meta just committed to more nuclear capacity than Switzerland's entire grid produces. The AI energy arms race has...

[OpenAI Launches GPT-5.4 'Thinking' with 1 Million-Token Context Window on March 5, 2026](#)

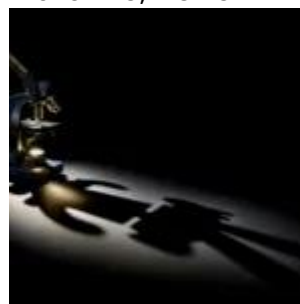
March 17, 2026



OpenAI just shipped a model that can hold 2,000 pages of text in working memory. Most enterprise RAG...

[University of Montreal Tests AI Against 100,000 Humans on Creativity—GPT-4 Beats 72% But Top 10% Still Win](#)

March 16, 2026



GPT-4 now outperforms the average human on standardized creativity tests. But before you fire your creative team, consider...

[31 Companies Hit Unicorn Status in January 2026—Highest Monthly Count Since June 2022, Adding \\$58.5B in Valuation](#)

March 7, 2026



The funding winter ended two months ago and most of the industry missed it. Fifty-four companies crossed the...

[SpaceX Acquires xAI for \\$250 Billion—Largest Private Acquisition in History Creates \\$1.25 Trillion AI-Space Infrastructure Giant](#)

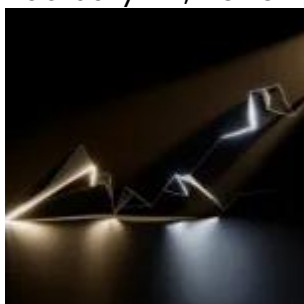
February 26, 2026



Elon Musk just closed the largest private acquisition in history—and it's not about rockets or AI. It's about...

[University of Montreal Study Proves AI Beats Average Humans on Creativity Tests—But Top 10% Still Outperform GPT-4](#)

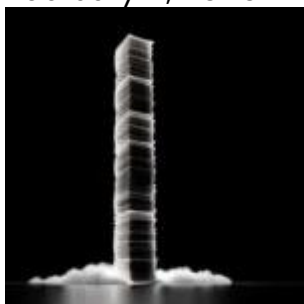
February 24, 2026



The world's largest creativity study just revealed an uncomfortable truth: half of humanity is now less creative than...

[Claude Opus 4.6 Scores 76% on Long-Context Retrieval—4X Better Than Its Predecessor at 18.5%](#)

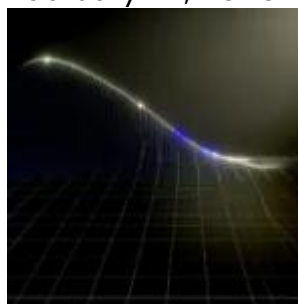
February 11, 2026



A 310% improvement in a single release isn't iteration—it's a discontinuity. Anthropic just proved that model performance can...

[Google Gemini 3.1 Pro Scores 77.1% on ARC-AGI-2—2.5x Jump Over Predecessor in Single Generation](#)

February 22, 2026



Google just doubled AI reasoning capability in 90 days while keeping the price identical. The assumption that frontier...

[OpenClaw Hits 150,000 GitHub Stars in 10 Weeks—Open-Source AI Assistant Overtakes Major Projects with 416,000+ npm Downloads](#)

February 6, 2026



A weekend project by an Austrian developer just outpaced a decade of GitHub growth patterns—150,000 stars in 10...

[ChatGPT's Market Share Drops to 61.3% as Gemini Surges 237% Year-Over-Year—The AI Chatbot Monopoly Era Ends](#)

February 4, 2026



ChatGPT lost 25 percentage points of market share in 12 months. The company that ate its lunch isn't...

[Cisco 2026 Study: 38% of Organizations Now Spend \\$5M+ on Privacy—Up from 14% in 2024 as AI Drives 90% to Expand Programs](#)

January 27, 2026



Privacy budgets just tripled. Not because of new regulations—because companies finally realized what AI does with their data.

[Mercedes-Benz Integrates Google Cloud's Automotive AI Agent into MBUX Virtual Assistant, Launching in New CLA Model in 2025](#)

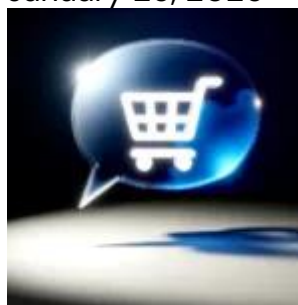
February 2, 2026



Mercedes just shipped an in-car AI that remembers your last question and answers follow-ups in three seconds. The...

[Microsoft Copilot Checkout Launches with PayPal, Shopify, and Stripe Integration: Shopping Sessions 33% Shorter, Purchases Up 194%](#)

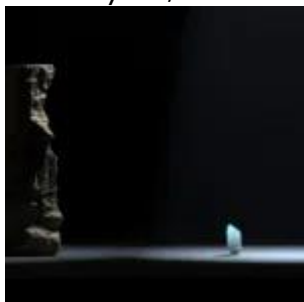
January 26, 2026



Microsoft just weaponized its conversational AI into a transaction layer—and the early numbers suggest a fundamental shift in...

[TII's Falcon-H1R 7B Outperforms 47B Models on Math Reasoning While Running on a 16GB Laptop](#)

January 22, 2026



A 7-billion parameter model just scored 88.1% on AIME-24 math reasoning, beating models with 47 billion parameters. The...

[Google's 12B TranslateGemma Outperforms Its Own 27B Model: Open Translation Hits 55 Languages with MetricX Score of 3.60](#)

January 20, 2026



Google's smaller translation model just beat its larger sibling on standardized benchmarks, forcing us to reconsider everything we...

[OpenAI Signs \\$10 Billion Cerebras Deal for 750 Megawatts of AI Inference Infrastructure Through 2028](#)

January 21, 2026



OpenAI just mass-ordered 750 megawatts of specialized silicon that isn't made by NVIDIA. The \$10 billion Cerebras deal...

[xAI Raises \\$20 Billion at \\$230 Billion Valuation, Plans 50 Million H100-Equivalent GPU Deployment by 2030](#)

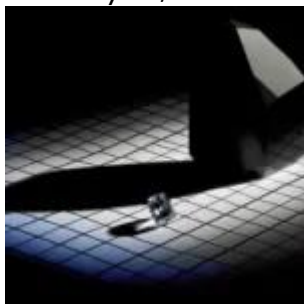
January 18, 2026



xAI just raised more in one round than most AI companies are worth—and plans to deploy 50× more...

[TII's Falcon-H1R 7B Outperforms 47B Models on Math Reasoning While Running on a 16GB Laptop](#)

January 15, 2026



A 7-billion parameter model just scored 88.1% on AIME-24 math reasoning, crushing NVIDIA's 47B Nemotron at 49.7%. The...

[The 8-Second War Plan: Why Air Force DASH-2's AI-Generated Courses of Action Just Made Human Military Oversight Mathematically Impossible](#)

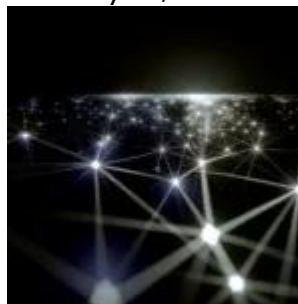
January 10, 2026



The Pentagon just proved that human commanders can't keep up with AI warfare—and they're deploying it anyway.

[Meta Announces \\$600 Billion AI Infrastructure Spend Through 2028, Creating 'Meta Compute' Division to Build Tens of Gigawatts of Data Center Capacity](#)

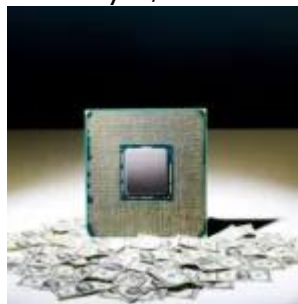
January 14, 2026



Meta just committed more capital to AI infrastructure than the entire GDP of Finland, Belgium, or Thailand combined—and...

[The Compute-Optimal Blind Spot: Why Most ML Teams Are Still Wasting 40% of Training Budget on the Wrong Scaling Trade-Off](#)

January 8, 2026



Your ML team probably burned through 40% of its training budget last quarter—and nobody noticed because the waste...