

The week agents moved forward.

Models, coding agents and the systems that make them useful.

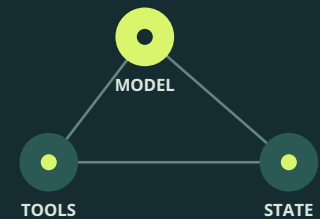
14 OCT 2026

First edition: 7-14 Oct | Research cutoff 16:47 IST

THE SIGNAL

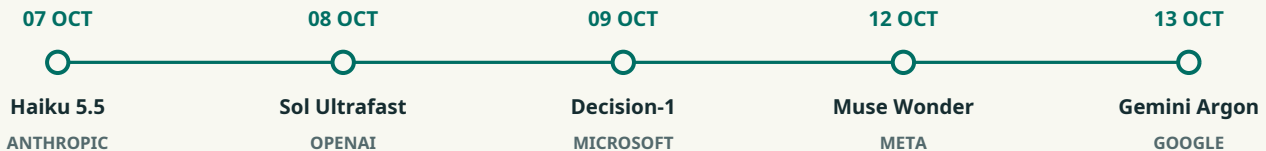
Longer runs. Smarter control.

New models emphasize sustained work; agent tools are improving context sharing, permissions and measurement.



01 MODEL RADAR

Launches + one serving-mode update



GOOGLE / 13 OCT

Gemini 4 Argon

A model built for long coding tasks, including migrations across a codebase.

Access: Paid Gemini API; Antigravity for Google AI Ultra subscribers.

Try: Evaluate a multi-file migration in your own repo.

[Official announcement >](#)

META / 12 OCT

Muse Wonder + Wonder Full

Models for coding and personal tasks that recall context and work through blockers.

Access: Wonder: Meta API + Muse Code. Full: Muse Code subscription.

Try: Check context continuity and recovery after failure.

[Official announcement >](#)

02 RESEARCH WORTH READING

New preprints; not peer-reviewed

13 OCT 2026 / V1

arXiv preprint

Same Attack, Different Judge

Tests how Jev, LLM judges and safety classifiers handle the same adversarial inputs across moderation, routing and phishing tasks.

Takeaway: Test an agent's routing or safety judge on your own workload, including benign inputs and false alarms.

[Original publication >](#)

09 OCT 2026 / V1

arXiv preprint

AgentSWE

Benchmarks whether coding agents can create, edit and improve other agents that meet real workflow requirements.

Takeaway: Freeze the delivered agent and test unseen acceptance cases, including its runtime behavior.

[Original publication >](#)

What to watch. What to try.

Four repositories, two ecosystem shifts and one useful experiment.

03 GITHUB WATCHLIST

Tryable open-source software; licences shown

01

github/gh-aw

MIT / Repository automation

[Repo](#) [Release notes >](#)

14 OCT / v0.91.12 / PRE-RELEASE

Tighter MCP controls and strict checks for unsafe pull_request_target roles. Model-routing rules also change. Review affected workflows before recompiling.

Try: Markdown task -> GitHub Actions agent.

02

OpenHands/ OpenHands

MIT / Agent control center

[Repo](#) [Release notes >](#)

12 OCT / v1.27.1 / UPDATE

Agent Canvas coordinates coding agents and automations. v1.27.0 adds OpenCode and Pi as local ACP agents; v1.27.1 fixes Docker bridge and theme issues.

Try: self-host a control center for your agents.

03

anomalyco/ opencode

MIT / Coding agent

[Repo](#) [Release notes >](#)

11 OCT / v1.19.0 / UPDATE

Adds explicit upgrade paths toward OpenCode 2 and tighter terminal-session checks. The /update command installs updates; automatic installation stops.

Try: local terminal agent with your provider.

04

letta-ai/ letta-code

Apache-2.0 / Stateful harness

[Repo](#) [Release notes >](#)

13 OCT / v0.34.13 / NEW CAPABILITY

A stateful agent can now delegate coding to an explicit OpenCode worker. OpenCode must be installed; the new webhook skill requires Letta Cloud.

Try: local mode with your model provider.

04 BEYOND CODING

What changed + comparison with existing tools

NOTION 3.8 UPDATE / 13 OCT

Shared context, visible agent edits

An existing-product update: per-person MCP revocation, column access controls and agent-attributed edit history. **Why it matters:** restrict sensitive fields and review agent changes. Column permissions require Business or Enterprise.

[Official update >](#)

GOOGLE + ANTHROPIC / 08-09 OCT

Cloud agents: what is different?

Both run agent work in the cloud. Google emphasizes an enterprise assistant using Gemini + Claude. Anthropic runs custom Claude agents; its 9 Oct update adds dynamic workflows. Google's broad availability is unverified.

[Official update >](#)

[Claude comparison >](#)

TRY GH-AW: TRIAGE THREE SAMPLE ISSUES / 30 MINUTES AFTER SETUP

Give gh-aw a crash report, a setup-instructions issue and a feature request.

Output: One report with categories, suggested priorities and follow-up questions.

Why: Test whether an agent saves time sorting issues; compare its reasoning with your own.

Guide: setup, workflow, commands and scorecard in try-gh-aw-issue-triage.md.

Primary sources linked throughout. Publisher claims are attributed; takeaways are editorial.

First issue establishes a baseline. Future daily editions can report changes since the previous brief.