

tech Gazette

The Monthly

October 2026

Software and AI, straight from the source. Through Oct 17; the month is still running.

The stories that moved it	2
What stops working next	3
Exploited, and down	4
Models, and what developers starred	5
Research worth the month	6

7,872 PIECES PRINTED	1,203 SOURCES	22 EXPLOITED VULNERABILITIES	55 DEADLINES AHEAD
--	-------------------------------------	--	--

Every line in this report is a piece a primary source published, in its words, with its link: the labs, companies, projects and people who did the thing. No algorithm chose it; nothing here is sponsored.

THE MONTH

The stories that moved it

The Weekly Reviews' top stories of the month, ranked by how many sources reported them and how much they change. Each links to where it was first said.

1**Claude Sonnet 5.5**

CLAUDE API RELEASE NOTES · AND 7 MORE SOURCES

We've launched Claude Sonnet 5.5 (claude-sonnet-5-5). It's available on the Claude API, Claude in Amazon Bedrock, Claude Platform on AWS, Claude on Google Cloud, and Claude in Microsoft Foundry. For its context window, o...

2**Gemini 4 Argon (gemini-4-argon) generally available (GA)**

GEMINI API CHANGELOG · AND 12 MORE SOURCES

Gemini 4 Argon (gemini-4-argon) generally available (GA): Released Gemini 4 Argon (gemini-4-argon), our most capable model built for complex coding, agentic, math, and science tasks. For API changes, pricing, and migrati...

3**Released GPT-6.1 Sol (gpt-6.1-sol) for complex coding and professional work at a lower cost than GPT-6 Astra.**

OPENAI API CHANGELOG · AND 7 MORE SOURCES

Feature · gpt-6.1-sol · v1/responses · v1/chat/completions. Released GPT-6.1 Sol (gpt-6.1-sol) for complex coding and professional work at a lower cost than GPT-6 Astra. Standard pricing per 1M tokens for prompts with up...

4**Claude Haiku 5.5 launch**

CLAUDE RELEASE NOTES · AND 9 MORE SOURCES

Claude Haiku 5.5 launch We just launched Claude Haiku 5.5, the cheapest, fastest, and most capable small model we've released, designed for high-volume, cost-sensitive tasks. For more information, see our blog post: Claud...

5**Added Ultrafast mode for GPT-6 Astra in the Responses API.**

OPENAI API CHANGELOG · AND 5 MORE SOURCES

Feature · gpt-6-astra · v1/responses. Added Ultrafast mode for GPT-6 Astra in the Responses API. Use gpt-6-astra with service_tier: "ultrafast" to reduce the time between generated output tokens. It is available to API c...

6**EmbeddingGemma 2: an open, lightweight multimodal embedding model**

GOOGLE DEEPMIND · AND 3 MORE SOURCES

7**Added Ultrafast mode for GPT-6.1 Sol in the Responses API.**

OPENAI API CHANGELOG · AND 3 MORE SOURCES

Feature · gpt-6.1-sol · v1/responses. Added Ultrafast mode for GPT-6.1 Sol in the Responses API. Use gpt-6.1-sol with service_tier: "ultrafast" to reduce the time between generated output tokens. It is available to all A...

8**Gemini Nano Banana 2.1 generally available (GA)**

GEMINI API CHANGELOG · AND 2 MORE SOURCES

Gemini Nano Banana 2.1 generally available (GA): Released Gemini Nano Banana 2.1 (gemini-nano-banana-2.1), the latest high-efficiency image generation and conversational editing model. An update to Nano Banana 2 (gemini-...

9**Introducing Mistral Large 4**

MISTRAL AI · AND 2 MORE SOURCES

10**Added computer use to the Agents API.**

OPENAI API CHANGELOG

Feature. Added computer use to the Agents API.

ACT ON THIS

What stops working next

Models being shut down and versions reaching end of support in the 90 days after Oct 17, with the replacement each vendor recommends. Only dates they state.

WHEN	WHAT STOPS	USE INSTEAD	STATED BY
Oct 20	gemini-2.5-flash shuts down	gemini-3.8-flash or gemini-3.5-flash-lite or gemini-3.1-flash-lite	Google Cloud (Vertex AI)
Oct 20	gemini-2.5-flash-lite shuts down	gemini-3.8-flash or gemini-3.1-flash-lite or Gemma 4	Google Cloud (Vertex AI)
Oct 20	gemini-2.5-pro shuts down	gemini-3.8-flash or gemini-3.5-flash	Google Cloud (Vertex AI)
Oct 22	gemini-omni-flash-preview shuts down	gemini-omni-1.1-flash	Google (Gemini API)
Oct 22	veo-3.1-fast-generate-preview shuts down	gemini-omni-1.1-flash	Google (Gemini API)
Oct 22	veo-3.1-generate-preview shuts down	gemini-omni-1.1-flash	Google (Gemini API)
Oct 22	veo-3.1-lite-generate-preview shuts down	gemini-omni-1.1-flash	Google (Gemini API)
Oct 23	deep-research-pro-preview-12-2025 shuts down	deep-research-preview-04-2026	Google (Gemini API)
Oct 23	ft-babbage-002 shuts down	gpt-5.6-terra	OpenAI
Oct 23	ft-davinci-002 shuts down	gpt-5.6-terra	OpenAI
Oct 23	ft-gpt-3.5-turbo shuts down	gpt-5.6-terra	OpenAI
Oct 23	ft-gpt-4 shuts down	gpt-5.6-sol	OpenAI
Oct 23	ft-gpt-4.1-nano-2025-04-14 shuts down	gpt-5.6-luna	OpenAI
Oct 23	ft-o4-mini-2025-04-16 shuts down	gpt-5.6-terra	OpenAI
Oct 23	gpt-3.5-turbo-0125 gpt-3.5-turbo, gpt-3.5-turbo-completions shuts down	gpt-5.6-terra	OpenAI
Oct 23	gpt-4-0613 gpt-4, gpt-4-0613-completions, gpt-4-completions shuts down	gpt-5.6-sol	OpenAI
Oct 23	gpt-4-1106-preview shuts down	gpt-5.6-sol	OpenAI
Oct 23	gpt-4-turbo gpt-4-turbo-2024-04-09, gpt-4-turbo-completions shuts down	gpt-5.6-sol	OpenAI

AND 37 MORE DEADLINES AT [TECHGAZETTE.COM.BR/DEADLINES/](https://techgazette.com.br/deadlines/)

SECURITY DESK

Exploited, and down

Exploited in the wild

[CVE-2026-104935: Ivanti Connect Secure Authentication Bypass by Spoofing Vulnerability](#)

CISA KEV · OCT 13

[CVE-2026-96951: Microsoft Windows Command Injection Vulnerability](#)

CISA KEV · OCT 13

[CVE-2026-105134: Ahsay AhsayCBS Command Injection Vulnerability](#)

CISA KEV · OCT 14

[CVE-2026-76519: Cisco Secure Email Gateway Relative Path Traversal Vulnerability](#)

CISA KEV · OCT 14

[CVE-2026-105133: Ahsay AhsayCBS Improper Authentication Vulnerability](#)

CISA KEV · OCT 14

[CVE-2026-0322: Palo Alto Networks PAN-OS Code Injection Vulnerability](#)

CISA KEV · OCT 14

[CVE-2026-21589: Atlassian Multiple Products Files or Directories Accessible to External Parties Vulnerability](#)

CISA KEV · OCT 15

[CVE-2026-102764: BeyondTrust Remote Support Authentication Bypass by Spoofing Vulnerability](#)

CISA KEV · OCT 16

[CVE-2025-61624: Fortinet Multiple Products Path Traversal Vulnerability](#)

CISA KEV · OCT 16

[CVE-2026-88779: Citrix NetScaler Improper Restriction of Operations within the Bounds of a Memory Buffer Vulnerability](#)

CISA KEV · OCT 4

[CVE-2015-5477: ISC BIND Data Processing Errors Vulnerability](#)

CISA KEV · OCT 8

[CVE-2016-3081: Apache Struts Command](#)

The longest incidents

[Service Disruption Impacting Chase connections](#)

PLAID · OCT 11

[Issues with spin up node in google asia-southeast1 region](#)

AIVEN · OCT 12

[Email API messages rejected with gw_reject//injection_failed](#)

TELNYX · OCT 12

[Increased build and deployment failures](#)

VERCEL · OCT 12

[Low Code Automation - Builds is facing issues](#)

BROWSERSTACK · OCT 12

[Model Performance or Availability Issue](#)

GROQ · OCT 12

[\[INCIDENT\]](#)

<https://www.contentful.com/developers/docs/is not working>

CONTENTFUL · OCT 12

[Low Code Automation - Builds is facing issues](#)

BROWSERSTACK · OCT 13

[2026-October-13 Service Incident](#)

SAUCE LABS · OCT 13

[Multiple US endpoints experiencing timeouts and latency](#)

ELEVENLABS · OCT 13

WHAT SHIPPED, WHAT ROSE

Models, and what developers starred

Models released (106)

New model: DeepSeek: DeepSeek V4.1 Flash 1017

OPENROUTER · OCT 17

New model: Elorian: Glimpse

OPENROUTER · OCT 17

Open model: deepseek-ai/DeepSeek-V4.1-Flash-1017

HUGGING FACE HUB · OCT 17

New model: Qwen: Qwen Image 2.1 Turbo

OPENROUTER · OCT 16

New model: Qwen: Qwen Image 2.1 Pro

OPENROUTER · OCT 16

New model: Perplexity: Decider Flash V1.2

OPENROUTER · OCT 16

New model: Perplexity: Decider V1.2

OPENROUTER · OCT 16

Open model: nvidia/NVIDIA-Nemotron-Labs-Dory1.0-4B-NVFP4

NVIDIA · OCT 16

Open model: nvidia/NVIDIA-Nemotron-Labs-Dory1.0-4B-BF16

NVIDIA · OCT 16

New model: Perplexity: Decider Flash Latest

OPENROUTER · OCT 16

Open model: tencent/Qwen3-8B-DSpine-b16

TENCENT HUNYUAN · OCT 16

New model: Reflection: Beam 1.0 501B A23B

OPENROUTER · OCT 16

AND 94 MORE MODELS AT
[TECHGAZETTE.COM.BR/MODELS/](https://techgazette.com.br/models/)

Rising on GitHub

+248,962 [morluto/rea](#) TYPESCRIPT

Reverse engineer anything with agents, from app behavior down to native binaries.

+41,883 [boykopovar/AnyPS5](#) C++

Tool for automatic PS5 executables porting to Linux and Windows

+21,116 [mattpocock/skills](#) SHELL

Skills for Real Engineers. Straight from my .agents directory.

+20,267 [storytold/artcraft](#) RUST

ArtCraft is an intentional crafting engine for artists, designers, and filmmakers

+16,775 [debpalash/VoiceStudio](#) PYTHON

VoiceStudio is the open-source, fully-local ElevenLabs alternative — voice cloning, voice design, video dubbing, dictation, transcription &

+14,507 [vectorize-io/hindsight](#) PYTHON

Hindsight: Agent Memory That Learns

+12,693 [heygen-com/hyperframes](#)

TYPESCRIPT

Write HTML. Render video. Built for agents.

+10,722 [paperclipai/paperclip](#) TYPESCRIPT

The open-source app everyone uses to manage agents at work

+9,024 [cathrynlavery/diagram-design](#) HTML

Editorial diagram design for Claude Code, Codex, GitHub Copilot, Cursor, Factory Droid, and Pi. 44 diagram types. Self-contained HTML + SVG.

+7,510 [NationalSecurityAgency/ghidra](#) JAVA

Ghidra is a software reverse engineering (SRE) framework

BY STARS GAINED: TYPESCRIPT 299,002 · PYTHON 53,121 · C++ 41,883 · RUST 26,121 · SHELL 21,116 · HTML 9,024

FROM THE PAPERS

Research worth the month

1,882 papers printed this month. A selection, each with its own abstract's words; the week's papers by field are at techgazette.com.br/research.

BHunter: Generating Blind-Spot Driving Scenarios for Robustness Testing of Perception Systems

ACM TOSEM · OCT 11

Distilling Transitional Patterns to Multimodal Large Language Models for Session-based Recommendation

ACM TOIS (INFORMATION SYSTEMS) · OCT 11

Privacy Risks and Leakage Pathways in Agentic and Multi-Agent AI Systems

ACM TOPS (PRIVACY AND SECURITY) · OCT 10

ChronoSteer: Bridging Large Language Model and Time Series Foundation Model via Synthetic Cross-Modal Alignment Dataset

ACM TKDD (KNOWLEDGE DISCOVERY FROM DATA) · OCT 11

Exploring Backdoor Vulnerabilities and Fairness Challenges in Generative Models: A Survey

ACM COMPUTING SURVEYS · OCT 12

One Word Opens the Gate: The Option-Channel Attack on Typed Decision Models as Agent Guardrails

ARXIV CS.AI · OCT 8

Revealing and Mitigating the Impact of LLM-Generated Content on Retrieval-Augmented Generation Systems

ACM TKDD (KNOWLEDGE DISCOVERY FROM DATA) · OCT 9

ARC: A Reasoning Recipe for Robot Foundation Models

ARXIV CS.RO · OCT 8

Reasoning-Informed Visual Editing

ARXIV (VISION, RETRIEVAL & HCI) CS.CV · OCT 8

Beyond Spatio-Temporal Priors: A Generalizable Approach for Dense Correspondence Matching

ARXIV CS.CV · OCT 8

tech Gazette

The paper, every morning. Software and AI news from the people who made it. Free, no algorithm, no account. It ends.

Read it at techgazette.com.br, or get it in your inbox:
techgazette.com.br/newsletter.

The Monthly is made from the paper's own archive: 7,872 pieces from 1,203 sources in October 2026. Titles are the sources' own. Share it freely, whole and unchanged.