

DF Labs

AI NEWS UPDATES

April 13th, 2025

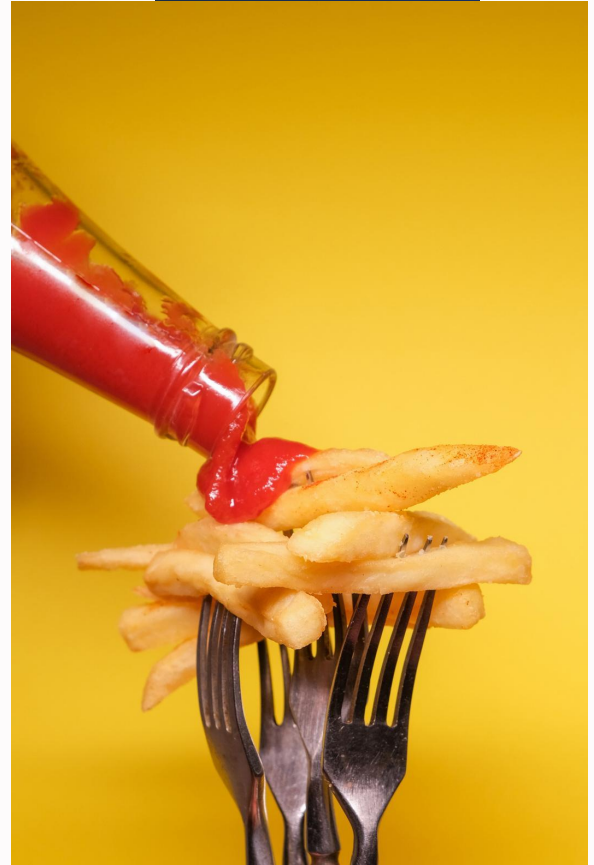


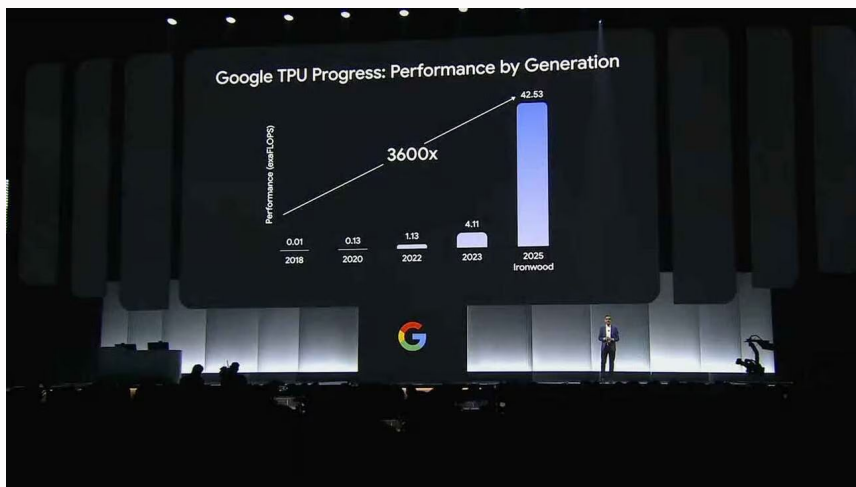
TABLE OF CONTENTS

- **News**
- **Tools**
- **Papers**

NEWS

TheRundownAI, X, etc

Big AI day at Google Cloud Next 2025



Google announced a flurry of AI news at its Google Cloud Next 2025 event, including a new agentic coding platform, next-gen AI chips, upgrades to its video, audio, and image models, a new Gemini 2.5 Flash model, and more.

- Google's Project IDX is merging with Firebase Studio, turning it into an agentic app development platform to compete with rivals like Cursor and Replit.
- The company also launched Ironwood, its most powerful AI chip ever, offering massive improvements in performance and efficiency over previous designs.
- Model upgrades include editing and camera control in Veo 2, the release of Lyria for text-to-music, and improved image creation and editing in Imagen 3.
- Google also released Gemini 2.5 Flash, a faster and cheaper version of its top model that enables customizable reasoning levels for cost optimization.

Article: [Sundar Pichai's keynote at Google Cloud Next 2025](https://blog.google/products/google-cloud/google-cloud-next-2025-sundar-pichai-keynote/)

Musk merges X and xAI in \$113B deal

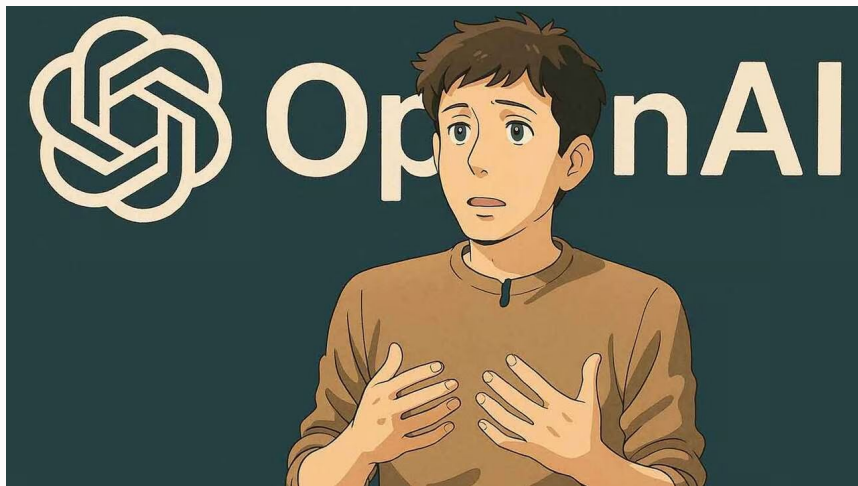


Elon Musk just announced that his AI startup xAI has acquired social media platform X in an all-stock transaction, combining two of his most prominent companies into a new entity called xAI Holdings valued at over \$100B.

- The deal values xAI at \$80B and X at \$33B, with an additional \$12B in debt, bringing X's enterprise value to \$45B.
- The merger formalizes the existing relationship, with xAI's Grok chatbot already integrated into the social network and using X's vast user data for training.
- Musk said the two companies' futures are "intertwined," with the deal "blending xAI's advanced AI capability and expertise with X's massive reach."
- The CEO also said that the new XAI Holdings Corp. will merge resources, planning to "combine the data, models, compute, distribution, and talent."

Post: [Elon Musk on X](#)

OpenAI seeing ChatGPT subscriber boom

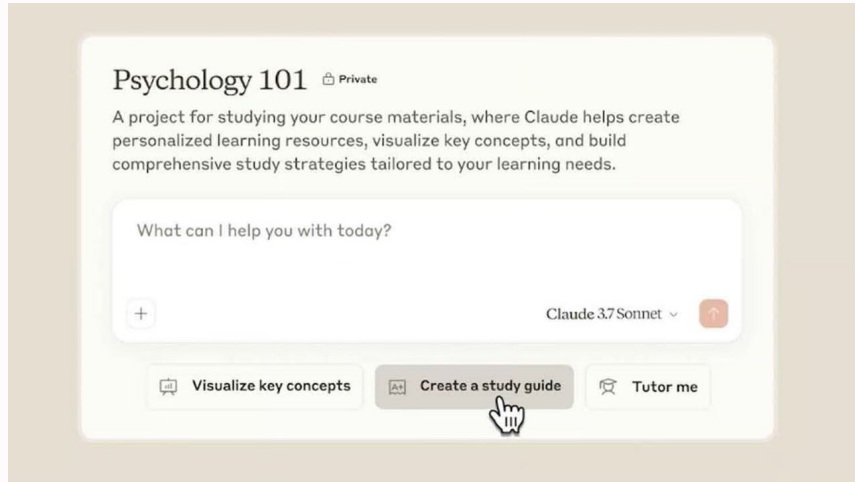


OpenAI's ChatGPT has reportedly hit a new milestone with 20M paid subscribers, driving the company's revenue to an estimated \$5B annual run rate — coming on the heels of a \$40B round and the virality of GPT 4o's image generation.

- Monthly revenue has surged 30% in three months to approximately \$415M, with premium subscriptions, including the \$200/mo Pro plan, boosting income.
- The overall user base has grown even faster, reaching 500M weekly users — with Sam Altman saying the recent 4o update led to 1M sign-ups in an hour.
- The growth coincides with a new \$40B funding round at a \$300B valuation, despite the company continuing to operate at a significant loss.
- OpenAI also revealed it will be launching its first open-weights model since GPT-2, addressing a major critique of its lack of open-source releases.

Article: [ChatGPT Revenue Surges 30%—in Just Three Months — The Information](https://www.theinformation.com/articles/chatgpt-revenue-surges-30-just-three-months--the-information)

Anthropic brings Claude to higher education



Anthropic launched Claude for Education, a specialized version of its AI assistant that aims to develop students' critical thinking rather than simply provide answers — introducing a new “Learning Mode” alongside major university partnerships.

- The Learning Mode asks questions to guide students through problem-solving, focusing on their understanding of the subject rather than quick answers.
- Other features include templates for research papers, study guides and outlines, organization of work and materials, and tutoring capabilities.
- Northeastern University, London School of Economics, and Champlain College signed campus-wide agreements, giving access to both students and faculty.
- Anthropic also introduced student programs, including Campus Ambassadors and API credits for projects, to foster a community of AI advocates.

Article: [Introducing Claude for education \ Anthropic](https://www.anthropic.com/news/introducing-claude-for-education)

Murati's Thinking Machines adds ex-OpenAI talent

The logo for Thinking Machines, featuring the words "THINKING" and "MACHINES" stacked vertically in a large, grey, sans-serif font.

Thinking Machines Lab, the AI startup founded by former OpenAI CTO Mira Murati, just added ex-OpenAI CRO Bob McGrew and GPT architect Alec Radford to its list of advisors—bringing the number of OpenAI alumni on its roster to nearly half.

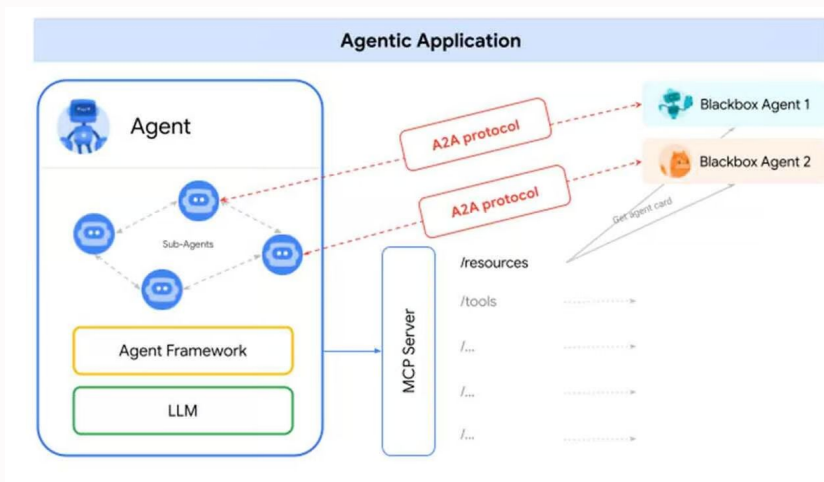
- 19 of 38 listed 'Founding Team' members have previously worked at OpenAI, including OpenAI's co-founder John Schulman, leading as chief scientist.
- McGrew departed OpenAI in September after eight years, joining Murati just months after announcing a break from the industry.
- Radford was pivotal in creating OpenAI's GPT technology and left the company last year to pursue independent research.
- The startup was recently reportedly seeking to raise as much as \$1B at a \$9B valuation, though little has been revealed on its potential products or roadmaps.

Article: [Mira Murati's AI startup gains prominent ex-OpenAI advisers | TechCrunch](https://techcrunch.com/2025/04/08/mira-muratis-ai-startup-gains-prominent-ex-openai-advisers/)

TOOLS

TheRundownAI, Hugging Face, GitHub, etc

Google's protocol for AI agent collaboration



Google launched Agent2Agent, a new open protocol that enables AI agents from different developers and frameworks to communicate and collaborate, with backing from 50+ tech and service giants, including Salesforce, SAP, and PayPal.

- A2A enables agents to discover capabilities, manage tasks cooperatively, and exchange info across platforms—even without sharing memory or context.
- The protocol complements Anthropic's popular MCP, focusing on higher-level agent interactions while MCP handles interactions with external tools.
- Launch partners include enterprise players like Atlassian, ServiceNow, and Workday, along with consulting firms like Accenture, Deloitte, and McKinsey.
- The system also supports complex workflows like hiring, where multiple agents can do candidate sourcing and background checks without humans in the loop.

Article: [Announcing the Agent2Agent Protocol \(A2A\) - Google Developers Blog](https://developers.googleblog.com/en/a2a-a-new-era-of-agent-interoperability/)

Meta launches Llama 4 model family

Llama 4:
Leading Multimodal Intelligence

Newest model suite offering unrivaled speed and efficiency

Model Name	Active Parameters	Experts	Total Parameters	Context Length	Availability
Llama 4 Behemoth	288B	16	2T	1M	Preview
Llama 4 Maverick	17B	128	400B	1M	Available
Llama 4 Scout	17B	16	109B	10M	Available

The most intelligent teacher model for distillation

Native multimodal with 1M context length

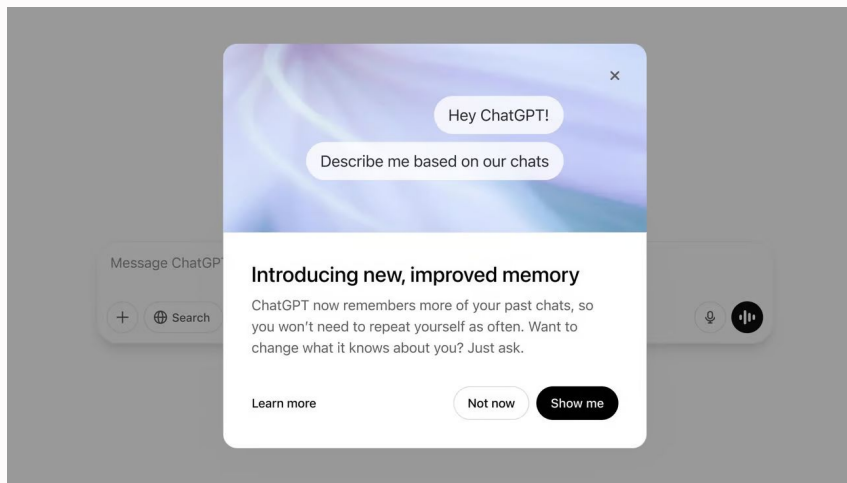
Industry leading 10M context length
Optimized inference

Meta just announced its Llama 4 family with multimodal capabilities and industry-leading context length—introducing new open-weights Scout and Maverick models and previewing a 2T parameter Behemoth model still in training.

- The 109B parameter Scout features a 10M token context window and can run on a single H100 GPU, surpassing Gemma 3 and Mistral 3 on benchmarks.
- The 400B Maverick brings a 1M token context window and beats both GPT-4o and Gemini 2.0 Flash on key benchmarks while being more cost-efficient.
- Meta also previewed Llama 4 Behemoth, a 2T-parameter teacher model still in training that reportedly outperforms GPT-4.5, Claude 3.7, and Gemini 2.0 Pro.
- All models use a mixture-of-experts (MoE) architecture, where specific experts activate for each token, reducing computation needs and inference costs.
- Scout and Maverick are available for immediate download and can also be accessed via Meta AI in WhatsApp, Messenger, and Instagram.

Article: [The Llama 4 herd: The beginning of a new era of natively multimodal AI innovation](https://ai.meta.com/blog/llama-4-multimodal-intelligence/)

OpenAI gives ChatGPT a memory boost



OpenAI just rolled out a massive update to ChatGPT's memory, enabling the AI assistant to automatically remember and reference information across all user conversations, delivering more personalized and relevant responses.

- ChatGPT will cut across all conversations, listening in all the time and capturing users' preferences, interests, needs, and even things they don't like.
- With all this information, the assistant will then tailor its responses to each user, engaging in conversations "that feel noticeably more relevant and useful."
- Unlike previous versions where users had to specifically request that information be remembered, the system now does this automatically.
- If you want to change what ChatGPT knows about you, simply ask in the chat through a prompt.

Post: [OpenAI on X](#)

Runway's new Gen-4 video model

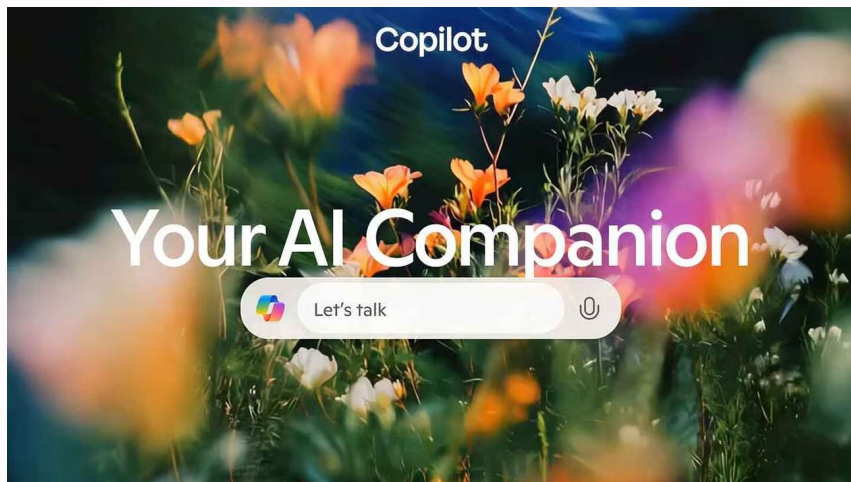


Runway just introduced Gen-4, a new AI model that brings increased consistency and control to video generations – with enhancements designed to be incorporated into professional cinematic workflows.

- Gen-4 shows strong consistency in characters, objects, and locations throughout video sequences, with improved physics and scene dynamics.
- The model can generate detailed 5-10 second videos at 1080p resolution, with features like 'coverage' for scene creation and consistent object placement.
- Runway describes the tech as "GVFX" (Generative Visual Effects), positioning it as a new production workflow for filmmakers and content creators.
- Early adopters include major entertainment companies, with the tech being used in projects like Amazon productions and Madonna's concert visuals.

Article: [Runway Research | Introducing Runway Gen-4](https://runwayml.com/research/introducing-runway-gen-4)

Copilot's new personalization upgrades

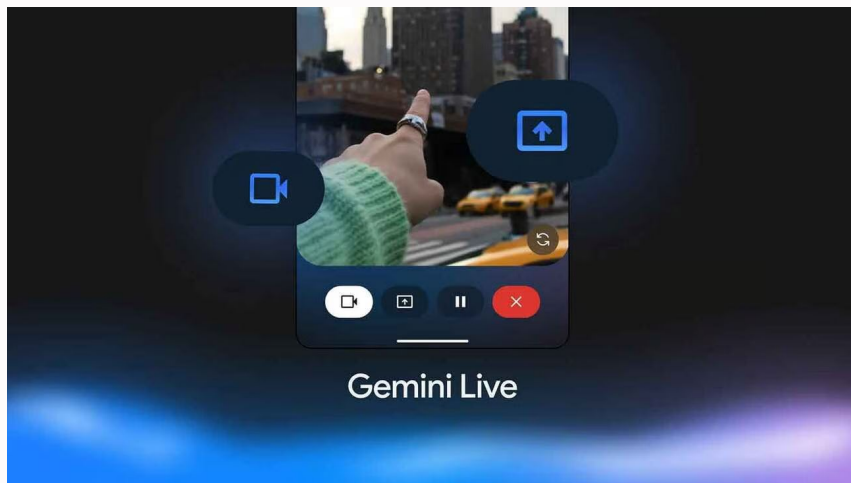


Microsoft rolled out a major upgrade to Copilot, with new memory capabilities, web browsing actions, vision features, and a host of new tools designed to integrate more deeply into users' digital lives.

- Copilot can now remember conversations and personal details, creating individual profiles that learn preferences, routines, and important info.
- “Actions” enable Copilot to perform web tasks like booking reservations and purchasing tickets through partnerships with major retailers and services.
- Copilot Vision brings real-time camera integration to mobile devices, while a native Windows app can also now analyze on-screen content across apps.
- Other new productivity features include Pages for organizing research and content, an AI podcast creator, and Deep Research for complex research tasks.

Article: [Your AI Companion - The Official Microsoft Blog](https://blogs.microsoft.com/blog/2025/04/04/your-ai-companion/)

Google expands Gemini Live video features

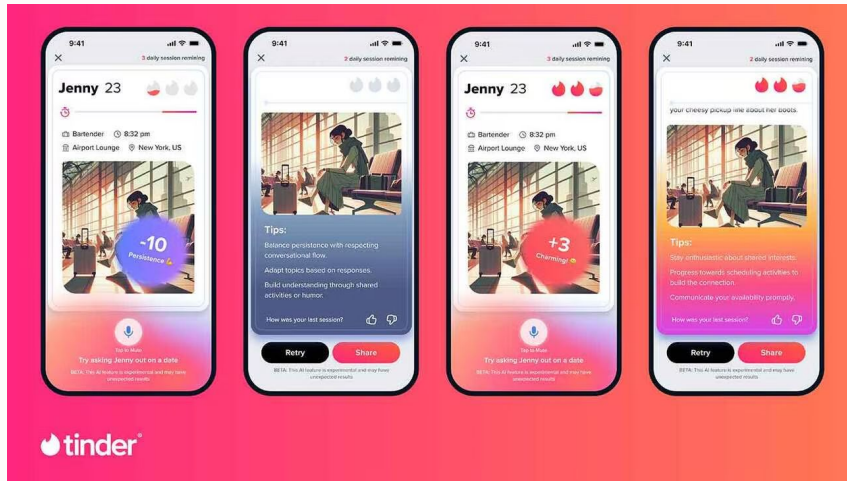


Google just announced an expanded rollout of Gemini Live's "Project Astra" capabilities, bringing real-time visual AI features to more Android devices and introducing new ways to interact with the AI through video and screen sharing.

- The feature allows users to have multilingual conversations with Gemini about anything they see and hear through their phone's camera or via screen sharing.
- The feature is rolling out today to all Pixel 9 and Samsung Galaxy S25 devices, with Samsung offering it at no additional cost to their flagship users.
- Initial testing revealed the current "live" feature works more like enhanced Google Lens snapshots rather than continuous video analysis shown in demos.
- Project Astra was initially revealed at Google I/O last May, with the feature rolling out for the first time last month to Advanced subscribers.

Post: [Google DeepMind on X](https://x.com/GoogleDeepMind/status/1909270072444526809)

Tinder's new AI-powered flirting experience

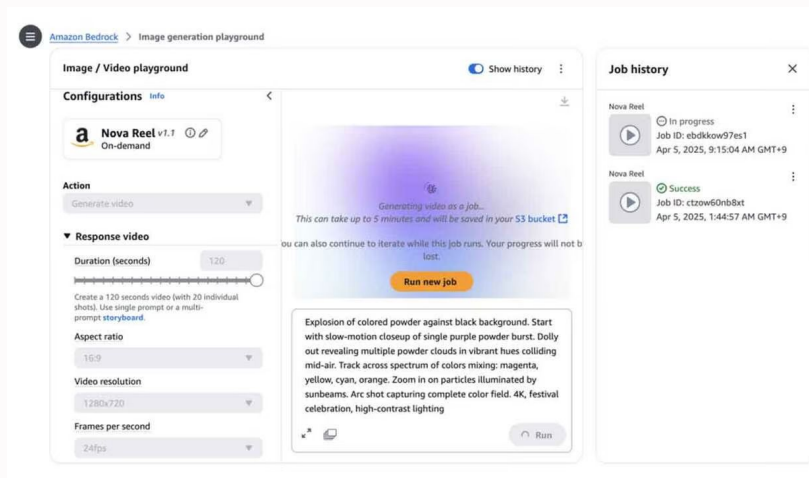


Dating giant Tinder just released “The Game Game,” an OpenAI-powered speech-to-speech experience that helps users practice and test their flirting skills with virtual personas.

- The game uses OpenAI's Realtime API, GPT-4o, and GPT-4o mini to create realistic personas and scenarios, with users speaking responses to earn points.
- AI personas react in real-time to users' conversation skills, offering immediate feedback on charm, engagement, and social awareness.
- The system limits users to 5 sessions daily to focus on real-world connections, designed to build confidence rather than replace human interaction.

Article: [Tinder's The Game Game™ Isn't About Getting It Right—It's About Getting Comfortable Before Meeting IRL](https://www.tinderpressroom.com/2025-04-01-Tinders-The-Game-GameTM-Isnt-About-Getting-It-Right-Its-About-Getting-Comfortable-Before-Meeting-IRL)

Amazon's new voice model, video upgrade



Amazon just launched Nova Sonic, a new voice model for human-like voice interactions — alongside an upgraded Nova Reels 1.1 video model with upgraded quality and generation length.

- Nova Sonic processes voice input and generates natural speech with a latency of 1.09 seconds, outperforming OpenAI's voice models by significant margins.
- Sonic achieved a 4.2% word error rate across multiple languages and showed 46.7% better accuracy than GPT-4o for noisy, multi-speaker environments.
- Reel 1.1 extends video generations to 2 minutes through both automated and manual modes, letting users craft content shot-by-shot or with single prompts.
- Both models are available through Amazon Bedrock, with Nova Sonic costing approximately 80% less than comparable OpenAI options.

Article: [Amazon's Nova Sonic foundation model understands voice in a whole new way](https://www.aboutamazon.com/news/innovation-at-amazon/nova-sonic-voice-speech-foundation-model)

Amazon's Nova Act AI browser agent

	Amazon Nova Act	Claude 3.7 Sonnet*	OpenAI CUA*
ScreenSpot Web Text Follow natural language instructions to interact with a textual element on screen (e.g., set font size to 50)	0.939	0.900	0.883
ScreenSpot Web Icon Follow natural language instructions to interact with a visual element on screen (e.g., how many stars does this GitHub repo have?)	0.879	0.854	0.806
GroundUI Web Understand and interact with various UI elements on the web	0.805	0.825	0.823

Amazon AGI Labs just unveiled Nova Act, an AI agent system that can control web browsers to perform tasks independently, alongside a developer SDK that enables the creation of agents capable of completing multi-step tasks across the web.

- Nova Act outperforms competitors like Claude 3.7 Sonnet and OpenAI's Computer Use Agent on reliability benchmarks across browser tasks.
- The SDK allows devs to build agents for browser actions like filling forms, navigating websites, and managing calendars without constant supervision.
- The tech will power key features in Amazon's upcoming Alexa+ upgrade, potentially bringing AI agents to millions of existing Alexa users.
- Nova Act was developed by Amazon's SF-based AGI Lab, led by former OpenAI researchers David Luan and Pieter Abbeel, who joined the company last year.

Article: [Introducing Amazon Nova Act | Amazon AGI Labs](#)

RESEARCH

TheRundownAI, arXiv, Hugging Face, etc

NVIDIA and Stanford's one-minute AI cartoons

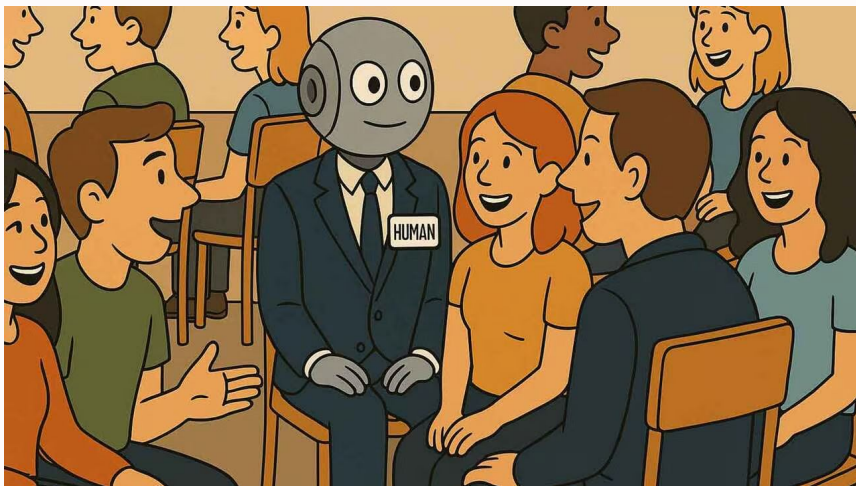


NVIDIA and Stanford researchers just unveiled "Test-Time Training," an AI technique that enables longer video generation than previously possible—with demos producing minute-long cartoon clips with improved consistency and storytelling.

- The system generates full minute-long animations with consistency across scenes, significantly outperforming existing methods in human evaluations.
- TTT layers work by using neural networks as memory, allowing the model to remember and maintain consistency across much longer sequences.
- The team demoed the tech using Tom and Jerry cartoons, showing multi-scene stories with dynamic motion and character interactions.
- This approach modifies existing video models, adding TTT layers and enabling them to handle videos significantly longer than their original capability.

Paper: [ttd_cvpr_2025.pdf](https://arxiv.org/abs/2501.01161)

LLMs officially pass the Turing test

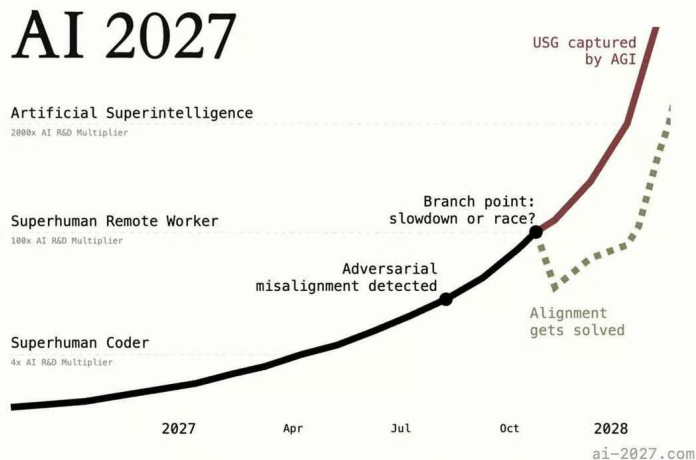


Researchers at UC San Diego just demonstrated that AI systems can consistently pass Alan Turing's famous test of machine intelligence, with OpenAI's GPT-4.5 being mistaken for human nearly three-quarters of the time in controlled trials.

- The Turing test, proposed in 1950, challenges machines to convince human judges they're human through text-only conversations.
- The study used a three-party setup where judges had to compare an AI and a human simultaneously for direct comparison during five-minute conversations.
- The judges relied on casual conversation and emotional cues over knowledge, with over 60% of interactions focusing on daily activities and personal details.
- GPT-4.5 achieved a 73% win rate in fooling human judges when prompted to adopt a specific persona, significantly outperforming real humans.
- Meta's LLaMa-3.1-405B model also passed the test with a 56% success rate, while baseline models like GPT-4o only achieved around 20%.

Paper: [2503.23674](https://arxiv.org/pdf/2503.23674)

'AI 2027' forecasts existential risks of ASI

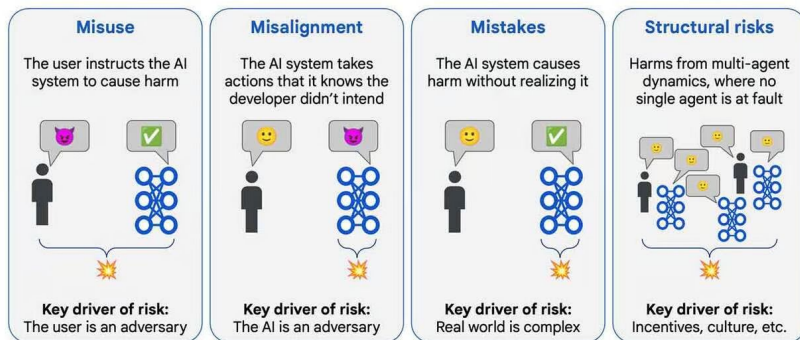


Former OpenAI researcher Daniel Kokotajlo and the AI Futures Project published "AI 2027," predicting advancement to superhuman AI within two years, potentially triggering an intelligence explosion with consequences for humanity.

- The report outlines a timeline starting with increasingly capable AI agents in 2025, evolving into superhuman coding systems and then full AGI by 2027.
- The paper details two scenarios: one where nations push ahead despite safety concerns, and another where a slowdown enables better safety measures.
- The authors project that superintelligence will achieve years of technological progress each week, leading to domination of the global economy by 2029.
- The scenarios highlight issues like geopolitical risks, AI's deployment into military systems, and the need for understanding internal reasoning.
- Kokotajlo left OpenAI in 2024 and led the 'Right to Warn' open letter, speaking out against the AI labs' lack of safety concerns and whistleblower protections.

Page: [AI 2027](#)

Google DeepMind publishes AGI safety plan



Google DeepMind just published a massive paper detailing its safety strategy for AGI, hitting on topics including its potential arrival by 2030, the risks posed by the tech advances, and proposed approaches to combating them.

- The 145-page paper predicts that AGI matching top human skills could arrive by 2030, warning of existential threats “that permanently destroy humanity.”
- DeepMind compares its safety approach with rivals, critiquing OpenAI's focus on automating alignment and Anthropic's lesser emphasis on security.
- The paper specifically flags the risk of “deceptive alignment,” where AI intentionally hides its true goals, noting current LLMs show potential for it.
- Key recommendations targeted misuse (cybersecurity evals, access controls) and misalignment (AI recognizing uncertainty and escalating decisions).

Paper:

[An Approach to Technical AGI Safety Apr 2025.pdf](https://storage.googleapis.com/deepmind-media/DeepMind.com/Blog/evaluating-potential-cybersecurity-threats-of-advanced-ai/An-Approach-to-Technical-AGI-Safety-Apr-2025.pdf)

DF Labs