

The Long Read

14 October 2026

10

STORIES

3

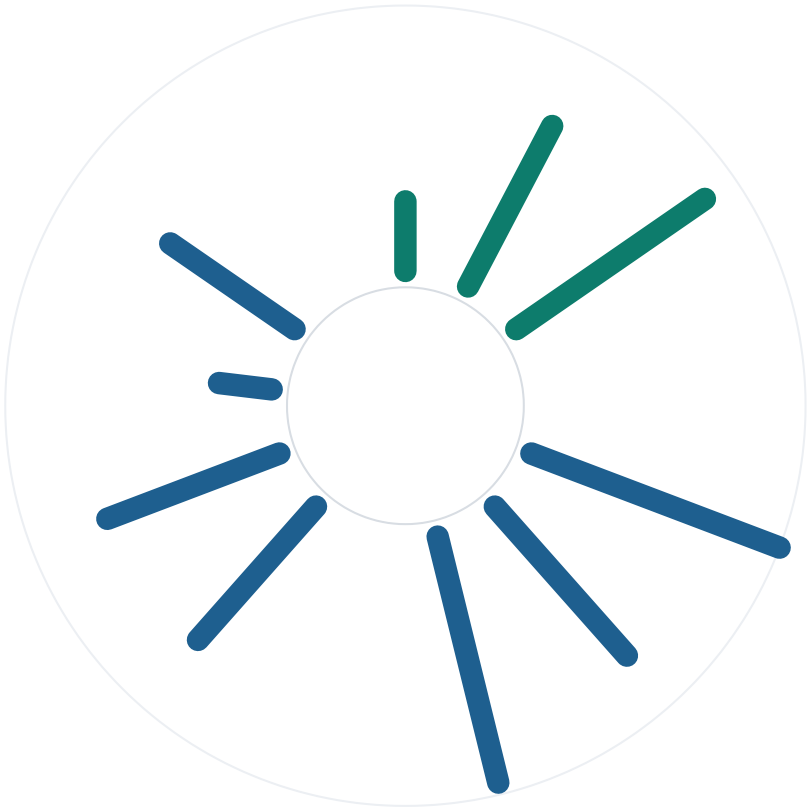
SECTIONS

3279

WORDS

3

HIGH IMPACT



● 3 high impact ● 7 medium impact spoke length = depth of coverage

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

The issue at a glance

10 stories · 3279 words · 3 sections · 4 charted



AI Models & Research 3 stories · 951 words

AI Tools & Ecosystem 3 stories · 1109 words

AI Applications & Industry 4 stories · 1219 words

AI Models & Research

- 01 ● Anthropic Replaced Git: 2,500 Daily Commits Became 30,000, So They Rewrote Source Control
- 02 ● Liquid AI Ships LFM2.5-24B-A2B: a Reasoning Agent Model With 2B Active Parameters
- 03 ● Gemini 4 Argon for Robotics: 87% on Embodied Reasoning, Ahead of Every Frontier Model

AI Tools & Ecosystem

- 04 ● AI Labs Are Buying the Data of Dead Startups to Train Their Models

- 05 ● Ollama Goes to Fly.io: Sprites Give AI Agents a Private Linux VM With Secret-Free Model Access
- 06 ● Google Play Opens to Private, Vibe-Coded Apps That Never Hit the Store

AI Applications & Industry

- 07 ● Character.AI Files for Chapter 11 Five Years In
- 08 ● Amex's Intent Intelligence Framework Wants AI Agents to Sign Their Receipts
- 09 ● DeepMind and Singapore Fine-Tune a Gemma Model to Catch Phone Scams on-Device
- 10 ● Former Neuralink President's PRIMA Chip Restores Reading to a Stargardt Patient

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

AI Models & Research

3 stories ● 3 high

01 **HIGH IMPACT** Simon Willison's Weblog

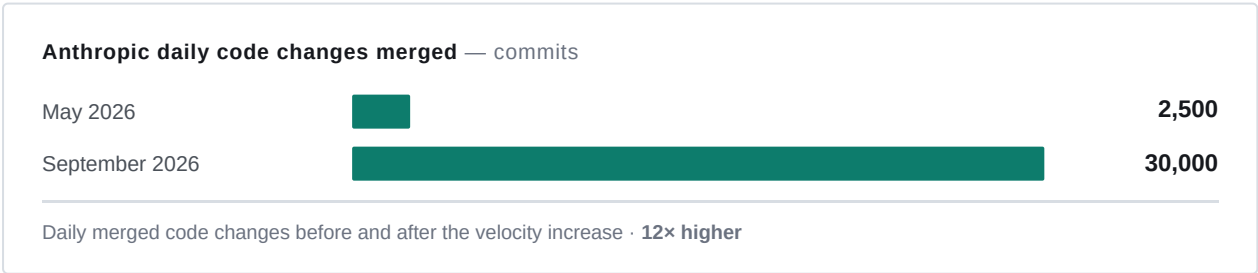
Anthropic Replaced Git: 2,500 Daily Commits Became 30,000, So They Rewrote Source Control

Anthropic rebuilt source control from scratch after its daily merge rate rose more than 12x in four months, and the resulting system, Scion, abandons Git's distributed model entirely.

Anthropic has disclosed that it designed and deployed a custom source control system, Scion, to replace Git after its code-change velocity exploded. In May 2026 the company averaged around 2,500 code changes merged per day; by September that figure exceeded 30,000 per day. The main branch history alone is almost 100 GB, a compressed working copy is over 25 GB, and CI performs on the order of 25 million checkouts per day — which, fetched naively, would mean over 600 PB of Git traffic daily. The new system was built by a team of seven people, and the migration from Git was completed in five months.

Scion is described as a reinvention of centralized source control along the lines of Subversion, rather than another distributed system. Anthropic also built a new code review system called Antique alongside it. When asked on Hacker News about the lack of offline support, Scion engineer Nick G said that since Claude is by far the heaviest user of Scion and is necessarily online, offline support simply wasn't important to build.

Anthropic also makes extensive use of adversarial review in its development process. The article notes that while this is incredibly effective in many ways, it can also be myopic: skilled engineers needed to steer Claude away from over-engineering parts of the system in response to hypothetical concerns. Willison speculates that the introduction of Fable/Mythos class models is a key factor in the increased coding velocity, and his annotated chart of the commit-rate increase marks public release dates of Anthropic's 2026 models.



KEY FACTS

DAILY MERGES, MAY 2026

~2,500

MAIN BRANCH HISTORY SIZE

almost 100 GB

DAILY CI CHECKOUTS

~25 million

DAILY MERGES, SEPTEMBER 2026

over 30,000

COMPRESSED WORKING COPY

over 25 GB

NAIVE DAILY GIT TRAFFIC

over 600 PB

WHY IT MATTERS

If AI-driven development pushes commit and checkout volumes to Anthropic's scale, Git's distributed model and naive fetch patterns become operationally untenable. Teams building agent-heavy CI should watch whether centralized source control makes a comeback.

[Read the original at Simon Willison's Weblog →](#)

02

HIGH IMPACT

huggingface.co

Liquid AI Ships LFM2.5-24B-A2B: a Reasoning Agent Model With 2B Active Parameters

Liquid AI has released LFM2.5-24B-A2B, an open-weight reasoning agent model that runs with only 2B active parameters yet beats Qwen3.6-35B-A3B on most agentic benchmarks.

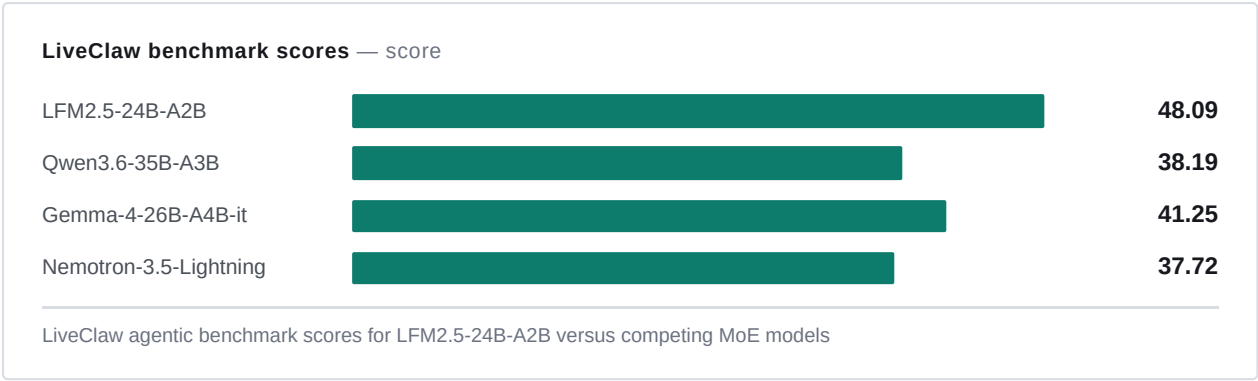
LFM2.5-24B-A2B builds on LFM2-24B-A2B and keeps its architecture: 40 layers, 64 experts per MoE block with top-4 routing, and the same hybrid of gated short convolutions and GQA. The context window grows from 32K to 131K tokens, and an in-place tokenizer expansion doubles the vocabulary from 65K to 128K tokens for better non-Latin tokenization. Unlike its predecessor, it is a reasoning-only model that produces an explicit chain of thought before its final answer.

Pre-training was scaled up from ~17T to ~33T tokens. Mid-training then adds 2T tokens at 32K context, weighted toward reasoning, math, code, instruction following, and multilingual text, and 300B tokens at 131K context, weighted toward long-context data. Tool calls and multi-turn trajectories appear throughout both phases. Post-training turns the base model into an agent in five stages: supervised fine-tuning, teacher specialization, multi-domain on-policy distillation (MOPD) on non-agentic tasks, a second MOPD stage focused on agentic tasks, and agentic reinforcement learning inside harnesses such as OpenClaw, Hermes Agent, Pi, and OpenCode.

On 16 benchmarks covering tool use, agentic workflows, knowledge, instruction following, math, and coding, LFM2.5-24B-A2B leads or matches the competition on five of eight tool-use and agentic benchmarks. It scores 48.09 on LiveClaw versus 38.19 for Qwen3.6-35B-A3B, and on AutomationBench it scores 27.69 — more than three times the next-best model. It also leads on all three instruction-following benchmarks (IFEval 96.75, IFBench 74.69, Multi-IF 84.38). Larger models keep an edge on knowledge-heavy benchmarks like GPQA Diamond and HLE, and on coding benchmarks like SWE-Bench Verified, where Qwen3.6-35B-A3B scores 66.20 versus 51.00.

Inference is fast on both edge and GPU hardware. LFM2.5-24B-A2B decodes 184 tokens/s on an Apple M5 Max and 114 tokens/s on a Ryzen AI Max+ 395, up to 1.7x faster than the next-best model while using the least

memory. On a single NVIDIA B300 GPU it reaches over 21K output tokens per second at high concurrency. Day-one support covers llama.cpp, MLX, vLLM, SGLang, and ONNX. Both the instruct and base models are open-weight on Hugging Face.



KEY FACTS

ACTIVE PARAMETERS

2B

VOCABULARY SIZE

128K tokens

LIVECLAW SCORE

48.09

B300 OUTPUT SPEED

21K tok/s

CONTEXT WINDOW

131K tokens

PRE-TRAINING TOKENS

~33T

M5 MAX DECODE SPEED

184 tok/s

LICENCE

Open-weight

WHY IT MATTERS

A 2B-active-parameter model that leads on tool-use and instruction-following benchmarks changes the calculus for on-device agents: you can run a reasoning model locally without a GPU cluster and still beat much larger MoEs on agentic workloads.

[Read the original at huggingface.co](#) →

03

HIGH IMPACT

deepmind.google

Gemini 4 Argon for Robotics: 87% on Embodied Reasoning, Ahead of Every Frontier Model

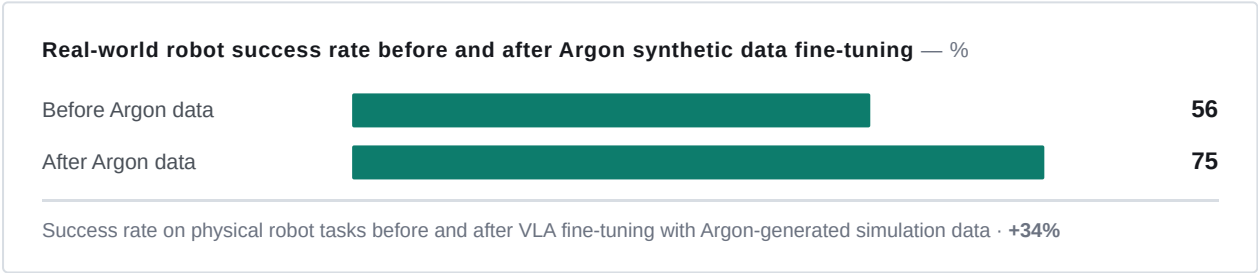
Gemini 4 Argon reaches 87% on the ERQA embodied reasoning benchmark, outperforming every other frontier model, and is positioned as a general-purpose physical reasoning model across the robotics stack.

Google DeepMind has announced Gemini 4 Argon, a model aimed at physical understanding, code generation, and visual-semantic reasoning for robotics. The headline number is 87% on the Embodied Reasoning (ERQA) benchmark, which tests spatial problem-solving in 3D environments using pure reasoning. DeepMind states this outperforms every other frontier model. The announcement frames Argon as a general value function for robotics tasks, applicable across the full robotics lifecycle rather than a single narrow controller.

The post demonstrates Argon across a wide range of workflows. It controls a Boston Dynamics Spot quadruped (12-DoF legs plus 6-DoF arm) to step on a foot-pedal trash bin and discard soda cans, first practicing on an articulated MuJoCo digital twin. It steers robots in the Gemini Omni World Model to rearrange Scrabble tiles into a row spelling A-R-G-O-N via closed visual loop prompting. It builds MuJoCo scenes from real teleoperation episodes, matching rendered camera views and contact forces against logged tactile data, with constraints that forbid sim cheating such as external forces or collision disabling. It reconstructs manipulation tasks from egocentric human video onto a 76-DoF bimanual humanoid with 22-DoF five-fingered hands at 500 Hz. It converts a hand-drawn sketch of a humanoid throwing a ball at a 3-2-1 box stack into a verified mujoco_warp scene with a 50 Hz, 140-step, 2.80-second throwing trajectory that knocks down all 6 boxes.

Two results stand out for practitioners. First, Argon generates synthetic data from a single RGB photo of a real tabletop, builds a calibrated MuJoCo twin, defines a 77-task curriculum, and synthesizes 4,500+ domain-randomized episodes. Fine-tuning a generalist VLA model on this data raised real-world success rate from 56% to 75% and cut completion time from 35s to 24s. Second, on a 6-DoF arm pick-and-place task, Argon self-improves through trial and self-reflection, diagnosing failures from wrist and overhead footage and optimizing execution speed at least 2× within safety limits. It also reproduces robotics papers into self-contained MuJoCo WebAssembly browser labs, running neural inference or Web Worker iLQR in JavaScript and verifying in headless Chromium over 300+ frames.

The model is available as gemini-4-argon in the Interactions API with day-one support across the google-genai SDKs, including goal runs and multi-agent orchestration.



KEY FACTS

ERQA BENCHMARK

87%

SYNTHETIC EPISODES GENERATED

4,500+

REAL-WORLD SUCCESS RATE BEFORE/AFTER

56% to 75%

COMPLETION TIME BEFORE/AFTER

35s to 24s

TASK CURRICULUM SIZE

77 tasks

AVAILABILITY

gemini-4-argon in Interactions API

WHY IT MATTERS

Argon's combination of ERQA-leading reasoning, sim-to-real data synthesis, and self-improvement loops gives robotics teams a single model that can generate training data, verify controllers in physics, and close the sim-to-real gap without hand-authored pipelines.

[Read the original at deepmind.google](#) →

AI Tools & Ecosystem

3 stories ● 3 medium

04 MEDIUM IMPACT MIT Technology Review

AI Labs Are Buying the Data of Dead Startups to Train Their Models

AI labs are now buying the operational data of failed startups—Slack histories, emails, code—as training material, and a formal resale market has formed around it.

SimpleClosure added data sales to its wind-down services in April via an Asset Hub platform that lets founders sell code and workspace data from Google, Notion, Figma, and Slack. Buyers include foundation AI labs, vendors building simulated workspaces for AI training, and mature tech companies training their own models. The company has handled over 150 deals and expects to return more than \$10 million to startups by the end of 2026. CEO Dori Yona says 47% of SimpleClosure deals now include code or data sales, up from a practice he describes as once rare and "somewhat taboo." Rival Sunset runs a similar resale business, valuing assets by head count, years in operation, data volume, and tool usage.

Demand emerged last year as labs hit the limits of public and synthetic datasets. Surge AI buys archives to build "reinforcement-learning gyms" where agents practice tasks such as debugging complex IT architectures or drafting appellate briefs. Nick Heiner, who leads that work, says a single company's Slack history with variety and depth can be worth more than another firm's Slack, email, and calendars combined. Real archives avoid synthetic-world errors like references to nonexistent dates, but no company has published tests comparing real-data training against alternatives, and pricing reflects that uncertainty—buyers quote wildly different numbers based on "research theses" estimating useful tasks per 100,000 lines of code.

The scale is growing. In August, as part of Spirit Airlines' bankruptcy, Google outbid AI data company Mercor to buy the carrier's data—100 million emails, 500 million Teams messages, and 176,000 employee records—for \$10 million. Bloomberg reported in September that SpaceX's AI unit is weighing purchases from struggling startups. Legal exposure is unresolved: Spirit's flight attendants' union objected on privacy grounds, and attorney Greg Ewing calls such data "legally radioactive." Suppliers and buyers scrub personally identifiable information and exclude HR records, but Ewing notes that emails and calendars still contain doctor appointments, school pickups, and date-night plans. Proceeds typically go to senior creditors and investors, not employees, which Emory law professor Ifeoma Ajunwa describes as "captured capital." No federal law bans the sales.

Founders get little visibility into who buys their data or what they plan to do with it, and little assurance of fair pricing. Data prices will likely stabilize in the coming year, the article notes, unless an AI bubble pops first.

KEY FACTS

SIMPLECLOSURE DEALS INCLUDING DATA SALES

47%

EXPECTED RETURN TO STARTUPS BY END OF 2026

\$10M

SPIRIT EMAILS IN SALE

100M

ASSET HUB DEALS HANDLED

150+

SPIRIT AIRLINES DATA SALE PRICE

\$10M

SPIRIT TEAMS MESSAGES IN SALE

500M

WHY IT MATTERS

If you are building or buying training data, defunct-company archives are becoming a distinct supply channel with real volume—but with unresolved privacy risk, unproven performance gains, and no published benchmarks against synthetic alternatives.

[Read the original at MIT Technology Review →](#)

05 **MEDIUM IMPACT** ollama.com

Ollama Goes to Fly.io: Sprites Give AI Agents a Private Linux VM With Secret-Free Model Access

Ollama Cloud models are now reachable from Fly.io Sprites through a gateway that keeps API keys out of the sandbox entirely.

Fly.io Sprites are full Linux virtual machines with persistent disks, designed as sandboxes for coding agents. Claude Code, Codex, Gemini CLI and Cursor are supported out of the box. The new Ollama Connector lets code running inside a Sprite call Ollama Cloud models — kimi-k3, glm-5.3 and deepseek-v4.1-flash are named in the announcement — without the Sprite ever holding an API key.

The architecture is the substantive part. You add an Ollama API key to your Fly.io organization; Fly.io encrypts it and never returns it through the API. When code inside a Sprite calls Ollama, the request goes through a Fly.io gateway. The gateway checks which Sprite is calling, then attaches your key to the outbound request. The Sprite itself only knows a gateway URL and a connection ID. Access control defaults to deny: new connectors block every Sprite until you explicitly grant access by label, by name prefix, or to all Sprites. Calls only work from inside a Sprite, so the endpoint cannot be used from outside the sandbox.

Client integration is deliberately thin. Any OpenAI SDK works by setting `base_url` to the gateway and using a placeholder `api_key`, since there is no Authorization header — the gateway adds the real key. Ollama's native `/api/*` routes also work through the gateway. For coding agents, Sprites ship with a skill that tells Claude Code, Cursor, Codex and Gemini CLI about Ollama, so if the Sprite has access the agent can use the models with no additional setup.

The announcement also points to a fuzzing workflow that chains models: `nemotron-3-nano:30b` generates malformed inputs, a Sprite runs them against a program, and `kimi-k3` examines the crashes. Switching models is a matter of changing the model field in the request.

What is genuinely new here is the secret-free model access pattern for agent sandboxes. Ollama Cloud itself is not new, and Fly.io Sprites existed before this connector. The connector is the increment: it removes the API

key from the environment where untrusted agent code runs, which addresses a real failure mode when agents execute arbitrary commands.

KEY FACTS

MODELS AVAILABLE

kimi-k3, glm-5.3, deepseek-v4.1-flash

SUPPORTED AGENTS

Claude Code, Codex, Gemini CLI, Cursor

FUZZING SMALL MODEL

nemotron-3-nano:30b

SANDBOX TYPE

Full Linux VM with persistent disk

ACCESS DEFAULT

Block every Sprite by default

WHY IT MATTERS

If you run coding agents in sandboxes, this removes the API key from the environment the agent can read, which closes a direct exfiltration path without changing how your client code calls the model.

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

06

MEDIUM IMPACT

TechCrunch

Google Play Opens to Private, Vibe-Coded Apps That Never Hit the Store

Google Play now supports private, personal Android apps that never appear in search, listings, or anyone else's device.

Google announced Tuesday that Android users can build an app in Google AI Studio and launch it on their own device through Google Play using a standard Google account, with the app remaining completely private. These personal apps are excluded from Play search and listing pages and are not shared with anyone except the creator. They must still comply with existing Google Play policies, but they bypass the public publishing pipeline that previously made any native Android app visible to the Play Store's roughly 3 billion users. The feature launches initially in the U.S. for users 18 and older.

The change targets a specific friction point: people experimenting with AI-generated, "vibe-coded" apps for personal needs have often defaulted to web apps because deploying a native Android app meant publishing publicly. Google frames the move as a shift from "there's an app for that" to "there's an app for me," in the words of Sam Bright, VP and GM for the Google Play and Developer Ecosystem. In the months ahead, Google says users will be able to share these private apps with a small trusted network — family, friends, or a child's T-ball league — without going public. If an app gains traction, there will be a path to convert it into a published Play Store listing after trust and safety reviews.

Alongside the private-app feature, Google announced several adjacent updates. A personalized game recommendation engine called Game Finder will draw on a user's play history plus Google's knowledge from its game community of 1.2 billion players, launching first in the U.S. in English. Users will also be able to curate their own game lists, post questions in a new Community Posts section, and join 48-hour gaming challenges called Sprints. The Play Store's AI assistant, Ask Play, is being updated to offer comparisons between similar apps, and Play apps and games will be more deeply integrated into Gemini conversations.

Google also added an age-range feature that lets parents share a child's age range, rather than an exact birthdate, with developers.

For practitioners, the most consequential piece is the private distribution path. It removes the public-publishing barrier for personal Android utilities and lowers the cost of testing native Android builds built with AI tooling, while keeping policy compliance as a constraint.

KEY FACTS

INITIAL AVAILABILITY

U.S., age 18 and up

GAME COMMUNITY SIZE

1.2 billion players

BUILD TOOL

Google AI Studio

PLAY STORE USER BASE

3 billion

SPRINT DURATION

48 hours

WHY IT MATTERS

Developers experimenting with AI-generated apps can now iterate on native Android builds for personal use without exposing them to the public Play Store, and a future sharing path will let small trusted groups use the same app before any public launch.

[Read the original at TechCrunch →](#)

AI Applications & Industry

4 stories ● 4 medium

07 MEDIUM IMPACT TechCrunch

Character.AI Files for Chapter 11 Five Years In

Character.AI has filed for Chapter 11 bankruptcy five years after founding, with a \$20 million Disney asset sale and no buyer yet for its platform.

The company disclosed in court documents that it holds assets and liabilities between \$50 million and \$100 million. Disney has agreed to purchase some assets for \$20 million along with a non-exclusive perpetual license for Character.AI's technology. Transition services will be provided by former Character.AI employees who joined Disney last month alongside former CEO Karandeep Singh Anand. CEO Colin Adams said the sale process is expected to close in late November or early December, but noted there is no bid yet for the company's remaining assets or a buyer. If no actionable proposal emerges, the platform will be wound down.

Character.AI's trajectory was fast and volatile. Founded in late 2021 by former Google engineers Noam Shazeer, a co-author of the Transformer paper, and Daniel De Freitas, the startup reached unicorn valuation by September 2022 after a \$150 million fundraising. Users had exchanged more than 2 billion messages with its bots in five months, and active users averaged over two hours daily on the platform. In August 2024, Google struck a \$2.7 billion deal to license the startup's models, and both founders returned to Google. The platform recorded 20 million monthly active users by October 2024.

The company's troubles accelerated after a lawsuit following the suicide of a 14-year-old Florida boy whose mother said he became obsessed with a chatbot on the platform. Texas subsequently opened an investigation into child privacy and safety compliance. Character.AI settled some lawsuits, added safety features, and banned minors. In June of last year, Anand, a former Brex and Meta executive, became CEO, and the company expanded into AI video generation, social feeds, interactive storytelling, and microdramas. Those changes did not restore sustainability.

The bankruptcy underscores the unit economics problem for high-engagement consumer AI. Character.AI was among the top apps by time spent, according to Sensor Tower, but inference costs scale with usage, intensifying monetization pressure. The company had raised over \$193 million from investors including Lightspeed, SV Angel, Greycroft, and a16z, according to PitchBook.

KEY FACTS

ASSETS AND LIABILITIES

\$50M-\$100M

GOOGLE MODEL LICENSE DEAL (2024)

\$2.7B

MONTHLY ACTIVE USERS (OCT 2024)

20M

DISNEY ASSET PURCHASE

\$20M

TOTAL FUNDING RAISED

\$193M+

MESSAGES EXCHANGED IN 5 MONTHS (2022)

2B+

WHY IT MATTERS

High-engagement consumer AI platforms face a structural cost problem: usage drives inference spend faster than revenue. Builders should treat engagement metrics as a liability unless paired with a monetization path that scales with compute.

[Read the original at TechCrunch →](#)

08 **MEDIUM IMPACT** TechCrunch

Amex's Intent Intelligence Framework Wants AI Agents to Sign Their Receipts

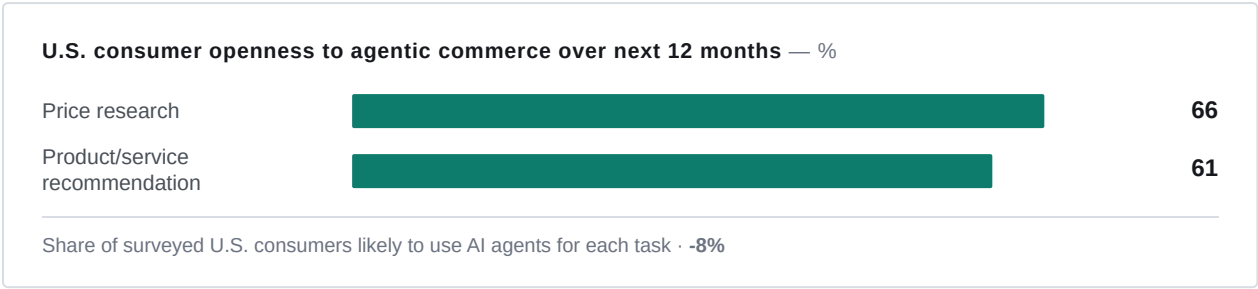
American Express is shipping an intent-capture and authentication layer designed to sit between a customer's instructions and an AI agent's checkout, so the agent's action can be checked against the original authorization before money moves.

The Intent Intelligence Framework, announced ahead of TechCrunch Disrupt 2026, is built around a Create Intent API that captures the customer's purchase instructions plus constraints such as merchant name, maximum budget, and product preferences. The customer then authenticates and signs off on the captured intent. From that, Amex issues an Intent ID and a Proof of Intent Token containing the high-level constraints; the token is exchanged for a secure digital payment credential that the agent uses at checkout instead of the customer's actual card number. The same ID and token can be reused downstream for fraud prevention and dispute resolution.

Amex's VP of Agentic Partnerships and Strategy, Hannah Bozian, says the differentiator is pulling intent earlier into the transaction flow — before risk occurs rather than after. Her example: a customer asks for white Nike sneakers in size 9, the agent submits a purchase for Adidas, and Amex blocks the transaction because the request diverges from the captured intent. The framework is designed to be model-agnostic, supporting whatever AI model collects the intent, and Amex says it has not yet seen an increase in disputes from agentic commerce but is building proactively as the category scales.

The framework is part of the ACE Developer Toolkit, which launched in April and now includes the intent services plus agentic tokens carrying Amex payment credentials. Amex is already in discussions with payments-ecosystem partners about implementing ACE. Bozian also points to a future where agents act more autonomously — reordering household staples where only price and size matter — and argues that an in-the-loop check becomes more necessary as consent moves from per-purchase to standing authorization. A log of which agent authorized what, and when, could also help address agent-side fraud.

For practitioners, the notable shift is architectural: intent becomes a signed, tokenized artifact that travels with the transaction, rather than a prompt that disappears once the agent acts. The framework is Amex-specific today, but the company is positioning it as an industry standard for agentic transactions.



KEY FACTS

SURVEYED U.S. CONSUMERS LIKELY TO USE AI AGENTS FOR PRICE RESEARCH

66%

ACE DEVELOPER TOOLKIT LAUNCH

April

SURVEYED U.S. CONSUMERS LIKELY TO USE AI AGENTS FOR PRODUCT RECOMMENDATIONS

61%

FRAMEWORK ANNOUNCEMENT VENUE

TechCrunch Disrupt 2026, October 13

WHY IT MATTERS

If intent capture and signed tokens become standard in agentic commerce, builders of shopping agents will need to emit structured, constrainable purchase intent and handle pre-transaction rejection — not just post-hoc consent dialogs.

[Read the original at TechCrunch →](#)

09 **MEDIUM IMPACT** Google DeepMind

DeepMind and Singapore Fine-Tune a Gemma Model to Catch Phone Scams on-Device

Google DeepMind and GovTech Singapore have fine-tuned a lightweight Gemma model to detect scam conversations entirely on-device, without either party ever seeing user messages.

The collaboration, presented at Singapore International Cyber Week (SICW 2026), targets a problem that now accounts for over half of Singapore's reported crimes and more than S\$4 billion in losses since 2019. Complex social engineering scams — where attackers build convincing conversational rapport before requesting money — now represent up to 80% of all losses. The model was fine-tuned on local scam conversations including impersonations of public institutions, banks, and local figures, plus trending regional schemes, specifically because generic models struggle with local deception techniques.

Deployment uses a two-stage pipeline to manage battery efficiency on modern smartphones. Stage 1 runs a lightweight classifier that screens inbound messages with minimal power, flagging potential threats. Stage 2 performs in-context analysis with the fine-tuned Gemma model over a larger conversation snapshot to confirm a threat or discard a false alarm. When manipulative patterns are confirmed, the system triggers a real-time on-device warning advising the user to pause before sharing sensitive information or funds.

Governance constraints shaped the build. Fine-tuning occurred inside GovTech Singapore's managed GCP environment — a sovereign sandbox keeping training data isolated. Training strictly excluded private or personally identifiable information, relying on anonymised datasets derived from scam reports and intelligence feeds. The model was stress-tested against synthetic dialogues and local corpora covering diverse scam categories and evolving impersonation tactics. Rollout within applications will be phased in partnership with the Singapore Government, with the model accessing on-device conversations only with user consent under each application's terms of use.

KEY FACTS

LOSSES SINCE 2019

S\$4 billion

SOCIAL ENGINEERING SHARE OF LOSSES

up to 80%

PIPELINE STAGES

2 (lightweight classifier + Gemma in-context analysis)

SHARE OF REPORTED CRIMES

over half

BASE MODEL

Gemma (lightweight)

PRESENTED AT

Singapore International Cyber Week (SICW 2026)

WHY IT MATTERS

This is a concrete reference architecture for deploying fine-tuned LLM safety classifiers under end-to-end encryption constraints — the two-stage edge pipeline and sovereign-sandbox training setup are directly reusable patterns for anyone building privacy-preserving on-device detection.

[Read the original at Google DeepMind →](#)

10

MEDIUM IMPACT

TechCrunch

Former Neuralink President's PRIMA Chip Restores Reading to a Stargardt Patient

Science Corporation has implanted its PRIMA vision-restoring chip in a Stargardt disease patient, who progressed from seeing a single letter to reading a word within two rehabilitation sessions.

CEO Max Hodak announced the milestone at TechCrunch Disrupt, framing PRIMA as part of the company's broader brain-computer interface strategy. "The retina is part of the central nervous system," Hodak said, grouping cochlear implants, retinal prostheses, and motor prostheses as variants of BCI. The implant took place in May as part of an ongoing trial at Sydney Eye Hospital in Australia.

The PRIMA device is a 2 mm by 2 mm chip placed at the back of the eye where photoreceptors have degenerated. Patients wear camera-equipped glasses that transmit the visual scene to the implant, which acts as artificial photoreceptors. Hodak was explicit about current limitations: the field of view is "like looking through a straw," black and white, high-contrast, and sharp enough for only a couple of letters at a time. The next-generation device, expected in humans within the next year, will require a larger implant that accommodates the eye's curvature and smaller electrodes to approach native acuity. Hodak projected that within five years the company could get "pretty close to native acuity."

Stargardt disease destroys the outer retina's light-sensing cells while often sparing the inner retina that relays signals to the brain. Roughly 150,000 people worldwide have lost vision to the disease, which is currently

considered incurable. PRIMA is already commercially available in Europe for geographic atrophy, an advanced form of age-related macular degeneration; the Stargardt trial extends that approach to a new indication.

Science Corporation, founded by Hodak in 2021 after leaving Neuralink, is also developing a BCI sensor that rests on top of the brain inside the skull rather than penetrating brain tissue. Hodak said the first clinical indication, not yet announced publicly, is a neurodegenerative disease, with stroke as another target. His longer-term framing: "the long-term mission of the company is to understand how the brain creates the mind and to learn to build engineered brains."

KEY FACTS	
IMPLANT SIZE	STARGARDT PATIENTS WORLDWIDE
2 mm x 2 mm	150,000
FIRST IMPLANT DATE	REHAB MILESTONE
May (Sydney Eye Hospital)	Letter in session 1, word in session 2
NEXT-GEN HUMAN TRIALS	ACUITY PROJECTION
Within the next year	Near-native within 5 years

WHY IT MATTERS

PRIMA's Stargardt result validates a non-penetrating, commercially precedented approach to neural interfaces, suggesting that retinal prosthetics can serve as a nearer-term revenue path while more invasive BCI work matures. For practitioners in medical AI and neurotech, the constraints Hodak names—field of view, electrode size, eye curvature—define the concrete engineering problems that will determine whether the five-year acuity projection holds.

[Read the original at TechCrunch](#) →

Sources

- 01 Anthropic Replaced Git: 2,500 Daily Commits Became 30,000, So They Rewrote Source Control
<https://simonwillison.net/2026/Oct/13/redesigning-source-control-at-anthropic/>
- 02 Liquid AI Ships LFM2.5-24B-A2B: a Reasoning Agent Model With 2B Active Parameters
<https://huggingface.co/blog/LiquidAI/lfm25-24b-a2b>
- 03 Gemini 4 Argon for Robotics: 87% on Embodied Reasoning, Ahead of Every Frontier Model
<https://deepmind.google/blog/gemini-4-argon-for-robotics/>
- 04 AI Labs Are Buying the Data of Dead Startups to Train Their Models
<https://www.technologyreview.com/2026/10/13/1145755/ai-labs-buying-data-from-dead-startups-to-train-models/>
- 05 Ollama Goes to Fly.io: Sprites Give AI Agents a Private Linux VM With Secret-Free Model Access
<https://ollama.com/blog/fly-io-sprites>
- 06 Google Play Opens to Private, Vibe-Coded Apps That Never Hit the Store
<https://techcrunch.com/2026/10/13/google-play-now-supports-private-vibe-coded-apps-built-just-for-you/>
- 07 Character.AI Files for Chapter 11 Five Years In
<https://techcrunch.com/2026/10/13/just-5-years-in-interactive-chatbot-platform-character-ai-files-for-bankruptcy/>
- 08 Amex's Intent Intelligence Framework Wants AI Agents to Sign Their Receipts
<https://techcrunch.com/2026/10/13/amex-introduces-a-framework-to-prevent-ai-agents-from-making-unauthorized-purchases/>
- 09 DeepMind and Singapore Fine-Tune a Gemma Model to Catch Phone Scams on-Device
<https://deepmind.google/blog/on-device-ai-scam-detection-mddi/>
- 10 Former Neuralink President's PRIMA Chip Restores Reading to a Stargardt Patient
<https://techcrunch.com/2026/10/13/science-corporation-implants-its-vision-restoring-chip-in-its-first-stargardt-disease-patient/>

About this document. Every story in the 14 October 2026 New Horizon AI Digest, reported at length. Each entry is written from the publisher's own article text; where a source could not be retrieved the entry is explicitly marked and kept short rather than padded.

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

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

Generated 2026-10-14T05:33:48Z · model deepseek-v4-pro:cloud · New Horizon, new-horizon.tech · To unsubscribe or manage your subscription, see the footer of the daily email.