THE AI TOOLCHAIN

NO. 097

Tail

THE DAILY RELEASE FIREHOSE

PUBLISHED NOVEMBER 21, 2026 · EVERY DAY

The daily firehose — everything the toolchain shipped today, already filtered.

▸ // HOW THIS ISSUE IS MADE SHOW

VIEW

ISSUE VIEW

scroll



SCROLL | EXPAND IN PLACE

INFORMATION DENSITY

everything



BRIEFED | EVERYTHING

Every tool in the issue, in full.

all 454 changes

Do you prefer **scroll**? 👍 👎

▸ // CONTENTS SHOW 101 of 104 tools

▾ // THE BRIEF HIDE SAME ISSUE, SAME PROMPT, TWO WRITERS: GPT-5.5 Opus 5

What stands out across today's releases, grouped by what it lets you do. Every tool named links to its entry below.

Three unrelated things stand out today: Snowflake Cortex AI can now score and tune AI function quality/cost in-database; Kiro moved the

vim ? help

same specs and MCP config across IDE, CLI and web; Brig can force risky GitHub automation into approved microVM-isolated workflows.

EVALUATE

Score and tune AI function behavior where the data already lives

`AI_FUNCTION_EVALUATION` and `AI_FUNCTION_OPTIMIZATION` make quality and cost checks callable inside Snowflake instead of exporting samples into a separate eval stack. Teams can turn model/function changes into measurable database workflows rather than spreadsheet-backed prompt reviews.

Snowflake Cortex AI

BUILD

Carry coding-agent context across local, remote and browser surfaces

Kiro's VS Code extension, CLI and web surface share specs, steering and MCP configuration, while the IDE adds remote SSH, Dev Container and cloud-session editing. That reduces the usual copy/paste drift between an IDE agent, a terminal run and a cloud workspace.

Kiro

GOVERN

Run high-risk automation with explicit approval and a harder sandbox boundary

Brigfiles now define target workflows with approval and config digesting, and `--isolation microvm` enforces microVM isolation. The practical win is that GitHub or repo-mutating automation can be made reviewable and confined instead of depending on every script behaving politely.

Brig

OPERATE

See why distributed AI jobs are stuck without leaving the CLI

Anyscale added commands for events, logs, metrics, node/GPU inspection, task/actor aggregation and scheduler introspection. When a Ray workload stalls or underuses GPUs, operators get the cluster facts in one place instead of stitching together dashboards, logs and node shells.

Anyscale

Does **GPT-5.5** read better? 👍 👎

BUILD

◆ AI Coding Agents

SST OpenCode

Feedback ⓘ changelog · Nov 21, 2026 CHANGELOG ↗

RANK

OpenCode is an open-source AI coding agent for terminal-based development tasks.

> OpenCode's latest changelog centers on background execution — compaction, subagents, and shell command timeouts no longer block the session — alongside an in-app upgrade path to OpenCode 2.0, GPT-6.2 prompt caching across every provider route, and a handful of UI additions.

└──▷ **WHAT SHIPPED · 11 FEATURES** most completely described first ? what's the number?

01 **Adjustable transcript verbosity** Not checked yet NEW 80

Transcript detail is now controlled with `session.verbosity` in your TUI config, or by toggling Verbosity between `medium` and `high` in settings, to see every tool call and thought step.

Increase transcript verbosity to see every tool call and thought step during a session.

```yaml
session:
  verbosity: high
```

— Exact config key and UI toggle values given changelog-20261121-d97dc5b7

02 **Timed-out shell commands move to background** Not checked yet 80

BREAKING

Shell commands that time out now move to the background instead of being killed, so long builds and test runs finish and report back. `timeout: 0` now means the default 120 seconds, a change from prior behaviour where it may have been treated as no timeout.

— Describes behaviour change and exact semantics of `timeout: 0`

changelog-20261121-d97dc5b7

03 **Jump to earlier prompt** Not checked yet NEW 75

The desktop app now supports jumping to any earlier prompt via Jump to message in the command palette or Cmd+Alt+G (Ctrl+Alt+G on Windows and Linux).

— Exact command-palette entry and keyboard shortcuts given changelog-20261121-d97dc5b7

04 **In-app upgrade to OpenCode 2.0** Not checked yet BREAKING

75

OpenCode 1 now offers an in-app upgrade to OpenCode 2.0, downloading it in the background and installing only on your choice of Restart to upgrade, Upgrade on quit, or Later. Plugins made for OpenCode 1 need an update to work with OpenCode 2.0.

— Exact upgrade options and plugin compatibility note given changelog-20261121-d97dc5b7

05 **Background compaction of session summaries** Not checked yet NEW

65

Session summaries can now be written in the background instead of pausing your next prompt, enabled with `compaction.background: true` in your config.

Prevent compaction from pausing your next prompt by letting the summary write itself in the background.

yaml

```yaml
compaction:
  background: true
```

— Names exact config key but mechanism described in one line changelog-20261121-d97dc5b7

06 **GPT-6.2 prompt caching across all routes** Not checked yet IMPROVED

65

GPT-6.2 prompt caching now works on every Responses route — OpenAI (API key and ChatGPT sign-in), GitHub Copilot, Zen, Azure and OpenRouter — by sending the system prompt as a leading developer message so it is cached from the second step on.

— Mechanism and route list named but requires no user action changelog-20261121-d97dc5b7

07 **In-app plugin updates** Not checked yet NEW

60

Plugins can now be updated without leaving the app via Settings › Extensions, clicking Check for updates, then Update on any plugin with a new version.

— Exact navigation path given but thin on mechanism changelog-20261121-d97dc5b7

THINNER COVERAGE BELOW

08 **25 MiB image history cap per session** Not checked yet IMPROVED

50

Sessions with many images now cap history at 25 MiB by replacing the oldest images with a placeholder, preventing oversized requests.

— Names exact limit and mechanism but is automatic, no action needed

changelog-20261121-d97dc5b7

09 **GPT 6.2 Sol and Astra models added** Not checked yet NEW

45

GPT 6.2 Sol and GPT 6.2 Astra, plus their fast variants, now appear in the model picker when signing in with ChatGPT.

– Names models and access path but no further detail changelog-20261121-d97dc5b7

10 **Background subagents by default** Not checked yet IMPROVED

40

Subagents now return immediately and run in the background by default, letting multiple subagents run concurrently while you continue working in the parent session.

– Explains behaviour but no named config or flag changelog-20261121-d97dc5b7

11 **Add comment toolbar in diff and file viewers** Not checked yet

NEW

40

Selecting code in diff and file viewers now shows an Add comment toolbar beside the selection.

– Names the UI control but little mechanism beyond that changelog-20261121-d97dc5b7

└──▷ **BREAKING ON UPGRADE**

! A foreground `timeout: 0` now means the default 120 seconds (previously it may have been treated as no timeout).

! Plugins made for OpenCode 1 need an update to work with OpenCode 2.0.

WAS THIS USEFUL?  

yetone magpie

Feedback ⓘ v0.1.4644 CODE ↗ RANK

Magpie is a macOS menu-bar app, TUI, and CLI that lists every AI coding agent on the machine and switches each one's model, such as Codex on DeepSeek or Claude Code on Kimi, by rewriting only that key in the agent's own config.

> This window magpie added reusable JavaScript skill code to `run_code` (with a `HubSkills` interface and `SaveSkill`), a `GET /v1/magpie/accounts` endpoint for listing and pinning accounts, caller-identity and subagent-identity headers for plugins, a masked-record search endpoint, automatic void-cut recovery for reasoning-only replies, an Anthropic egress region check, grouped relay-model effort variants, unserved routing-group visibility, provider-scoped plugin panels, and a Traditional Chinese site localization.

└──▷ **WHAT SHIPPED · 10 FEATURES** most completely described first ? what's the number?

01 **Skill code support in run_code** Not checked yet NEW 100

Agents can call library skills directly from `run_code` as skills.<name>(args), invoking the skill's `code.js` default export with the supplied args and the same tool access. Scripts run via `run_code` can be saved back to the library as a new named skill using `SaveSkill`, which writes the script as the skill's `code.js` and never overwrites an existing skill or folder. The `Script` struct now carries a `Skills []Skill` field (each with `name` and `code` JSON fields) alongside the existing `Tools []string` and `Args` fields, and a new `HubSkills` interface (`Skills(agent string) []HubSkill`, `SaveSkill(agent, name, about, code string) error`) lets a hub expose per-agent skill code to `run_code`. Skill `code.js` files are normalised for embedding via `SkillSource`, which rewrites `export default` to a return statement so the file body runs as an async function expression inside a script.

— Full mechanism, named structs, interfaces, and call syntax v0.1.4629

02 **Account listing and request pinning** Not checked yet NEW 90

Call `GET /v1/magpie/accounts` to list every configured account and API key across all providers, including accounts that do not appear in the quotas view, with each entry's `pin` value ready for use. Pin any request to a specific account by passing its `<provider>/<id>` pin value in the `X-Magpie-Account` request header.

— Runnable endpoint and header with exact pin format v0.1.4621

03 **Search masked records by value or placeholder** Not checked yet NEW 90

Search masked records by original value or placeholder using the new `q` query parameter on the `masking/records` endpoint, with results capped at `searchMax` (200) characters and matches returned in the `hits` field of each record; records whose values the masker no longer knows are flagged as `unknown` rather than treated as misses.

— Names endpoint, query param, limit and response field v0.1.4628

04 **Caller and subagent identity headers for plugins** Not checked yet

NEW

85

Plugins now receive the requesting agent's identity via X-Magpie-Caller-Agent, X-Magpie-Caller-Name, and X-Magpie-Caller-Via headers on each fetch, requests forwarded through a chain of Remote magpie nodes carry the full hop list in the X-Magpie-Via-Chain header, and magpie's plugin host strips all X-Magpie-Caller-* headers before forwarding requests upstream to vendor endpoints; in addition, plugin fetch requests for Codex subagents now carry the spawned thread ID via X-Magpie-Subagent (populated from Codex's thread-id / turn metadata), and OpenCode subagent requests carry the same header derived from `x-opencode-parent-session-id` / `x-opencode-session-id`, extending the subagent-awareness already present for Claude Code.

— Names all headers and the stripping safeguard Names exact header and source fields across two agents

v0.1.4616 v0.1.4636

05 **Automatic void-cut recovery for reasoning-only replies**

Not checked yet NEW

80

Adds `voidRooms` tracking in the gateway so that when a provider's API default cuts a no-`max_tokens` request to pure reasoning output, the gateway retries that model with its actual output limit (`voidRoom`), falling back to `voidRoomUnknown` (32768 tokens) when the model's own limit is not known.

— Mechanism and exact fallback value, but runs automatically v0.1.4626

06 **Egress region check for Claude requests** Not checked yet NEW

80

magpie now checks whether the outbound exit IP is in an Anthropic-supported region by querying Cloudflare's trace page through each configured proxy, reporting one of three states — `ok`, `unsupported`, or `unknown` — without touching Anthropic or the user's account. It can optionally hold Claude requests when the check finds an unsupported region, acting as a canary for a dropped proxy or a VPN node in the wrong country.

— Names the three states and mechanism, no config key given v0.1.4622

07 **Unserved routing group visibility** Not checked yet NEW

70

Agent model pickers now list every routing group the agent lacks alongside a human-readable reason via `UnservedGroup.Why`, using constants such as `WhyGroupOff`, `WhyGroupUnticked`, `WhyGroupTurnedOff`, `WhyGroupNoServer`, `WhyGroupAgentCant`, `WhyGroupNarrowed`, `WhyGroupNotPicked`, and `WhyGroupTakenOut`.

— Names all reason constants but no direct user command `v0.1.4618`

---

08 **Grouped relay models with effort variants** Not checked yet

IMPROVED

65

Relay-added models whose IDs share a base and differ only by a trailing effort word (`low`, `medium`, `high`, `minimal`, or `xhigh`) are now grouped as one model with effort variants, so the model picker and editor chips show one entry instead of one per effort level.

— Names exact effort suffixes and UI effect `v0.1.4623`

---

09 **Traditional Chinese site localization** Not checked yet NEW

65

magpie's site now serves a Traditional Chinese home page at `/zh-tw/`, using Taiwan vocabulary (帳號, 閘道, 選單列, 終端機, 設定檔) with docs links routing to the Simplified Chinese docs. Browser language detection now routes `zh-TW`, `zh-HK`, `zh-MO`, `zh-Hant`, and `zh-Hant-TW` tags to `/zh-tw/` instead of `/zh/`.

— Gives exact path and detection tags, a clear destination `v0.1.4640`

THINNER COVERAGE BELOW

---

10 **Provider-scoped plugin panels** Not checked yet IMPROVED

50

Plugin panels now carry a `Providers` field (a list of provider IDs) in the `pluginPanel` struct, so a plugin's UI rows appear only in the editors of the providers it supplies rather than in every provider's editor.

— Names the struct field but describes only internal change `v0.1.4644`

WAS THIS USEFUL? 👍 👎

Alibaba OpenCodeReview

Feedback ⓘ v1.14.0 NOTES ↗

RANK

OpenCodeReview is a secure, fast, efficient, battle-tested at Alibaba's scale. Hybrid architecture code review tool: deterministic pipelines + LLM Agent, precise line-level comments, built-in multi-language ruleset (NPE, thread-safety, XSS, SQL injection), OpenAI & Anthropic compatible.

> OpenCodeReview's v1.14.0 release expands its ruleset and file support with new SQL, Dhall, and Ruby review rules, CUDA/HIP source file review, severity-ordered CLI output, and an optional behavior and impact analysis step.

└──▷ **WHAT SHIPPED · 4 FEATURES** most completely described first ? what's the number?

01 **CUDA and HIP source file review support** Not checked yet NEW

55

Added review support for CUDA (`.cu` / `.cuh`) and HIP (`.hip`) source files.

— Names exact file extensions supported, no further mechanism given. v1.14.0

02 **Severity-ordered CLI findings** Not checked yet IMPROVED

40

Review and scan findings can now be ordered by severity in the CLI output.

— Clear CLI behavior change but no flag name given. v1.14.0

03 **SQL, Dhall, and Ruby review rules** Not checked yet NEW

35

Added built-in review rules for SQL, Dhall, and Ruby languages, extending the multi-language ruleset.

— Only mentioned in summary, no rule details or flags given. v1.14.0

04 **Optional behavior and impact analysis step** Not checked yet NEW

30

Reviews can now include an optional step that performs behavior and impact analysis.

— No config key or flag named to enable this step. v1.14.0

WAS THIS USEFUL? 👍 👎

Yuval Sarel cones

Feedback ⓘ v0.4.25 NOTES ↗ RANK

Cones is a terminal dashboard for managing coding-agent sessions across projects.

> cones added Linux platform support with systemd timers and native clipboard tooling, reworked worktree organization with a new breaking config default, and layered fold/unfold navigation, state counts and inline prompts onto the session-list dashboard.

└──▷ **WHAT SHIPPED · 8 FEATURES** most completely described first ? what's the number?

01 Worktree folder display and config

Not checked yet BREAKING

90

Worktree sessions now sit inside their repository's folder by default and show the worktree name dim after the title (for example, `fix pager · in ⑂ agent-a1f`); set `worktrees: folder` in config to list every worktree as its own top-level folder instead of nesting it under its repository. The old `nested` and `flat` values for the `worktrees` config key are replaced by the new default value `repo`, and existing configs using `nested` or `flat` will read as `repo` on upgrade. Worktree headings also align with titles, show the last path part of the worktree, and display `├` and `└` tree lines under their folder.

List every worktree as its own top-level folder in the session list rather than grouping them under their repository.

yaml

```
worktrees: folder
```

— Names config key, file, old/new values and exact display format v0.4.22 v0.4.23

02 Fold/unfold navigation for session list headings

Not checked yet IMPROVED

88

Session list headings now go two levels deep — a folder or tag at top, with tags inside a folder — connected by tree lines (`├─` and `└─`) after the state and harness marks. Fold a tag or folder with `←` when nothing is typed; unfold with `→` or Enter — the folded heading keeps its session counts and stays folded across restarts. `←` on a heading folds it, and `←` on a folded tag also folds its parent folder.

— Exact keys and persistence behaviour given v0.4.25 v0.4.23

03 Linux platform support

Not checked yet NEW

80

cones now runs on Linux: process discovery reads `/proc`, background jobs install as systemd user timers, and notifications, clipboard and image paste use `notify-send`, `wl-copy` / `xclip` or OSC 52, and `wl-paste` / `xclip`.

— Names implementation tools but no setup command v0.4.23

04 **Quiet exit for piped commands** Not checked yet IMPROVED

65

Commands `ls`, `show`, `search`, `skill`, and the job list now exit quietly when piped into `head` instead of panicking.

— Names exact commands and fixed failure mode v0.4.22

THINNER COVERAGE BELOW

05 **Inline waiting-session prompts** Not checked yet NEW 45

Waiting rows now show a Claude background job's question inline on the row, with the input notification quoting it on one line; the composer placeholder on a waiting row now points at Enter.

— Describes behaviour without exact UI path v0.4.22

06 **Session heading state counts** Not checked yet NEW 35

Dashboard session-list headings now count working, shell, input, done, and failed sessions after the heading name.

— Names the counted states but little else v0.4.22

07 **Mouse drag indicator for session rows** Not checked yet NEW

30

Mouse-press on a session now highlights it until release and underlines the row it would drop on.

— One-line visual change, no further mechanism v0.4.25

08 **Quick-reference help overlay** Not checked yet NEW 30

cones adds a quick-reference help overlay accessible via `?`.

— Only mentioned in summary, no further detail given v0.4.25

└──▷ **BREAKING ON UPGRADE**

! The `worktrees` config values `nested` and `flat` are replaced by `repo` (the new default); existing configs using `nested` or `flat` will read as `repo` on upgrade.

WAS THIS USEFUL? 👍 👎

Cline

Feedback ⓘ v4.1.27 NOTES ↗ CODE ↗ RANK

Cline autonomous coding agent as an SDK, IDE extension, or CLI assistant.

> Cline shipped full-content file reads without the old 8k cap, TUI-driven Cline Cloud sessions with new one-shot run-mode CLI flags, and an SDK update adding a UX-verification subagent, cloud artifact and session-resume tooling, hub-based provider/MCP management, and a much larger model catalog.

└──▷ **WHAT SHIPPED · 14 FEATURES** most completely described first ? what's the number?

01 Cloud session management, TUI access, and artifact tooling

Not checked yet **NEW** 88

Set `CLINE_CLOUD_SESSIONS=1` to run tasks in Cline Cloud from the TUI, toggle between Local and Cloud execution with Ctrl+O, and pick the repo and branch with Ctrl+K; the SDK adds artifact tooling for cloud sessions via createPublishArtifactsTool({ publish }), which publishes artifacts via signed upload policies using `createArtifactPublisher` from `@cline/core/cloud` (artifact IDs default to the source filename), plus HubCloudSessionClient.getArtifactDownloadUrl(sessionId, artifactId) to retrieve a signed artifact download URL without starting the sandbox, and exports confirmed cloud handoff publication directly from `@cline/core/cloud` for direct hosts; `session.resume` now resumes a saved session by ID using hub history instead of requiring `initialMessages`, a cloud task can be created and sent in one call via `initialMessage` or launched from a Git snapshot via `localGit`, and session history can be retrieved from the hub in bounded chunks, including archived cloud history.

Run a headless task in Cline Cloud from the TUI, choosing the repo, branch, and cloud model interactively.

```
$ CLINE_CLOUD_SESSIONS=1 cline
```

Download an artifact from a cloud session by getting a signed URL without spinning up the sandbox.

```javascript
const url = await
HubCloudSessionClient.getArtifactDownloadUrl(sessionId,
artifactId);
```

Resume a previously saved agent session and continue work without supplying initial messages.

```javascript
const runtime = await session.resume(savedSessionId);
```

Create a cloud task and send its first message in a single call, optionally launching from a Git snapshot.

```javascript
await cloudClient.createTask({ initialMessage: 'Run the full test suite and open a PR for any failures', localGit: snapshotId });
```

— Exact env var and keybindings with a runnable example Named functions with a runnable code example for downloads Named methods and params with runnable code examples

cli-v3.0.75 sdk/sdk/v0.2.1

02 **One-shot run mode controls** Not checked yet BREAKING 88

One-shot (non-interactive) CLI runs now default to yolo mode instead of act mode; pass `--act` to restore act-mode behavior and get team tools. The first turn of a one-shot run can now be interrupted with `--timeout` or Ctrl+C, which previously had no effect. Pass `--auto-approve false` to keep a one-shot run in act mode with auto-approve off.

Run a one-shot CI prompt that needs team tools — passing `--act` overrides the new yolo-mode default.

```
$ cline --act "Run the full test suite and open a PR for any failures"
```

Run a timed one-shot prompt in CI and ensure it can be interrupted if it stalls on the first turn.

```
$ cline --timeout 120 "Lint the codebase and report errors"
```

— Named flags with runnable examples and a clear before/after cli-v3.0.75

03 **UX verification subagent and report tool** Not checked yet NEW 87

Adds a `ux_verifier` UX verification subagent, registered when tools and spawning are enabled and disableable via `spawn_agent`, alongside an opt-in `report_ux_verification` tool built with createUxVerificationReportTool() that produces pass/fail/blocked results with captions and screenshot references.

— Names subagent and tool functions but lacks a usage example sdk/sdk/v0.2.1

04 **MCP tool read-only hints and badges** Not checked yet IMPROVED

76

The MCP tab now shows a `read-only` badge on tools whose server declares them read-only, and the SDK carries these tool hints (`readOnlyHint` and others) through stdio and URL clients and the manager cache via `McpToolDescriptor.annotations`.

— Names UI badge and the SDK field carrying the hint sdk/sdk/v0.2.1 v4.1.27

05 **Full file content from read_files tool** Not checked yet IMPROVED

75

The `read_files` tool now delivers full file content to the model — up to 2000 lines or about 48k characters — instead of being capped at 8k characters.

— Clear numeric before/after; not a user-invoked command v4.1.27

06 **Model catalog, defaults, and picker updates** Not checked yet

NEW INTEGRATION RETIRED INTEGRATION UPDATED INTEGRATION IMPROVED

75

The model catalog grew from 6,457 to 7,621 models, adding 94 providers from models.dev; the extension updated defaults for 67 providers (mostly to Kimi K3.1 Pro, Gemini 4 Neon including the `gemini` and `vertex` providers, or DeepSeek V4.1 Flash 1017), while the SDK's bundled `cline` provider default changed to Nemotron 3.5 Super VL and `CLINE_DEFAULT_MODEL_ID` became `anthropic/claude-sonnet-5.5` (with `@cline/shared` adding `ANTHROPIC_DEFAULT_MODEL_ID` for Anthropic's bare ID format); new users with no saved model now start on the first free model instead of Claude Sonnet 5.5; and a provider's model list now refreshes automatically after you save its API key or base URL, showing an error in the model picker when the refresh fails.

— Describes behavior and failure case but no named surface Rich numbers and named env vars across extension and SDK Simple behavior change, no further mechanism given

v4.1.27 sdk/sdk/v0.2.1 cli-v3.0.75

07 **Reasoning effort controls** Not checked yet IMPROVED

62

The reasoning effort dropdown in the extension gained a Default entry that clears the saved effort setting and shows the effort level that will actually be used for a new task; the SDK's session config and hub RPC schema now also accept a `max` reasoning effort level.

— Names dropdown entry and new schema value, no usage example sdk/sdk/v0.2.1 v4.1.27

THINNER COVERAGE BELOW

08 **Interleaved reasoning placeholder field** Not checked yet NEW

59

Models that require interleaved reasoning now receive a placeholder `reasoning_content` on assistant messages that have none, carried as `ModelInfo.metadata.interleavedReasoningField`.

— Names the field but gives no usage example `sdk/sdk/v0.2.1`

09 **Hub commands for provider, model and MCP management**

Not checked yet NEW

51

Manage providers, models, MCP servers, and MCP OAuth as hub commands, with token authentication required for provider management.

— Names four managed surfaces but no exact command syntax `sdk/sdk/v0.2.1`

10 **CLI stops reading local .env and bunfig.toml** Not checked yet

50

BREAKING

The CLI no longer reads `.env` or `bunfig.toml` from the directory it is launched in.

— Clear breaking change but no migration guidance given `cli-v3.0.75`

11 **Removal of CLINE_CAPTURE_* request logging** Not checked yet

DEPRECATED

50

The `CLINE_CAPTURE_*` provider request capture, which wrote full request bodies to disk, has been removed.

— Names the removed env var family and its prior behavior `sdk/sdk/v0.2.1`

12 **LM Studio API key field** Not checked yet NEW

42

LM Studio settings now include an API key field to authenticate against servers with Require Authentication turned on.

— Thin UI description, no further mechanism given `v4.1.27`

13 **Privacy settings link in Account view** Not checked yet NEW

37

The Account view now includes a Privacy settings link to the Cline web app's data-use preferences.

— Simple UI addition, no further detail given `v4.1.27`

14 **Third-party notices per platform package** Not checked yet NEW

35

Each platform package now ships a `THIRD_PARTY_NOTICES.txt` listing embedded third-party code.

– Names the file but gives no further detail cli-v3.0.75

└──▷ **BREAKING ON UPGRADE**

! One-shot runs (prompt argument or piped stdin) now start in yolo mode by default instead of act mode; pass `--act` to restore the previous behaviour when team tools are needed.

! The CLI no longer reads `.env` or `bunfig.toml` from the directory it is launched in.

! The `CLINE_CAPTURE_*` provider request capture, which wrote full request bodies to disk, has been removed.

desktop-v0.0.50 NOTES ↗ CODE ↗ STABLE

Cline adds Max reasoning effort in the thinking picker, lifts the 8k `read_files` cap to 2000 lines/48k chars, adds a Stop button for `/compact`, and adds Privacy and Third-party notices links.

└──▷ **GET THIS VERSION** ▾

```
$ git clone --branch desktop-v0.0.50
https://github.com/cline/cline.git
# already have the repo? check out this version:
$ git checkout desktop-v0.0.50
```

Read files in full up to 2000 lines or about 48k characters — `read_files` results are no longer cut to 8k characters before reaching the model. Not checked yet

Select Max reasoning effort in the thinking picker for models whose top level is max. Not checked yet

Stop a manual `/compact` mid-run with a Stop button instead of watching it sit on Thinking... indefinitely. Not checked yet

Access a Privacy settings link to Cline's data-use preferences from the Account view. Not checked yet

View third-party open-source licenses under Settings → About → Third-party notices. Not checked yet

WAS THIS USEFUL? 👍 👎

repoprompt-ce

Feedback ⓘ v1.10.0 NOTES ↗ RANK

Repoprompt-ce is a native macOS context engineering app for AI coding agents, with an MCP CLI (community edition of RepoPrompt).

> RepoPrompt CE's v1.10.0 release adds Google Antigravity model support across its core surfaces, a provider handover preview for mid-session switching, and an experimental Command Center for transcript history, alongside Oracle selection controls, extended idle auto-wake, and bundled model updates.

└──▷ **WHAT SHIPPED · 6 FEATURES** most completely described first ? what's the number?

01 Google Antigravity model support across core surfaces

Not checked yet NEW 67

RepoPrompt CE can use Google Antigravity models via the installed `agy` CLI in Agent Mode, delegated role defaults, Built-in Chat, and Oracle, and Antigravity now appears as a provider option in onboarding and the Setup Wizard. Users can opt in to Antigravity usage tracking to see weekly and five-hour quotas by model group, including remaining usage and reset times.

Connect your installed Antigravity CLI so Agent Mode can use Google Antigravity models for agentic tasks.

📍 1. Sign in to your Google account with the `agy` CLI on your machine. 2. In RepoPrompt CE, go to Settings › Setup Wizard and select Antigravity as a provider. 3. Open Agent Mode and choose an Antigravity model from the model selector.

— Names CLI, four integration points, and quota mechanism but no limits v1.10.0

02 Provider handover preview when switching mid-session

Not checked yet NEW 60

Switching AI providers now shows a handover preview of what state will carry over, plus recovery controls for cases where the current session is busy or waiting for input.

Review and switch your active AI provider mid-session without losing context — useful when a session is busy or awaiting input.

📍 1. While in an active session, open the provider switcher. 2. Review the handover preview showing what state will carry over. 3. Use the recovery controls if the session is busy, then confirm the switch.

— Describes mechanism and edge-case handling with a usage walkthrough v1.10.0

THINNER COVERAGE BELOW

03 **Bundled model updates: Codex 0.170.0 and Claude Fable 5.2**

Not checked yet IMPROVED 48

Bundled Codex is updated to 0.170.0, adding GPT-6.2 model options and a configurable Codex context window; Claude Fable 5.2 is now also available.

— Names exact versions and new model options, no config instructions v1.10.0

04 **Oracle selection for new conversations** Not checked yet NEW

40

Users can pick which Oracles answer a new conversation; Oracle requests run in the background and the displayed Oracle count tracks the latest request.

— Explains behaviour but no UI path or config surface given v1.10.0

05 **Experimental Command Center for transcript history**

Not checked yet NEW 40

A new experimental Command Center lets users browse saved history, export transcripts, and reopen linked workspaces or sessions.

— Names capabilities but no access path or export format v1.10.0

06 **Extended idle auto-wake intervals** Not checked yet IMPROVED

40

Idle auto-wake now offers intervals through 24 hours and better respects attention, pending prompts, and active compaction.

— States new limit but no setting location given v1.10.0

WAS THIS USEFUL?  

Alibaba Qwen Code

Feedback ⓘ v0.28.3 NOTES ↗ RANK

Qwen Code is an open-source AI coding agent that lives in your terminal.

> Qwen Code's v0.28.3 release adds a wave of managed-channel integrations (GitHub, Teams, WeCom, QQ, generic webhooks), plus ACP Mods support, idle MCP connection persistence, a configurable vision bridge timeout, an SSE connection admission budget, Web Shell cross-session messaging, a Traditional Chinese locale, and an opt-in CLI screen-reader mode.

└──▷ **WHAT SHIPPED · 10 FEATURES** most completely described first ? what's the number?

01 **Managed-channel adapters for GitHub, Teams, WeCom and QQ**

Not checked yet NEW 80

Adds a GitHub managed channel; a Generic Webhook channel that posts notification, prompt, and scheduled-task results as JSON to a configured HTTPS endpoint with optional Bearer authentication; an experimental WeCom managed-channel adapter that routes private-chat messages into bound Managed Agent Sessions and returns replies over an authenticated WebSocket; a Microsoft Teams channel extension with personal chats, group mentions, thread isolation, and `createTeamsIngress` host webhook setup; and an off-by-default QQ bot option giving agents tools to send local files, images, videos, and voice messages.

— Names every adapter and its mechanism, but no CLI/config flags given v0.28.3

02 **Configurable vision bridge timeout** Not checked yet NEW

80

Set a vision bridge timeout with `--timeout` on `/model --vision`, accepting seconds or ms-style values and applying immediately.

Override the vision bridge timeout mid-session when working with slow vision endpoints.

```
$ /model --vision --timeout 30s
```

— Runnable command with flag and value format shown v0.28.3

03 **SSE connection admission budget for daemon streams**

Not checked yet NEW 80

Apply the `--max-sse-connections` admission budget across all daemon stream surfaces, with shared occupancy accounting and logged refusal when the limit is exceeded.

— Flag and mechanism named but no example command v0.28.3

04 **Idle MCP connections persisted across sessions** Not checked yet

NEW 75

Keep idle MCP connections alive across session suspensions with the `qwen.managed-agent.runtime-broker.mcp-parking.enabled` opt-in setting so reopened sessions reuse them, skipping reconnect overhead.

Persist idle MCP connections across managed-agent session suspensions so the next wake skips reconnect overhead.

```yaml
qwen.managed-agent.runtime-broker.mcp-parking.enabled: true
```

— Config key, effect and example provided v0.28.3

05 **Inject channel dispatch mode** Not checked yet NEW 70

Adds an `inject` channel `dispatchMode` that delivers incoming messages to the running turn so the agent can read them at the next tool-batch boundary.

— Named config value and mechanism, no usage example v0.28.3

06 **Mods support in ACP sessions** Not checked yet NEW 65

Enable Mods in ACP sessions for `qwen --acp`, Zed, and the VS Code companion by turning on `experimental.mods`.

— Exact config flag given, but behaviour of Mods unexplained v0.28.3

07 **Cross-session messaging and Goal markers in Web Shell**

Not checked yet NEW 60

Adds `/send` for cross-session messaging in Web Shell and a status-tinted Goal marker on Web Shell sidebar session rows to flag sessions with active Goals.

— Command and UI location named, little mechanism detail v0.28.3

THINNER COVERAGE BELOW

08 **Traditional Chinese UI locale** Not checked yet NEW 50

Adds a Traditional Chinese UI locale (`zh-TW`, `zh-HK`, `zh-MO`, `zh-Hant`) to Web Shell and VS Code companion chrome.

— Locale codes named but no activation steps given v0.28.3

09 **Screen-reader mode for CLI** Not checked yet NEW 50

Adds opt-in native screen-reader presentation for the CLI with keyboard navigation, stable completed-reply announcements, and semantic history while Ink remains the

default renderer.

— Describes behaviour but no toggle name or flag given v0.28.3

10 **Workspace artifact library in Managed Agent Web UI**

Not checked yet NEW 45

Adds a workspace artifact library to the Managed Agent Web UI for browsing, previewing, and downloading artifacts across sessions.

— UI feature described at surface level only v0.28.3

WAS THIS USEFUL? 👍 👎

codealta

Feedback ⓘ 4.5.0 NOTES ↗ RANK

Codealta is an agentic Development Environment

> CodeAlta shipped a one-command cross-platform installer with version pinning and headless fallback, per-project secrets and environment variables, a local Pseudonymizer plugin for PII redaction, multi-machine phone access, and a new disk-usage view in Settings > Storage, alongside automation tiers, shared worktree sessions and script notifications — while removing the `wait` tool from Code Mode scripts.

└──▷ **WHAT SHIPPED · 11 FEATURES** most completely described first ? what's the number?

01 One-command installer with version pinning and TUI fallback

Not checked yet NEW 95

A single shell command installs or updates CodeAlta: `curl -fsSL https://codealta.io/install.sh | sh` on macOS/Linux or `irm https://codealta.io/install.ps1 | iex` on Windows (PowerShell). Flags choose which apps to install — `--desktop --tui` for a terminal-first setup without the CLI, `--all` for Desktop+TUI+CLI, and `--all --no-modify-path` to install everything without touching the shell profile — while `--version <v>` / `-Version <v>` pins to a specific release and `--prerelease` / `-Prerelease` grabs the latest preview build. On Linux without a desktop session (e.g. over SSH), the installer automatically falls back to CodeAlta TUI instead of CodeAlta Desktop.

Install only the TUI and Desktop apps on macOS/Linux — useful when you want a terminal-first workflow alongside the desktop app but no headless CLI.

```
$ curl -fsSL https://codealta.io/install.sh | sh -s -- --desktop --tui
```

Install all three apps (Desktop, TUI, CLI) on macOS/Linux while keeping your shell profile unchanged.

```
$ curl -fsSL https://codealta.io/install.sh | sh -s -- --all --no-modify-path
```

Install CodeAlta Desktop, TUI and CLI in one shot on macOS or Linux, including .NET.

```
$ curl -fsSL https://codealta.io/install.sh | sh -s -- --all
```

— Every install flag and fallback condition is named and runnable. 4.5.0

02 **Per-project secrets and environment variables** Not checked yet

NEW

80

Set per-project environment variables in `.alta/config.toml` under the `[env]` table, referencing secrets by name without committing their values, e.g. `GH_TOKEN = "${secret:ACME_GH_TOKEN}"`; sub-projects in a monorepo inherit them. A new Settings > Secrets and environment page organizes this into Overview, Secrets, Variables, Environment, Vaults, Sign-ins and Activity tabs.

Scope a repository token to only the project that needs it, so it never reaches sessions of other repos.

```toml
# .alta/config.toml
[env]
GH_TOKEN = "${secret:ACME_GH_TOKEN}"
```

— Names the exact config key and gives a runnable example. 4.5.0

03 **Pseudonymizer plugin for PII redaction** Not checked yet NEW

70

A built-in Pseudonymizer plugin replaces personal data (names, locations, IBANs) with tokens such as `[PERSON_001]`, `[LOCATION_002]` and `[IBAN_001]` before sending to a model provider, then restores real values in the response; detection runs locally in French, English or several other languages.

— Mechanism well explained but no activation steps given. 4.5.0

04 **Shared worktree sessions** Not checked yet NEW

65

Start a new session in an existing worktree from the CLI with `alta session create --worktree-path` so multiple sessions can share one worktree.

Spin up a second agent session alongside an existing one in the same worktree, sharing the same checkout.

```
$ alta session create --worktree-path ./worktrees/feature-x
```

— Runnable command, limited explanation of session interplay. 4.5.0

05 **Disk usage view in Settings > Storage** Not checked yet NEW

60

A new `Settings > Storage` view shows disk usage per project, worktree and session, with git state details, and lets you selectively clean build output (`bin` , `obj` , `node_modules`) from chosen checkouts.

— Named UI path and target folders, no command given. 4.5.0

THINNER COVERAGE BELOW

06 **Automation tiers and Adaptive model selection** Not checked yet

NEW

55

Create automation tiers from the CLI with `alta automation create --tier` to assign model tiers, and use Adaptive in Runs for issues, scripts and automations.

— Command named but tier/Adaptive mechanics unexplained. 4.5.0

07 **Multi-machine phone access** Not checked yet NEW

55

The phone page now lists sessions from every machine in the fleet; enabling 'Show my other machines on the phone' lets you view, prompt and start sessions on remote machines, including headless CodeAlta Node machines.

.

— Named toggle gives a starting point, mechanism is prose. 4.5.0

08 **Notifications from scripts and automations** Not checked yet NEW

55

A new notify() function in Code Mode scripts raises a CodeAlta notification; automations also gain a `Notify me when a run ends` option that fires on completion or only on failure.

— Function and option named, usage context otherwise thin. 4.5.0

09 **Tab switching shortcuts** Not checked yet NEW

45

Use `Ctrl+Tab` to switch back to the last-used tab; hold `Ctrl` to see all tabs in the pane ordered by recent use.

— Simple shortcut description, minimal mechanism. 4.5.0

10 **Richer plugin metadata cards** Not checked yet IMPROVED

40

Plugin cards in `Settings > Plugins` now display a picture, icon, color, author, version and tags; custom plugins support `[Plugin]` attributes for icon and color, alongside a `plugin.webp` image.

— Lists new fields shown, no deeper mechanism or steps. 4.5.0

11 **`wait` tool removed from Code Mode** Not checked yet BREAKING

35

The `wait` tool is removed from Code Mode scripts; scripts that relied on it must be updated to wait for what they start themselves.

— States the removal but no migration example. 4.5.0

└──▷ **BREAKING ON UPGRADE**

! The `wait` tool is removed from Code Mode scripts; scripts that relied on it must be updated to wait for what they start themselves.

WAS THIS USEFUL? 👍 👎

pstack-claude

Feedback ⓘ v0.9.146 NOTES ↗ RANK

pstack-claude is a port of Poteto's pstack agent framework to Claude Code, Codex, Copilot, Pi, and other AI coding assistants.

> pstack-claude's v0.9.144–146 window centers on tightening watch-pr's CI verdict accuracy (new `--expect-app` flag, check-suites source tagging, approval prompts, stricter GitHub CLI version requirement) alongside new DeepSeek Harness support, session-routing control, OMP skill editing, and Windows/TLC fixes for the statechart command.

└──▷ **WHAT SHIPPED · 7 FEATURES** most completely described first ? what's the number?

01 **`--expect-app` flag for watch-pr** Not checked yet NEW

93

`watch-pr` now accepts a `--expect-app <slug>` flag naming a GitHub App whose queued-but-not-yet-running check suite counts as a pending check rather than clean CI, preventing false READY verdicts on repos like `microsoft/react-native-windows` where Azure Pipelines can sit queued for hours.

— Flag, mechanism and concrete repo example all given v0.9.146

02 **`PSTACK_SESSION_HOOK=off` session routing control** Not checked yet NEW

85

Setting `PSTACK_SESSION_HOOK=off` turns off session routing without a user sheet, checked in `session-start.sh` and `session-start.ps1` before the sheet — covering Claude Code, Codex, GitHub Copilot, Muse, and Antigravity runtimes that pass the variable to their hooks.

— Exact env var and script paths given, mechanism explained v0.9.146

03 **GitHub CLI 2.48+ required for watch-pr suite lookup**

Not checked yet BREAKING

83

The suite lookup now requires GitHub CLI 2.48 or later; a CLI that rejects `--slurp` produces an immediate exit 7 with an upgrade instruction, saving 12 minutes of retries at the default settings.

— Exact version, flag and exit code given, clear migration note v0.9.144

04 **DeepSeek Harness (DSH) skills-only install support**

Not checked yet NEW

74

pstack can run as a skills-only install under DeepSeek Harness (DSH) via the shared `~/.agents/skills/` install path, with a generated `dsh-tools.md` mapping Claude

Code tool names, subagent policy, and model roles to DSH.

– Install path and mapping file named, scope clear v0.9.145

05 **watch-pr verdict accuracy and CI reading improvements**

Not checked yet IMPROVED 74

Each failed or pending check row from the check-suites API now carries `source: "check-suites"` in the verdict JSON, while the top-level CI source still names the primary read. A zero-run GitHub Actions suite waiting for approval now tells the reader to approve its workflows, while other providers requiring action point the reader to their check details. `watch-pr` now verifies the head and run count before deciding a suite is covered, skips a suite with a positive count even when GitHub returns a null status, and reads suites as soon as reported checks finish alongside outstanding review-thread pages.

– JSON field and behaviours named but only internal logic v0.9.144

06 **Statechart command Windows and TLC output handling**

Not checked yet IMPROVED 65

The `/statechart` command now reads TLC output with `\r\n` line endings as well as `\n` in the run reader, DOT reader, and message reader. On native Windows, `statechart/SKILL.md` now documents the `node` command that runs `statechart.ts`, and `java` and `mise` start without a console window.

– Specific readers and files named but no single action v0.9.145

07 **Reflect writes edited skills to OMP folders** Not checked yet

IMPROVED 62

On OMP, `reflect` now writes edited pstack skills to `.omp/skills/<skill>/` for project copies or the active profile's agent skills folder for personal ones, and warns before approval that the edited copy is a separate skill beside `pstack:<skill>`.

– Named path given but no runnable command v0.9.146

└──▷ **BREAKING ON UPGRADE**

! The suite lookup now requires GitHub CLI 2.48 or later; a CLI that rejects `--slurp` exits immediately with code 7.

WAS THIS USEFUL?  

Anthropic Claude Code

Feedback ⓘ v2.1.330 NOTES ↗

CODE ↗ RANK

Claude Code is Anthropic's terminal coding agent that plans, edits, and tests code in local repositories.

> Claude Code v2.1.330 adds Enterprise Admin Console sign-in support, self-hosted runner start-directory control and new orchestration metrics, JSON diagnostics for plugin installs, and automatic chart rendering via the built-in Easel mod.

└──▷ **WHAT SHIPPED · 5 FEATURES** most completely described first ? what's the number?

01 Enterprise Admin Console sign-in and link controls

Not checked yet NEW 75

After `claude auth login`, Claude Code can sign in to organizations managed in Anthropic's Enterprise Admin Console and will apply the organization's settings and use its cloud provider or gateway. Setting the environment variable `CLAUDE_CODE_FULL_SIGN_IN_LINK=1` prints the full sign-in URL instead of the short link, for when the short link errors or a script needs to read it.

Print the full sign-in URL in a CI script where the shortened link fails or needs to be parsed programmatically.

```
$ CLAUDE_CODE_FULL_SIGN_IN_LINK=1 claude auth login
```

— Names exact command, env var, and org-level effects v2.1.330

02 Self-hosted runner start-directory control

Not checked yet NEW 65

Setting `--start-dir` or the `SELF_HOSTED_RUNNER_START_DIR` environment variable lets Claude begin work in a specific folder of a repository on a self-hosted runner, while settings and hooks still load normally.

— Names flag and env var but no runnable example given v2.1.330

03 New self-hosted runner timing metrics

Not checked yet NEW 65

Two new metrics are exposed on the self-hosted runner's `/metrics` endpoint: `claude_code_self_hosted_runner_order_to_registered_seconds`, measuring seconds from work order issued to registration accepted on orchestrator-spawned runners, and

`claude_code_self_hosted_runner_order_to_first_session_seconds`, measuring seconds from work order to first session handed over on runners started for a session.

— Names endpoint and both metrics precisely, no usage sample v2.1.330

04 **JSON diagnostics for plugin install command** Not checked yet

IMPROVED 60

`claude plugin install --json` output now includes `commandPrinted` and `commandFailed` fields, letting scripts detect whether a plugin's install command wrote to stderr or exited with an error.

Check whether a plugin's install command had any output or failed, useful when scripting plugin deployment across many runners.

```
$ claude plugin install my-plugin --json | jq '{commandPrinted, commandFailed}'
```

— Includes a runnable command demonstrating the new fields v2.1.330

THINNER COVERAGE BELOW

05 **Automatic chart rendering via Easel mod** Not checked yet IMPROVED 55

Easel, a built-in mod, now draws charts on request and automatically under MCP metrics results on Pro, Max, and where available Team and Enterprise plans, on macOS and Linux; automatic charts can be turned off in `/config`.

— Names plans, platforms, and config toggle location, no command v2.1.330

WAS THIS USEFUL? 👍 👎

The Open Engine Zeroshot

Feedback ⓘ v16.2.0 NOTES ↗

RANK

Zeroshot orchestrates executor-verifier agent loops that implement and validate software changes while logging workflow state to SQLite.

> Zeroshot's v16.2.0 release adds node-scoped human steering, lifecycle scripts for local environments, and documented gateway protocol options across its agent lanes.

└──▷ **WHAT SHIPPED · 3 FEATURES** most completely described first ? what's the number?

01 **Node-targeted steering command** Not checked yet NEW 82

Human guidance can now be directed at a single graph node using `zeroshot steer --node NAME`, routing prompt and tool-hook delivery only to that admitted agent across mapped instances and future visits, rather than broadcasting to all active participants.

Steer a specific graph node when you want human guidance to reach only one agent rather than all active participants in a run.

```
$ zeroshot steer --node NAME
```

— Exact command and mechanism given, fully actionable v16.2.0

02 **Local environment lifecycle scripts** Not checked yet NEW 72

Local Linux and macOS runs can now execute `setup` and `startup` scripts defined via `--environment FILE`, keeping background services alive to graph nodes for the life of the attempt, with an optional `teardown` script run on orderly cleanup.

— Flag and script hooks named but no runnable example shown v16.2.0

THINNER COVERAGE BELOW

03 **Gateway protocol options in setup catalog** Not checked yet NEW 55

Wire protocol descriptions and allowed values for the Pi, Copilot, and registered OpenCode and Hermes lanes are now exposed under `setupLanes[].fieldOptions.GATEWAY_API` in the setup catalog.

— Config path named but described only as documentation, thin on behavior v16.2.0

WAS THIS USEFUL? 👍 👎

All Hands AI OpenHands

Feedback ⓘ v1.43.0 NOTES ↗

RANK

OpenHands is an open-source software-development agent that writes code, runs commands, and browses the web.

> OpenHands' v1.43.0 release focuses on UI and workflow polish: automations can now be created from the setup wizard, a new Settings → Experimental page exposes feature flags like conversation reactions, chat gains decision-card and message-edit rendering, and the conversation list is now virtualised for smoother scrolling.

└──▷ **WHAT SHIPPED · 4 FEATURES** most completely described first ? what's the number?

01 Experimental features settings page

Not checked yet NEW

60

A new Settings → Experimental page lets users enable and manage experimental feature flags, including conversation reactions, without touching server config.

Turn on conversation reactions and other experimental capabilities for your user account without touching server config.

📍 In the OpenHands UI, go to Settings → Experimental and toggle the feature flags you want to enable (e.g. Conversation reactions).

– Exact UI path and toggle steps provided v1.43.0

THINNER COVERAGE BELOW

02 Automations creation via setup wizard

Not checked yet NEW

45

Automations can now be created directly from the setup wizard, and each automation run row displays a conversation indicator icon.

– Names UI surface but no exact navigation steps given v1.43.0

03 Decision cards and inline message-edit rendering in chat

Not checked yet NEW

45

When experimental features are enabled, chat now renders experimental decision cards and shows the agent's message edits inline.

– Describes behavior but depends on an unlisted specific flag v1.43.0

04 Virtualised conversation message list

Not checked yet IMPROVED

The conversation message list is now virtualised, enabling smooth scrolling in long conversations that previously strained the browser.

— Explains change but no user action or config involved `v1.43.0`

WAS THIS USEFUL?  

flow-next

Feedback ⓘ flow-next-v9.18.0 NOTES ↗ RANK

flow-next is a workflow plugin that takes a bug, idea or ticket to a verified pull request with specs, cross-model review, live QA, and receipts

> flow-next's biggest addition this window is full Google Antigravity CLI support — a dedicated installer, bridge recipe, and read-only review wrapper — alongside new cross-model reviewer connections and several reviewer-backend behavior changes, some of them breaking.

└──▷ **WHAT SHIPPED · 7 FEATURES** most completely described first ? what's the number?

01 Antigravity CLI agent loop and installer

Not checked yet NEW

93

Running `./scripts/install-antigravity.sh` installs flow-next into Google's Antigravity CLI, mapping all 21 agents to Antigravity's model tiers, rewriting tool restrictions as allowlists, and validating every skill and agent with `agy plugin validate`; start a session with `agy --settings '{"toolPermission":"always-proceed"}'` (or `agy --settings always-proceed`) and run `/flow-next:setup` to take the full plan-build-review-land loop through to a landed pull request with every `/flow-next:` command available. A plain `agy plugin install` of the repository instead of the dedicated script loads skills but skips 19 of the 21 agents and misidentifies the host.

Bootstrap a full Antigravity CLI environment with all 21 flow-next agents pre-loaded and every `/flow-next:` command ready to use.

```
$ bash install-antigravity.sh
```

Run the entire plan-build-review-land loop without stopping at approval gates, useful in automated CI or hands-off environments.

```
$ agy --settings always-proceed
```

Set up flow-next in a project running inside Google's Antigravity CLI after installing with the dedicated script.

```
$ ./scripts/install-antigravity.sh
agy --settings '{"toolPermission":"always-proceed"}'
# then, in your project conversation:
/flow-next:setup
```

— Names exact install script, flags, and command sequence `flow-next-v9.18.0`

02 **Cursor reviewer default effort lowered** Not checked yet BREAKING

88

A bare `cursor` reviewer now runs Cursor's GPT-5.6 Sol at medium effort instead of high; restore the previous default with `flowctl config set review.backend cursor:gpt-5.6-sol-high`.

Restore the previous high-effort Cursor reviewer default after upgrading, so existing review workflows are not affected by the effort change.

```
$ flowctl config set review.backend cursor:gpt-5.6-sol-high
```

— Exact before/after and runnable restore command given `flow-next-v9.18.0`

03 **Symlinked review criteria now fail closed** Not checked yet

88

BREAKING

On a checkout without symlink support, a `.flow/criteria.md` that git tracks as a symlink now fails reviews closed instead of silently reviewing without criteria; fix by running `git config core.symlinks true` and `git checkout -- .flow/criteria.md`, or by copying the target's content into the file.

— Exact failure mode and two concrete remediation commands `flow-next-v9.18.0`

04 **Antigravity as read-only review backend** Not checked yet NEW

75

Another host can use Antigravity as a review-only command through an `agy` wrapper that runs only from the absolute path set in `AGY_PATH` and ignores `PATH`; in a live review the wrapper reached SHIP and resumed its prior conversation on re-review.

— Names AGY_PATH env var and describes observed behavior `flow-next-v9.18.0`

05 **Cross-model reviewer connections behind review gates**

Not checked yet NEW

65

Connect Gemini, Grok, Kimi, DeepSeek, or a custom reviewer behind every review gate using a local wrapper, with automatic re-review triggered when cited files change between spec and build.

– Names supported reviewers but no config key or command given `product docs`

06 **Bridge recipe to drive Antigravity from other hosts**

`Not checked yet` **NEW** 60

Claude Code, Codex, or another host can drive Antigravity using the bridge recipe documented in `flowctl usage`, which hands Antigravity a self-contained implementation task with push and history rewrite denied.

– Points to flowctl usage doc but no exact command shown `flow-next-v9.18.0`

07 **Copilot reviewer fallback chain** `Not checked yet` **IMPROVED** 60

A bare `copilot` reviewer now falls back through GPT-6 Astra and then GPT-6 Sol when GPT-6.2 Sol is unavailable, matching the existing `codex` fallback chain.

– Clear mechanism but no user action required or available `flow-next-v9.18.0`

└──▷ **BREAKING ON UPGRADE**

! A bare `cursor` reviewer now runs GPT-5.6 Sol at medium effort instead of high; restore the old default with `flowctl config set review.backend cursor:gpt-5.6-sol-high`.

! On a checkout without symlink support, a `.flow/criteria.md` that git tracks as a symlink now fails reviews closed instead of silently reviewing without criteria; run `git config core.symlinks true` and `git checkout -- .flow/criteria.md`, or copy the target's content into the file.

! `flowctl config set review.backend command` and any command spec containing whitespace are now refused at set time; rename the reviewer using the `command:<label>` form instead.

! On Antigravity, install with `./scripts/install-antigravity.sh` rather than a plain `agy plugin install` of the repository; a plain install loads skills but skips 19 of the 21 agents and misidentifies the host.

WAS THIS USEFUL? 👍 👎

juggler

Feedback ⓘ v0.8.2 NOTES ↗ RANK

The Juggler Code Agent

> Juggler v0.8.2 adds a full git client, multi-provider web search, manual skill invocation control, and prompt history, alongside several window-management and pinboard UI refinements.

└──▷ **WHAT SHIPPED · 9 FEATURES** most completely described first ? what's the number?

01 Skill invocation control via disable-model-invocation

Not checked yet NEW 75

Setting `disable-model-invocation: true` in a skill prevents the model from invoking it autonomously; it can then be started manually with `$name`.

Prevent the model from autonomously invoking a sensitive skill — reserve it for manual use only.

```yaml
disable-model-invocation: true
```

— Names exact config key, manual invocation syntax, and includes example. v0.8.2

02 Built-in git client

Not checked yet NEW 75

Juggler's git integration now covers staging, commits, branches, worktrees, push/pull, a commit graph, reword/squash/drop, blame, reflog undo, conflict resolution, and auto-generated commit messages.

— Lists many git operations but no exact commands or UI path. v0.8.2

03 Multi-provider web search including Exa's WebSearch tool

Not checked yet NEW 70

Web search now supports Tavily, Brave, SearXNG, DeepSeek, Bocha, and Zhipu as providers, selectable in your preferred order. Exa now answers `WebSearch` directly, replacing its separate `exa_search` tool, and existing keys carry over.

— Names six providers and the exact tool-name change for Exa. v0.8.2

04 Popped-out pinboard window behavior

Not checked yet IMPROVED 70

Popped-out pinboard windows now have their own zoom buttons, reopen only on the machine that popped them out, work in windows connected to another machine's

server, and stay open when a conversation is deleted, dropping only the Plan and Todo tabs.

— Details four distinct behavior changes for popped-out windows. v0.8.2

05 **Prompt history panel** Not checked yet NEW 65

Adds a prompt history button to search, reuse, or delete past prompts; use ⌘S or Ctrl+S to set a draft aside.

— Gives exact keyboard shortcuts and button purpose. v0.8.2

THINNER COVERAGE BELOW

06 **Settings → Connectivity panel** Not checked yet NEW 55

Adds Settings → Connectivity to show where the server is and display connected clients relative to it.

— Names exact settings path but little mechanism beyond display. v0.8.2

07 **Pinboard tab switching shortcut** Not checked yet NEW 55

Use Ctrl+Tab to switch between pinboard tabs while the pinboard is open or popped out.

— Gives exact keyboard shortcut, minimal further detail. v0.8.2

08 **New window opening methods** Not checked yet NEW 50

Opening a project from the project chooser now opens it in a new window, launching Juggler while it's already running opens a new window, and the Windows taskbar menu adds a New Window option.

— Describes three window-opening entry points without deeper mechanism. v0.8.2

09 **Windows auto-update reopen and fallback** Not checked yet IMPROVED 45

On Windows, an update that completes now reopens Juggler automatically; a failed update offers the installer instead.

— Describes before/after behavior without further detail. v0.8.2

WAS THIS USEFUL?  

mvschwarz OpenRig

Feedback ⓘ v0.7.20 NOTES ↗ RANK

OpenRig is a network manager for AI coding agents: build persistent teams with roles, shared context and owned work.

> OpenRig's v0.7.20 release focuses on daemon resilience — a tmux-outage delivery hold and automatic Claude usage-limit retry scheduling — alongside pod-bound workflow instances, an opt-in Claude status-line throttle, and new seat-pool, provider-binding, evidence-reference and queue-inspection capabilities.

└──▷ **WHAT SHIPPED · 9 FEATURES** most completely described first ? what's the number?

01 **Tmux-outage delivery hold and status visibility** Not checked yet

NEW 78

When the tmux server a seat depends on disappears, the daemon holds automatic delivery, probes read-only every 5 s backing off to 60 s, and resumes delivery when the server answers. Outage status is visible across `rig status`, `rig ps`, and `rig doctor` while delivery is held.

— Detailed mechanism with exact backoff intervals and named commands v0.7.20

02 **Full task resolution view in queue show** Not checked yet NEW

75

Read a handed-over task's resolution — including the return route and result — with `rig queue show <id> --full`.

Inspect the full result and return route of a handed-over task after your agent completes it.

```
$ rig queue show <id> --full
```

— Runnable command with worked example v0.7.20

03 **Evidence reference on workflow continuation steps** Not checked yet

NEW 73

Use `--evidence-ref` on a workflow continuation step to supply a required receipt, as shown in `rig workflow show`, `rig workflow route`, and `rig workflow resume`.

— Names flag and three commands it appears in v0.7.20

04 **Seat-pool declaration with sizing planner** Not checked yet NEW

60

Declare a seat-pool block on a model member node, stored via migration `178_seat_pool`; a sizing planner for the pool is included.

— Names migration id but no usage example given v0.7.20

THINNER COVERAGE BELOW

05 **Automatic retry on Claude usage-limit wake refusal**

Not checked yet NEW

50

When a parked row's wake is refused at a Claude seat's usage limit, the daemon now retries the wake automatically at the stated reset time.

— Explains trigger and behaviour but gives no command v0.7.20

06 **Per-seat provider bindings view** Not checked yet NEW

45

View per-seat provider bindings with `rig provider`.

— Names command but no further mechanism described v0.7.20

07 **Fresh seat selection on start** Not checked yet NEW

45

Select fresh seats when starting with `rig start`.

— Names command but gives no mechanism or options v0.7.20

08 **Opt-in Claude status-line collection throttle** Not checked yet

NEW

35

An opt-in per-seat throttle limits how often the Claude status line is collected.

— No config key, flag or interval named v0.7.20

09 **Pod-bound workflow instances** Not checked yet NEW

35

A workflow instance can be bound to a pod so it stays on that one pod across its lifecycle.

— No command or flag given for enabling this v0.7.20

WAS THIS USEFUL? 👍 👎

Earendil Works Pi

Feedback ⓘ v1.6.0 NOTES ↗ RANK

Pi is an AI coding agent toolkit that provides a unified LLM API, agent loop, TUI interface and CLI.

> Pi added GPT-6.2 Astra and Sol as default models across the OpenAI, Azure, and Codex providers, expanded its third-party package ecosystem with sandboxing, Datadog, vision, and OAuth plugins, and shipped several breaking changes to its pi-ai and extension APIs alongside TUI and SDK improvements.

└──▷ **WHAT SHIPPED · 19 FEATURES** most completely described first ? what's the number?

01 **pi-ai global API moved to compat entrypoint** Not checked yet

90

BREAKING

Pi's global API functions `stream`, `complete`, `completeSimple`, `getModel`, `getModels`, `getProviders`, `registerApiProvider`, and `getEnvApiKey` are no longer on the root entrypoint of pi-ai; extension sources that typecheck against pi-ai's published types must switch imports to `@earendil-works/pi-ai/compat` (slated for removal) or migrate to the new createModels() / provider-factory API. Inherited selective-provider entrypoints are also removed in favor of root packages with explicit provider factories.

— Exact functions, import paths and migration path given product docs

02 **New Qwen/DeepSeek models and per-plan API keys** Not checked yet

NEW

88

Adds `deepseek-v4.1-flash`, `deepseek-v4.1-flash-1017`, and `qwen4-max-preview` to the Qwen Token Plan Individual providers, alongside new per-plan environment variables `QWEN_TOKEN_PLAN_TEAM_API_KEY`, `QWEN_TOKEN_PLAN_TEAM_CN_API_KEY`, `QWEN_TOKEN_PLAN_INDIVIDUAL_API_KEY`, and `QWEN_TOKEN_PLAN_INDIVIDUAL_CN_API_KEY`; the existing `QWEN_TOKEN_PLAN_API_KEY` and `QWEN_TOKEN_PLAN_CN_API_KEY` still work.

Use per-plan API keys to authenticate against the Qwen Token Plan Individual gateway without sharing a single team key across environments.

```
$ export QWEN_TOKEN_PLAN_INDIVIDUAL_API_KEY=sk-...
pi
```

— Exact model ids, env vars and runnable example given v1.6.0

03 **GPT-6.2 Astra and Sol models added** Not checked yet NEW

85

GPT-6.2 Astra and GPT-6.2 Sol are added to the `openai`, `azure`, and `openai-codex` providers; Azure serves them with a 1.05M context window, Codex serves GPT-6.2 Sol with 500K and GPT-6.2 Astra with 272K, and both reject reasoning effort `none` and `minimal` so thinking level `off` is unavailable. GPT-6.2 Sol (`gpt-6.2-sol`) is now the default model for the `openai`, `azure`, and `openai-codex` providers, replacing GPT-6.1 Sol at the same price.

— Exact context windows, providers and default change named `v1.6.0`

---

04 **find tool path patterns now relative to search dir**

`Not checked yet` `BREAKING`

83

The find tool's path patterns are now relative to the search directory, like grep's `glob`: `src/*.ts` finds `src/a.ts` but no longer `pkg/src/a.ts`; use `**/src/*.ts` to match at any depth. Existing patterns that relied on matching through parent folders above the search path will silently return fewer results.

— Exact pattern examples and migration syntax given `v1.6.0`

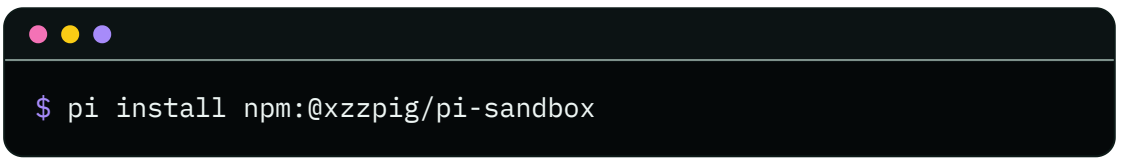
---

05 **New community packages and integrations** `Not checked yet` `NEW`

81

Adds new third-party community packages: `@xzzpig/pi-sandbox` for OS-level sandboxing with interactive permission prompts; `pi-deltabox` for agent-native sandbox sessions with diff, commit, and abort support (`pi install npm:pi-deltabox`); `@datadog/pi-plugin` for querying Datadog logs, metrics, traces, and dashboards through natural conversation (`pi install npm:@datadog/pi-plugin`); `pi-polish` for Claude Code-style output with folded tool calls and colored diff edits; `@xzzpig/pi-starline` for a Starship-inspired statusline and Opencode-style TUI; `@heyhuynhgiabuu/pi-pretty` for pretty terminal output with syntax-highlighted file reads, colored bash output, and tree-view directory listings (`pi install npm:@heyhuynhgiabuu/pi-pretty`); `@getpipher/vision` to delegate image analysis to a vision model only when the active model is text-only, passing images through natively for multimodal models (`pi install npm:@getpipher/vision`); `pi-telephone` for calling agent sessions across machines on a Tailscale network; `pi-runline` for turning any API or command into a callable agent action; and `@gotgenes/pi-anthropic-auth` for Anthropic OAuth compatibility (`pi install npm:@gotgenes/pi-anthropic-auth`).

Isolate agent actions at the OS level with interactive permission prompts, so the agent cannot touch files or network resources you have not approved.



Connect Pi to your Datadog account to query logs, metrics, and traces through natural conversation during an incident investigation.

```
$ pi install npm:@datadog/pi-plugin
```

– Two named packages with install command and capabilities Named package with runnable install command Three named packages with features, one install command Named package, mechanism, and install command Names package but no install command given Names package but no install command or example Named package with install command, no further detail

product docs

06 **Extension render() returns readonly cached arrays** Not checked yet

77

BREAKING

Extension components' render() now returns `readonly string[]`, and Container returns the same cached array while its children are unchanged. Extensions that modify render output in place — including super.render() in a Container subclass or an embedded component's output — must copy it first.

– Exact method signature and required migration step v1.6.0

07 **Extension events for compaction and prompt lifecycle**

Not checked yet IMPROVED

73

`session_before_compact` and `session_compact` events now include `reason` and `willRetry` fields, letting extensions distinguish manual `/compact`, threshold auto-compaction, and overflow retry flows. New `ui_prompt_start`, `ui_prompt_end`, and `session_compact_failed` events let handlers distinguish active agent work from user-facing prompt waits and expose compaction failure reasons, retry state, source, and error message.

– Names exact event names and fields but needs extension code product docs

08 **TUI keybinding and display controls** Not checked yet NEW

73

Pi 0.80.0 adds Ctrl+J as a default newline keybinding alongside Shift+Enter. Terminal hyperlink, image, and truecolor capability auto-detection can now be overridden via new environment variables and advanced settings. `fullscreenCopyOnSelect` disables automatic fullscreen selection copy, with Ctrl+X to copy the active selection when disabled. A new clickable label jumps to the latest transcript message using the `tui.altScreen.bottom` shortcut while scrolled up in fullscreen.

Disable automatic fullscreen selection copy so accidental mouse drags do not overwrite the clipboard.

```yaml
fullscreenCopyOnSelect: false
```

— Names exact keys and config field with example value product docs

09 **Meta Muse login via /login meta** Not checked yet NEW 70

Log in to Meta (Muse subscription) via `/login meta` with automatic Model API key refresh, plus `META_API_KEY` environment variable support.

Log in to Meta's Muse subscription so Pi can auto-refresh your Model API key without manual rotation.

```
$ /login meta
```

— Exact command and env var with runnable example product docs

10 **Transcript notice for dropped reasoning items** Not checked yet NEW 68

Adds a transcript notice when OpenAI drops reasoning items from a request (for example "OpenAI dropped 2 reasoning items: prefix_mismatch (details in session)"), shown when `showCacheMissNotices` is on and only when more items are dropped than in the previous response.

— Names config flag and example message, no UI path v1.6.0

11 **Chat-template thinking controls for OpenAI-compatible providers**

Not checked yet NEW 65

OpenAI-compatible custom providers can now map Pi thinking levels into `chat_template_kwargs`, enabling vLLM and Hugging Face chat-template models such as DeepSeek to use provider-native thinking controls.

— Names config field and target models but no setup steps product docs

12 **llama.cpp extension default port changed** Not checked yet IMPROVED 60

The llama.cpp extension now connects to `http://127.0.0.1:9931` by default, matching the current llama-server default port, instead of port 8080.

— Exact before/after ports named v1.6.0

THINNER COVERAGE BELOW

13 **Persistent thinking-level selector** Not checked yet NEW

58

Select and persist thinking levels with the new `/thinking` selector; save the selected model as the global default with `Ctrl+S`.

– Names command and keybinding but lacks level list product docs

14 **clear_queue RPC for queued messages** Not checked yet NEW

55

Use the `clear_queue` RPC to retrieve and remove queued steering and follow-up messages.

– Names exact RPC but no call signature shown product docs

15 **GoogleThinkingLevel type renamed** Not checked yet BREAKING

55

The `GoogleThinkingLevel` type is renamed to `GoogleApiThinkingLevel`; code referencing this type must be updated.

– Exact before/after type names given product docs

16 **Anthropic message threads on by default broadened** Not checked yet IMPROVED

50

Anthropic message threads are now on by default for any provider whose `baseUrl` is on `api.anthropic.com`, not only the `anthropic` provider.

– States scope change but no config example v1.6.0

17 **Resume externally managed sessions in SDK** Not checked yet NEW

48

Resume externally managed session entries through SessionManager.inMemory() in the SDK.

– Names method but no usage detail product docs

18 **Native PowerShell tool** Not checked yet NEW

45

A new native `powershell` tool runs Windows command execution, configurable through the SDK and config.

– Names tool but no config keys given product docs

19 **Image MIME type detection helper** Not checked yet NEW

45

Detect image MIME types from files using the new detectSupportedImageMimeTypeFromFile() public library export.

— Names export but no usage example `product docs`

└──▷ BREAKING ON UPGRADE

! Pi's global API functions (`stream` , `complete` , `completeSimple` , `getModel` , `getModels` , `getProviders` , `registerApiProvider` , `getEnvApiKey`) are no longer on the root entrypoint; extension sources typechecking against pi-ai's published types must switch those imports to `@earendil-works/pi-ai/compat` (the compat entrypoint and loader alias will be removed in a future release).

! Inherited selective-provider entrypoints are removed; use the root packages with explicit provider factories instead.

! The `GoogleThinkingLevel` type is renamed to `GoogleApiThinkingLevel` ; code referencing this type must be updated.

! Extension components: render() now returns `readonly string[]` , and `Container` returns the same cached array while its children are unchanged. Extensions that modify render output in place — including super.render() in a `Container` subclass or an embedded component's output — must copy it first.

! The find tool's path patterns are now relative to the search directory: `src/*.ts` no longer matches `pkg/src/a.ts` . Use `**/src/*.ts` to match at any depth. Existing patterns that relied on matching through parent folders above the search path will silently return fewer results.

WAS THIS USEFUL? 👍 👎

kane-cli

Feedback ⓘ 0.10.7 NOTES ↗ RANK

Kane CLI by TestMu AI (Formerly LambdaTest)

> Kane CLI 0.10.7 adds a full bug-tracking command set tied into RCA output, named explorations for easier retrieval, Windows support for remote test runs, and custom header passthrough in test runs.

└──▷ **WHAT SHIPPED · 4 FEATURES** most completely described first ? what's the number?

01 Bug tracking commands feeding RCA output

Not checked yet NEW

86

New `kane-cli bugs` commands let you record known bugs against a test so the next RCA run reads them back; `kane-cli bugs categories` lists, adds, and manages org-specific root-cause categories; `kane-cli bugs retire` withdraws previously submitted bug feedback. Known bug IDs and org category names are now carried through to `failure.yaml` so triage output matches recorded data.

List your org's RCA categories before tagging a failure, so you apply a consistent label from your own taxonomy.

```
$ kane-cli bugs categories
```

— Names all subcommands and the output file they affect 0.10.7

02 Named explorations

Not checked yet NEW

65

You can label an exploration at creation time with `kane-cli explore run --name <name>`, making runs easy to find in exports.

Start a named exploration so the run is easy to find in exports and retries keep their identity.

```
$ kane-cli explore run --name checkout-happy-path
```

— Clear flag and command but limited mechanism detail 0.10.7

THINNER COVERAGE BELOW

03 Windows support for remote test runs

Not checked yet IMPROVED

55

`testrun --remote` now works from Windows, sending POSIX-style paths to the HyperExecute grid.

— Names flag and platform but no further mechanism 0.10.7

04 **Custom headers in test runs** Not checked yet IMPROVED 45

Custom headers configured for a test are now included in `kane-cli test` runs.

— Brief description, names the command but no config details 0.10.7

WAS THIS USEFUL? 👍 👎

agentbro

Feedback ⓘ v3.2.2 NOTES ↗

Agentbro is a macOS Dynamic Island for AI coding agents: Claude Code, Codex, Gemini CLI, Cursor, hooks, skills, and local agent management.

agentbro: AgentBro now shows QwenWork CN Flash task titles and status — including running and completed — in the Island for tasks run with the Flash engine.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v3.2.2
https://github.com/shirenchuang/agentbro.git
# already have the repo? check out this version:
$ git checkout v3.2.2
```

AgentBro now surfaces DeepSeek Harness conversation previews in the Island, using your prompt as the session title and showing previews of the prompt and assistant reply in session details. Not checked yet

└──▷ BREAKING ON UPGRADE

! To update an existing DeepSeek Harness integration, you must quit DeepSeek Harness, remove and reinstall its Hook in AgentBro, then reopen Harness — the old Hook will not work automatically.

WAS THIS USEFUL? 👍 👎

Amazon Kiro Feedback ⓘ

General: Enterprise Management Preview in Europe (Frankfurt) (2026-11-20)

BLOG ↗ RANK

Kiro is an agentic development environment that uses spec-driven workflows to plan, build, and maintain software.

> Kiro expanded beyond its IDE this window with a VS Code extension, a CLI surface, and a Web surface sharing specs, steering, and MCP configuration, while the IDE gained remote SSH, Dev Container, and cloud-session editing, the CLI (v3.5) added VPC-aware cloud sessions, organization MCP sign-in, and a configurable login portal, and the Enterprise Management preview rolled out to Europe (Frankfurt).

└──▷ **WHAT SHIPPED · 9 FEATURES** most completely described first ? what's the number?

01 New Kiro surfaces: VS Code, CLI, and Web Not checked yet NEW

95

Adds three new Kiro surfaces: a VS Code extension (available on the Visual Studio Marketplace and Open VSX for macOS, Windows, and Linux, requiring VS Code 1.145 or later) that brings the Kiro agent into VS Code sharing the same specs, steering, and MCP configuration as the Kiro IDE, storing specs, steering files, skills, hooks, Powers, and MCP configuration in the `.kiro` workspace folder, letting specs become actionable tasks via editor actions and per-task CodeLens, starting sessions in Supervised mode with the option to switch to Autopilot or Approve for me, enabling workflows and memory in VS Code Settings, and adding context to chat and spec sessions via the # shortcut in the sidebar; a CLI surface for terminal-native and headless workflows sharing the same specs, steering, and MCP configuration as the other Kiro surfaces; and a Web surface offering zero-setup sandboxed sessions.

Enable Kiro's workflow and memory capabilities after installing the extension, so agents can track state across sessions.

> 📍 In VS Code, open Settings (Cmd/Ctrl+,), search for Kiro, then enable Workflows and Memory toggles.

Switch from Supervised to Autopilot mode so Kiro agents run and apply changes without manual approval at each step.

> 📍 In the Kiro sidebar, open the mode selector at the top of the chat panel and choose Autopilot or Approve for me.

– Names install source, version requirement, config folder, and mode switches. Only a one-line announcement, no commands or config shown. Brief description, no setup steps or limits given.

Extensions: Introducing Kiro for VS Cod…

02 **VPC connection visibility for cloud sessions** Not checked yet

NEW 95

Display VPC connection details for cloud sessions by setting `--no-interactive --cloud`, which prints a `[cloud] vpc <name> · <vpc-id> · <region>` line on stderr, or use `--output-format stream-json` to emit a `cloudSession` record with `vpcConnection`; the connection name, VPC ID, and Region also appear in `/session-id`.

– Exact flags, output formats and fields are all named.

CLI: Cloud Sessions in Your VPC, Organi…

03 **Organization MCP sign-in and OAuth client ID in CLI**

Not checked yet NEW 90

Set up organization-wide MCP server sign-in in V3 sessions with `kiro-cli mcp idp setup`, supplying your identity provider's issuer and client ID so one sign-in covers every server configured to use it. Add an HTTP MCP server with a specific OAuth client ID using the `--oauth-client-id <client-id>` flag on `kiro-cli mcp add`.

– Both commands and flag given verbatim with config inputs.

CLI: Cloud Sessions in Your VPC, Organi…

04 **Configurable login portal via auth.portalUrl** Not checked yet

NEW 84

Choose which portal browser sign-in opens by setting `auth.portalUrl`; administrators can enforce this via the `AuthPortalUrl` policy or `auth_portal_url` in `managed-settings.json`.

Configure a shared organization login portal so all users authenticate through your company's SSO entry point instead of the default Kiro portal.

```yaml
auth.portalUrl: https://sso.example.com/kiro
AuthPortalUrl: https://sso.example.com/kiro
```

– Exact config key, policy name and example file given.

CLI: Cloud Sessions in Your VPC, Organi…

05 **Enterprise Management preview in Europe (Frankfurt)**

Not checked yet NEW 80

Enterprise administrators can now use the Enterprise Management preview with Kiro profiles in the Europe (Frankfurt) (`eu-central-1`) region, in addition to US East (N. Virginia) (`us-east-1`). Existing profiles in Europe (Frankfurt) using IAM Identity Center or an external identity provider can migrate to Controls when Migrate to the new experience appears on the Kiro Settings page — this move is one-way and cannot be reversed. Newly created Kiro profiles in Europe (Frankfurt) now have Controls enabled from the start.

Migrate an existing Frankfurt profile using IAM Identity Center to the new Controls experience.

> 📍 In the Kiro console, go to Settings, then select your Europe (Frankfurt) profile. When Migrate to the new experience appears, follow the prompt to move to Controls. Note: this migration is one-way and cannot be reversed.

— Names region codes, migration path and UI entry point.

General: Enterprise Management Preview …

---

06 **New IDE connection options** Not checked yet NEW 75

Adds three ways to connect the IDE to remote or containerized environments: Connect to Host lets you connect the IDE to your own Linux machine over SSH so files, terminals, and the agent run on the remote machine while the editor stays local and the agent keeps working after you disconnect; Reopen in Container opens a project inside its Dev Container locally or on an SSH host, building the repository's `.devcontainer` configuration with Docker so you use the repo's tools without installing them locally; and Open in IDE opens a cloud session in a new IDE window, attaching the editor and chat to the agent already running in that cloud session without copying the repository locally.







Connect to a remote Linux dev machine over SSH so the agent can run builds and tests there while you edit locally.

> 📍 In the IDE, open the remote menu and select `Connect to Host`, then choose your configured SSH host or select 'Add New SSH Host...' to register one. A new Kiro window opens on the remote machine's files.

Isolate your project's toolchain inside a Dev Container so you don't install dependencies on your local machine.

> 📍 With a project that has a `.devcontainer` folder open, open the remote menu and select `Reopen in Container`. Kiro builds the container with Docker and reopens the project inside it.

Pick up an in-progress cloud session in a full IDE window without cloning the repository locally.

> 📍 In the Sessions view, find the cloud session you want, then click `Open in IDE`. A new Kiro window opens on the session's files under Projects [Kiro Cloud] with chat already attached to the running agent.

– Exact UI command named, behaviour explained, but no flags/config. Names exact command and devcontainer mechanism. Exact UI action named with clear behaviour, no limits given.

IDE: Remote SSH, Dev Containers, and Cl…

---

07 **Memory access control in V3 sessions** Not checked yet NEW

70

Control memory access level in V3 sessions via `/memories config`, which now sets `read & write`, `read only`, or `off` with a separate `Automatic memory updates` checkbox.

– Names command and exact options, no example shown.

CLI: Cloud Sessions in Your VPC, Organi…

---

08 **Error type fields in V3 stream-JSON output** Not checked yet

IMPROVED

60

Error records in a V3 `--no-interactive` stream-JSON run now name the backend error type in `errorType` and `retryErrorType` fields.

— Names exact fields but no usage example given. CLI: Cloud Sessions in Your VPC, Organi…

THINNER COVERAGE BELOW

09 **Unified Powers section in Kiro Settings** Not checked yet IMPROVED

35

Installed and available powers now share one Powers section under Kiro Settings.

— States the UI location but no further mechanism.

IDE: Remote SSH, Dev Containers, and Cl…

NEW SUPPORTED PLATFORM Grok 4.7 in AWS GovCloud (US)

└──▷ ALSO FROM THESE RELEASES





Models: Grok 4.7 now available in AWS GovCloud (US): Grok 4.7 is available with experi

BLOG ↗ STABLE

> Grok 4.7 adds a 500K-token context window to the Kiro IDE and CLI, with tiered credit multipliers of 2.7x (up to 200K tokens) and 5.4x (above 200K tokens).
>
> Grok 4.7 is now available with experimental support in the Kiro IDE and CLI across AWS GovCloud (US-East) and AWS GovCloud (US-West), selectable from the model selector after updating to the latest version. Not checked yet
>
> Grok 4.7 offers a 500K-token context window, with a 2.7x credit multiplier for requests up to 200K tokens and a 5.4x multiplier for requests exceeding 200K tokens. Not checked yet

---

WAS THIS USEFUL? 👍 👎

Cognition Devin Desktop

Feedback ⓘ

announced · Nov 19, 2026 BLOG ↗ RANK

Devin Desktop is an agent-native code editor for developing software locally, remotely, and in containerized environments.

> Devin Desktop shipped a Databricks Migrations integration for converting and reconciling legacy data platform code, alongside support for the new Gemini 4 Neon model.

└──▷ **WHAT SHIPPED · 2 FEATURES** most completely described first ? what's the number?

01 **Databricks Migrations integration** Not checked yet NEW

73

Devin Desktop now supports Databricks migrations, converting code from legacy sources — including Teradata, Oracle, SQL Server, Informatica, Redshift, and Snowflake — to Databricks SQL, PySpark, Spark, and Declarative Pipelines, with output governed in Unity Catalog. It also runs pre-cutover reconciliation across schema, row counts, aggregates, and row-by-row diffs to verify every source against its Databricks copy before migration completes. Devin Desktop is now available on the Databricks Marketplace, letting customers apply purchases toward their existing Databricks Universal Commit.

— Names sources, targets, Unity Catalog, marketplace; no commands or UI path

launch-20261119-8e8ff481

THINNER COVERAGE BELOW

02 **Gemini 4 Neon model support** Not checked yet NEW

55

Devin Desktop and Devin CLI now support Gemini 4 Neon as a selectable model, scoring 46.2% on the FrontierCode 2.0 benchmark at approximately $4.38 per task.

— Names model and benchmark score but no selection steps launch-20261118-049cdcef

WAS THIS USEFUL? 👍 👎

Sourcegraph Amp

Feedback ⓘ announced · Nov 17, 2026 BLOG ↗

RANK

Amp is Sourcegraph's agentic coding tool for the terminal and editor, running multi-step edits with subagents and shared team threads.

> Amp shipped a native Windows desktop app, extending its native app lineup to Windows users.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **Native Windows desktop app** Not checked yet NEW 30

Amp now has a native Windows desktop app, joining its existing native app lineup.

— Names the platform but no details on features, download, or mechanism.

launch-20261117-7dae0abc

WAS THIS USEFUL? 👍 👎

◆ AI App Builders

Vercel v0

Feedback ⓘ updated · Nov 21, 2026 NOTES ↗ RANK

The v0 generative UI platform builds web applications from natural-language prompts.

> v0 introduces enterprise Spend Caps, letting team owners set per-member credit spending limits for each billing cycle.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **Spend Caps for enterprise team members** Not checked yet NEW

60

Enterprise team owners can set per-member credit spending limits for a billing cycle via Settings → Members, applying either a default limit to all members or customizing a limit per individual.

Cap how much shared-credit-pool usage any member can consume in a billing cycle, so enterprise budgets don't overrun.

> 📍 In v0, go to Settings → Members, then set a default spending limit for all members or select an individual to customise their cap.

— Names exact UI path and behavior but no numeric limits given snapshot-20261121

WAS THIS USEFUL? 👍 👎

Lovable

Feedback ⓘ Nov 19, 2026 (2026-11-20) BLOG ↗ RANK

Lovable builds web apps and sites from natural-language prompts.

> Lovable's latest releases add seven payment connectors spanning global and regional markets, two new GPT-6.2 model options, and a change to keep building through unanswered questions instead of blocking.

└──▷ **WHAT SHIPPED · 7 FEATURES** most completely described first ? what's the number?

01 **Payment connectors for GoCardless, Mercado Pago, Razorpay, Paystack and Square**

Not checked yet NEW 85

Added `GoCardless` for one-off bank payments, recurring direct debits, and monthly or yearly subscriptions (suited to memberships, rent collection, invoicing); Mercado Pago for Checkout Pro, Pix, and card payments with confirmed-order marking and refunds (Latin America); Razorpay for UPI, card, netbanking, and wallet payments via Razorpay Checkout or payment links, refunds, and free-trial subscriptions (India); Paystack for card, bank transfer, and mobile money payments, subscriptions, and refunds (Nigeria, Ghana, Kenya, South Africa, Côte d'Ivoire); and Square for payment links, card charges, refunds, order and customer management, and Square Terminal checkouts. All can be added from the Connectors panel or by asking Lovable directly in a project.

Connect Razorpay to a new project without leaving the Lovable chat — useful when building a membership or event-ticketing app targeting India.

📍 In the Lovable project chat, type: Connect Razorpay to this project, then follow the prompts. Alternatively, go to Connectors, search for Razorpay, and click Connect.

Connect GoCardless to collect recurring direct debits for a subscription or rent-collection app.

📍 In the Lovable project chat, type: Connect GoCardless to this project, then follow the prompts. Alternatively, go to Connectors, search for GoCardless, and click Connect.

Add payment collection to an app serving markets in Nigeria or Kenya by wiring in Paystack through the Connectors panel.

📍 1. In your Lovable project, open Connectors. 2. Select Paystack and follow the prompts to authenticate. 3. Ask Lovable: 'Add a checkout button that uses Paystack to accept card and mobile money payments.'

— Names five connectors, payment types, regions and add-via-chat path

Nov 20, 2026 (2026-11-20) Nov 19, 2026 (2026-11-20)

02 **Google Maps Platform billing moves to Lovable credits**

Not checked yet

65

BREAKING

New Google Maps Platform connections set up with Managed by Lovable now use Lovable credits instead of a separate billing account. Managed connections created before November 19, 2026 stop working after December 15, 2026 unless the creator accepts the updated usage terms in the project chat or on the connection.

— Names exact deadline and required action to avoid breakage Nov 19, 2026 (2026-11-20)

03 **GPT-6.2 Astra and Sol model options** Not checked yet NEW

60

Apps built in Lovable can now use `openai/gpt-6.2-astra` for complex reasoning, coding, research, and document creation at lower cost than GPT-6 Astra, and `openai/gpt-6.2-sol` for everyday reasoning, coding, and professional work; both accept text and image input and respond with text.

— Names exact model identifiers but no selection UI or command shown

Nov 20, 2026 (2026-11-20)

04 **Custom audience publishing by email domain** Not checked yet NEW

60

Published apps can be restricted so only users with a verified corporate email domain can access them, without adding each person to the workspace, by selecting Custom audience in the publish dialog and entering the verified domain.

Restrict a published app so only staff with a verified corporate email address can access it, without adding each person to your workspace.

📍 1. Open the publish dialog for your app. 2. Under Audience, select Custom audience. 3. Enter your verified email domain (e.g. yourdomain.com) and publish.

— Exact publish-dialog steps given, mechanism described but plan scope unclear

Nov 20, 2026 (2026-11-20)

THINNER COVERAGE BELOW

05 **Continues building past unanswered questions** Not checked yet

IMPROVED

55

Lovable now continues building whatever does not depend on an unanswered question, picks a default for blocked steps, and reworks only what needs to change if a different choice is made later — but still waits before irreversible actions such as deleting data.

— Explains mechanism and limits but no user-facing action to take

Nov 19, 2026 (2026-11-20)

06 **Free first-year domain claim on paid upgrade** Not checked yet

NEW

55

Workspace admins and owners who upgrade from Free to a paid plan can now claim one domain free for the first year, covering extensions such as `.app` and `.dev`, if claimed within 90 days of first payment.

— Gives eligibility window and extensions but no claim steps Nov 20, 2026 (2026-11-20)

07 **Switch Google accounts without signing out** Not checked yet

IMPROVED

40

Users who log in with Google can now switch to another Google account from the user menu on lovable.dev without signing out.

— Clear UI location named but minimal mechanism detail Nov 20, 2026 (2026-11-20)

└──▷ ALSO FROM THESE RELEASES

. . . .

└──▷ BREAKING ON UPGRADE

! Managed Google Maps Platform connections created before November 19, 2026 stop working after December 15, 2026 unless the creator accepts the updated usage terms in the project chat or on the connection.

WAS THIS USEFUL? 👍 👎

◆ Speech & Voice AI

ElevenLabs

Feedback ⓘ announced · Nov 20, 2026 BLOG ↗ RANK

ElevenLabs is a voice AI platform with text-to-speech, speech-to-text, voice agents, dubbing and music generation, available in the app and through an API.

> ElevenLabs shipped a major Ads Engine expansion — TikTok integration plus AI-generated video ad variants, soundtracks and multi-market localization — alongside a healthcare push for ElevenAgents with real-time medical transcription, HIPAA-eligible zero-retention calls, an agent-building assistant, and guardrails/testing tools.

└──▷ **WHAT SHIPPED · 8 FEATURES** most completely described first ? what's the number?

01 **Scribe v3 Live Medical transcription for agent calls**

Not checked yet NEW 80

ElevenAgents now supports Scribe v3 Live Medical, a real-time medical transcription model optimised for drug names and clinical terms in live agent calls, with mid-call language switching so a caller who gives a member ID in English and continues in Spanish is transcribed correctly throughout, and support for adding up to 10,000 custom terms — plan names, therapies, devices, locations — to improve accuracy for organisation-specific vocabulary.

— Model name, term limit, and language example all given launch-20261118-29ab9a68

02 **Zero Retention Mode for HIPAA-eligible deployments**

Not checked yet NEW 75

Teams can enable HIPAA-eligible deployments by signing a BAA and turning on Zero Retention Mode so audio and transcripts are not stored on ElevenLabs infrastructure.

— Exact config toggle and precondition (BAA) named launch-20261118-29ab9a68

03 **AI-Generated Video Ad Content** Not checked yet NEW 67

Ads Engine now generates new AI-driven video ad content: alternative hook openings with voiceovers from a digital voice replica of the original speaker, lip-synced to the new audio and captioned to match, plus original soundtracks via the `ElevenLabs Music API` that analyze the tone, pace, and emotion of the visuals, add sound effects on key moments, and are cleared for broad commercial use.

— Clear multi-step mechanism but no concrete API or UI path Names Music API and mechanism, no runnable step shown

launch-20261120-5b9157bb

04 **Video Ad Localization** Not checked yet NEW 60

Ads Engine adds video ad localization via Dubbing v2, translating and dubbing each ad with lip-syncing to the new audio and optional routing to a native-speaking localization professional for accuracy review, along with translated caption re-rendering that removes burned-in originals and preserves the original ad's timing and brand fonts in the new language.

— Names Dubbing v2 and review option, no config detail Specific mechanism described, no actionable control surface

launch-20261120-5b9157bb

THINNER COVERAGE BELOW

05 **Healthcare Receptionist agent template** Not checked yet NEW 55

ElevenAgents adds a Healthcare Receptionist template with system prompt, call flow, and integration setup pre-configured so teams can reach a working agent in minutes.

— Named template with clear starting point, minimal detail launch-20261118-29ab9a68

06 **ElevenAgents Architect assistant** Not checked yet NEW 52

ElevenAgents adds the `ElevenAgents Architect` built-in AI assistant, letting teams describe agent behaviour in plain language and have changes edited and tested before going live.

— Names the assistant but no UI path or command shown launch-20261118-29ab9a68

07 **Agent Guardrails & Testing** Not checked yet NEW 52

ElevenAgents adds custom guardrails, letting teams write plain-language rules — for example, routing any symptom question to a clinician — in addition to built-in topic guardrails, and simulated call testing so teams can verify that an agent follows its configured rules before launch.

— Concrete example given, no exact configuration syntax Bare description with no mechanism or UI path

launch-20261118-29ab9a68

08 **TikTok integration in Ads Engine** Not checked yet NEW 38

Ads Engine now connects to TikTok ad accounts in addition to Google, Meta, and LinkedIn, pulling existing creative and performance data from the connected platform.

— Names four platforms but no setup steps given launch-20261120-5b9157bb

WAS THIS USEFUL?  

◆ Coding Agent Add-ons

planning-with-files

Feedback ⓘ v3.29.0 NOTES ↗ RANK

planning-with-files is a persistent file-based planning system for AI coding agents with crash-proof markdown plans, session recovery, and context-rot protection.

> planning-with-files added checkbox-only re-attestation enrollment, letting plan snapshots be re-verified when only checkbox state changes.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **Checkbox-only re-attestation enrollment** Not checked yet NEW

40

Plans can now be enrolled for checkpoint verification that accepts only checkbox changes, validated against the same snapshot hash, using a shell helper.

— Describes mechanism but no exact command or flag name given v3.29.0

WAS THIS USEFUL?  

Paradev

Feedback ⓘ desktop-v2.0.11 NOTES ↗ RANK

Paradev is a desktop workspace for running CLI coding agents like Claude Code in parallel, with terminals, diff viewer, and browser previews.

> Paradev's latest desktop release focuses on billing usage visibility and cloud credential handling, alongside smaller additions for admin operations, Vercel Workflow environments, subagent navigation, and theming.

└──▷ **WHAT SHIPPED · 7 FEATURES** most completely described first ? what's the number?

01 **Per-member cloud usage tracking in Billing** Not checked yet NEW

70

Paradev adds a usage card on the Billing page showing each member's own cloud usage. Under the hood it adds a `cloud_usage_daily` table kept in sync by a trigger on `cloud_usage_events`, a backfill function scoped to one billing day, and a `billingDay` group-by so usage is served from the daily rollup, bucketed in Pacific-time billing days.

– Names tables, trigger and group-by but UI path only, no command desktop-v2.0.11

THINNER COVERAGE BELOW

02 **Credential rule push to cloud boxes** Not checked yet IMPROVED

50

Paradev now pushes credential rules to cloud boxes immediately on change and before expiry, replacing the previous 30-minute credential poll.

– Clear before/after mechanism but no config surface named desktop-v2.0.11

03 **GB-hour rate in periods.list API** Not checked yet IMPROVED

50

Paradev's `periods.list` API now returns the organisation's GB-hour rate alongside period data.

– Names exact API but gives no further detail on usage desktop-v2.0.11

04 **Subagent navigation UI** Not checked yet NEW

45

Paradev now shows when you are inside a subagent view and provides a navigation path to get back to the parent.

– Describes behaviour but no exact UI controls named desktop-v2.0.11

05 **Vercel Workflow environment builds** Not checked yet NEW

40

Paradev now builds environments on Vercel Workflow with a thin integration path.

— Names the integration but mechanism left vague `desktop-v2.0.11`

06 **Paradev admin connector** `Not checked yet` NEW 25

Paradev adds the Paradev admin connector, extending its connector surface to cover administrative operations.

— Named but no detail on what operations it covers `desktop-v2.0.11`

07 **Custom theme background hooks** `Not checked yet` NEW 25

Paradev adds custom theme background hooks to the desktop app, providing stable extension points for theming.

— Bare mention of hooks with no usage path `desktop-v2.0.11`

WAS THIS USEFUL? 👍 👎

◆ Image & Video Generation

Comfy Org ComfyUI

Feedback ⓘ announced · Nov 18, 2026

BLOG ↗ RANK

ComfyUI is a node-graph interface, API, and backend for running diffusion models, chaining generation steps into reusable visual workflows.

> ComfyUI's Comfy Cloud now supports Custom Builds, letting teams package and share custom nodes, models, and LoRAs (including private node packs) for cloud GPU execution without local setup.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **Custom Builds in Comfy Cloud** Not checked yet NEW 91

Comfy Cloud now supports Custom Builds, packaging a team's custom nodes, models, LoRAs, and their dependencies (including private node packs) into a shared setup available on every paid plan. Builds can be packaged starting from a workflow or a Comfy Desktop snapshot via the Developer Platform, with command-line packaging also supported, and workflows created with a Custom Build can also be run via the API for developer automation alongside artist use. GPU selection ranges from RTX PRO 6000 (48 GB or 96 GB) up to H100, H200, and B200 (180 GB) with per-second billing while workers are running, and concurrent job limits scale by plan: 2 simultaneous jobs on Standard and Creator plans, up to 20 on Team.

Bring a new artist onto a project so they immediately have the team's custom nodes, models, and LoRAs without local setup.

📍 In Comfy Cloud, go to Developer Platform › Custom Builds › New Build, select Start from a Comfy Desktop snapshot, resolve any dependency conflicts, choose a GPU (e.g. H100), and click Deploy. Team members then open the editor, pick the Custom Build from the setup selector, and start running workflows.

— Names GPUs, plan limits, and exact UI steps for setup launch-20261118-78c5c1c0

WAS THIS USEFUL? 👍 👎

AGENT

◆ AI Agent Frameworks

Arcade

Feedback ⓘ updated · Nov 21, 2026 NOTES ↗ RANK

Arcade provides an actions runtime that lets AI agents securely call external services with authorization and governance controls.

> Arcade's latest release adds chart-level controls for blocking toolkits and forwarding OpenFGA telemetry, broadens tool execution history, and ships a cluster of security hardening changes (credential revocation on suspension, OAuth browser binding, log scrubbing, internal-only service defaults) alongside chart value removals in 1.15.

└──▷ WHAT SHIPPED · 10 FEATURES most completely described first ? what's the number?

01 Chart options for toolkit blocking and OpenFGA telemetry

Not checked yet NEW 90

New Helm chart values let operators block named toolkits from being served fleet-wide via `workerDefaults.supervisor.disabledToolkits` (with a per-worker override), and forward OpenFGA's own internal traces to the configured collector via `rbac.openfga.telemetry.enabled`, which is off by default.

Prevent a specific toolkit from being loaded by any worker, e.g. to block an untrusted or temporarily broken toolkit fleet-wide.

yaml

```yaml
workerDefaults:
  supervisor:
    disabledToolkits:
      - telegram
```

Send OpenFGA's internal traces to your configured collector so you can correlate authorization decisions with tool execution spans.

yaml

```yaml
rbac:
  openfga:
    telemetry:
      enabled: true
```

— Exact config keys with runnable YAML snippets provided snapshot-20261121

02 Removed deprecated chart values in 1.15

Not checked yet BREAKING

70

Chart 1.15 removes the `openaiApiKey` and `openaiApiKeySecretRef` chart values, removes the `workers[].http` chart value, and removes `supportImages.busybox` (deprecated since chart 1.12).

— Lists exact removed chart keys for migration snapshot-20261121

03 **OIDC discovery routing fix for sign-in** Not checked yet IMPROVED

65

Sign-in now follows `oidcDiscoveryUrl` to the discovery host, so the Experience API's OIDC token exchange works even when the Coordinator is unreachable at its public hostname from inside the cluster.

— Names exact config key and affected internal components snapshot-20261121

THINNER COVERAGE BELOW

04 **Execution history records client IP and retains failed runs**

Not checked yet IMPROVED 55

Tool execution history now records the caller's client IP, which is included in responses, and persists executions that time out, are cancelled, or are refused by a pre-execution hook through retention sweeps.

— Behavior described but no field name or endpoint given snapshot-20261121

05 **Legacy MCP handshake fallback for compatibility** Not checked yet

IMPROVED 55

Falls back to the legacy MCP initialize handshake when a remote server's discovery answers with a bare HTTP 500 or 501, restoring compatibility with servers such as OpenMetadata.

— Names exact HTTP codes and affected server, no user action needed snapshot-20261121

06 **Credential scrubbing from logs and traces** Not checked yet

IMPROVED 50

Credentials embedded in tenant- or client-supplied URLs (MCP/worker, OAuth resource/redirect/JWKS) are now scrubbed from logs, traces, audit records, and Sentry above debug level.

— Names affected URL types but no reader action required snapshot-20261121

07 **Reduced secrets mounted into pods** Not checked yet IMPROVED

50

Worker and dashboard pods no longer mount the full application-secrets Secret — unused DB, Redis, and signing-key values are excluded.

— Names excluded secret types but offers no user action snapshot-20261121

08 **Internal-only networking for core services** Not checked yet

IMPROVED

40

OpenFGA, the usage collector, and toolkit workers now default to accepting connections only from calling Arcade components.

— States the change but gives no config to adjust it snapshot-20261121

09 **Suspension revokes all issued credentials** Not checked yet

IMPROVED

40

Suspension now revokes already-issued API keys, account tokens, and OAuth sessions across the organization.

— Clear scope but no command or flag to trigger it snapshot-20261121

10 **OAuth flows bound to originating browser** Not checked yet IMPROVED

30

OAuth authorization flows are now bound to the browser that started them, preventing cross-browser code completion attacks.

— Security fix described with no config surface snapshot-20261121

└──▷ **BREAKING ON UPGRADE**

! `openaiApiKey` and `openaiApiKeySecretRef` chart values are removed in chart 1.15.

! `workers[].http` chart value is removed in chart 1.15.

! `supportImages.busybox` chart value is removed (deprecated since chart 1.12).

WAS THIS USEFUL?  

platypus

Feedback ⓘ v4.6.0 NOTES ↗ RANK

Platypus is a self-hosted AI Agents for your whole team — on your infrastructure, your models, around the clock.

> Platypus v4.6.0 adds job-level control over which models the Provider uses for task execution and memory extraction.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **Per-job model settings for task and memory extraction**

Not checked yet NEW 42

Job settings now let you configure the Provider's task model and memory extraction model on a per-job basis, rather than relying on a single global setting.

— Names the setting area but no exact config key or UI path v4.6.0

WAS THIS USEFUL? 👍 👎

claw-orchestrator

Feedback ⓘ v7.29.0 NOTES ↗ RANK

Claw-orchestrator is a run Claude Code, Codex, Antigravity, Grok Build and OpenCode behind one runtime, and a run only counts as done when checks it didn't write pass. Cross-engine fan-out, councils, session handoff, spend caps, MCP server, ACP agent.

> claw-orchestrator refines its forum mention handling in v7.28.1, filtering out non-member mentions and flagging likely-misdirected ones.

└──▷ WHAT SHIPPED · 1 FEATURE most completely described first ? what's the number?

01 **Member-scoped forum mention filtering** Not checked yet IMPROVED

67

`clawo forum post` now filters mentions to known forum members only — non-member `@word` tokens such as `@0.88` are excluded from post metrics — and flags names that look intended for a recipient but are not actually on the forum.

Review a post for non-member mention names that claw-orchestrator flagged as likely intended recipients not on the forum.

```
$ clawo forum post <post-id>
```

— Names exact command and filtering behaviour with example. v7.28.1

NEW INTEGRATION Gemini 4 Neon engine registration

UPDATED SUPPORTED PLATFORM Tested engine versions

v7.29.0 NOTES ↗ LATEST STABLE

claw-orchestrator adds image input to `session_send` and wires new flags across Claude Code, Codex, Antigravity, Grok, and OpenCode engines.

└──▷ GET THIS VERSION ▾

npm

```
$ npm install -g @enderfga/claw-orchestrator@7.29.0
```

```
git

$ git clone --branch v7.29.0 https://github.com/Enderfga/claw-orchestrator.git
# already have the repo? check out this version:
$ git checkout v7.29.0
```

signed build provenance · checked 21 Nov Wrong command?

Send images in `session_send` by passing an `images` list of absolute paths (PNG, JPEG, GIF, or WebP) for Claude Code, Codex, `codex-app`, Grok, and OpenCode engines; any other engine refuses the message. Not checked yet

Support new Claude Code settings `strictMcpConfig`, `settingSources`, `systemPromptSnapshot`, `pluginDir`, `autoCompactWindow`, `appendSubagentSystemPrompt`, and `disableSlashCommands`, plus `--no-chrome` unless `chrome` is true. Not checked yet

Support new Codex settings `codexConfig`, `codexFeatures`, and `ignoreRules`, plus forking via `codex exec fork` and the `thread/fork` endpoint; `codex_review` now returns usage data. Not checked yet

Support `jsonSchema`, `addDir`, `noSessionPersistence`, and native system/developer instructions for `codex-app`. Not checked yet

Support new Antigravity settings `jsonSchema`, `addDir`, `agent`, and `disableSlashCommands`. Not checked yet

+2 more

└─▷ **BREAKING ON UPGRADE**

! Grok `allowedTools` and `disallowedTools` are now permission rules mapped to `--allow` and `--deny`, matching Claude Code behaviour — existing Grok tool-permission configs must be updated accordingly.

WAS THIS USEFUL? 👍 👎

HeyGen HyperFrames

Feedback ⓘ v0.9.74 NOTES ↗ RANK

HyperFrames is an open-source HTML-to-video renderer that runs in AI-agent workflows.

> HyperFrames added in-place text editing on sealed preview frames in Studio.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **In-place text editing on sealed preview frames** Not checked yet

NEW 40

Sealed preview frames in Studio now support in-place text editing sessions, letting users edit text directly on a rendered preview.

— Names the UI area but no mechanism or limits v0.9.74

WAS THIS USEFUL? 👍 👎

Vercel AI SDK

Feedback ⓘ ai@7.0.179 NOTES ↗ RANK

The Vercel AI SDK provides TypeScript APIs for model generation, structured output, tool use, and streaming application interfaces.

> Vercel AI SDK added WebSocket-based Live model support through the Gateway, including direct client connections via short-lived tokens, alongside telemetry hooks for `experimental_extract` and a security fix closing a replay hole in automatic tool-call denials.

└──▷ **WHAT SHIPPED · 3 FEATURES** most completely described first ? what's the number?

01 WebSocket Live model support via Gateway

Not checked yet NEW

95

gateway.experimental_realtime() connects to OpenAI Live models over WebSocket for `openai/gpt-live-1` and `openai/gpt-live-1.1`, speaking a normalized Live protocol (`session-start`, PCM16 audio, `context-append`, `session-close` with final usage); pass `{ api: 'live' }` for other Live model IDs or `{ api: 'realtime' }` to force the realtime route. gateway.experimental_realtime.getToken() mints short-lived Live-bound client secrets so a browser client can connect directly to a Live model without an application relay using `experimental_useRealtime`, passing the token through `api.token` when the model declares the `client-secret-websocket` connection.

Start a Live model WebSocket session via the Gateway for real-time audio interaction with GPT Live models.

typescript

```
import { gateway } from '@ai-sdk/gateway';

const liveModel = gateway.experimental_realtime('openai/gpt-live-
1');
// Or for other Live model IDs:
const otherLiveModel =
gateway.experimental_realtime('openai/other-live-model', { api:
'live' });
// Or to force the realtime route:
const realtimeModel = gateway.experimental_realtime('openai/gpt-
live-1', { api: 'realtime' });
```

Mint a short-lived Live-bound client secret so a browser client can connect directly to a Live model without an application relay.

```typescript
import { gateway } from '@ai-sdk/gateway';

const token = await
gateway.experimental_realtime.getToken('openai/gpt-live-1');
// Pass token.api.token to the client for use with
experimental_useRealtime
```

— Names the protocol messages, options and includes runnable code `ai@7.0.179`

02 **Automatic tool-call denials no longer replayable** Not checked yet

75

BREAKING

Automatically denied tool calls in `generateText` and `streamText` with `experimental_toolApprovalSecret` configured are no longer signed, so a client cannot replay a denial response to force execution of a rejected tool call. Signatures issued before this change for automatic denials remain valid until the secret is rotated; rotate the secret passed to `experimental_toolApprovalSecret` to invalidate them, after which approval requests still awaiting a user decision must be re-requested.

— Names config key and exact migration step to rotate secret `ai@7.0.178`

THINNER COVERAGE BELOW

03 **Telemetry and lifecycle callbacks for experimental_extract**

Not checked yet IMPROVED

33

`experimental_extract` gains telemetry, runtime context, and lifecycle callbacks for observability and control over extraction calls.

— Names the API but no concrete callback or config names `ai@7.0.178`

└──▷ **BREAKING ON UPGRADE**

! With `experimental_toolApprovalSecret` configured, signatures previously issued for automatically denied tool calls remain valid; to invalidate them, rotate the secret passed to `experimental_toolApprovalSecret` — approval requests still awaiting a user decision must then be re-requested.

WAS THIS USEFUL? 👍 👎

cotal

Feedback ⓘ v0.172.0 NOTES ↗ RANK

Cotal is the open standard for agent coordination

> Cotal's v0.172.0 release adds remote approval of Claude Code tool permissions and tightens diagnostics and error messaging across extension management, restore flows, Linux self-update, and OpenCode integration.

└──▷ **WHAT SHIPPED · 6 FEATURES** most completely described first ? what's the number?

01 **Extension package diagnostics in `cotal ext list`** Not checked yet

IMPROVED 75

`cotal ext list` now reports missing or corrupt extension package metadata without loading or repairing packages; a corrupt or unreadable manifest can be fixed separately with `cotal ext seed --repair`.

Check extension package health without triggering a repair or load cycle.

```
$ cotal ext list
```

Repair a corrupt or unreadable extension manifest when version or update checks fail.

```
$ cotal ext seed --repair
```

— Exact commands given for both diagnosis and repair v0.172.0

02 **Remote approval for Claude Code tool permissions** Not checked yet

NEW 70

A Claude Code seat launched by Cotal on a signed-in space now lets its owner approve or deny Claude's tool permissions from another session while Claude's own dialog stays open, using `MeshAgent.prompts` to expose the agent's prompt desk to a connector.

— Named API surface and mechanism, but no step-by-step usage shown v0.172.0

03 **Cleanup command surfaced after restore** Not checked yet IMPROVED 65

After a restore completes via `cotal up --foreground --restore`, the cleanup command is now printed directly in terminal output so it can be copied and run immediately.

After a restore completes, see the cleanup command and copy it directly from the terminal output.

```
$ cotal up --foreground --restore
```

— Exact command given, but mechanism of cleanup itself unexplained v0.172.0

04 **Linux self-update runtime and recovery handling** Not checked yet

IMPROVED 65

Linux updates now use the validated per-user runtime when `XDG_RUNTIME_DIR` is unset, and report a partial self-update with a recovery command if the systemd user manager cannot activate the daemon.

— Names env var and failure mode but not the recovery command itself v0.172.0

05 **Clearer error for unsupported OpenCode versions** Not checked yet

IMPROVED 60

Spawning OpenCode with an unsupported `opencode` binary on PATH now prints a single line naming the file, its version, the supported version lines, and how to install a supported OpenCode.

— Describes message content but gives no runnable example v0.172.0

THINNER COVERAGE BELOW

06 **Scoped connection instructions for Cotal tools** Not checked yet

IMPROVED 35

Connection instructions now recommend only the Cotal tools that the connection can call, rather than listing the full set.

— Thin prose description with no named surfaces or example v0.172.0

WAS THIS USEFUL?

Cloudflare Agents

Feedback ⓘ `@cloudflare/think@0.24.0`

NOTES ↗ RANK

The Cloudflare Agents SDK builds stateful agents on Workers and Durable Objects, with persistent state, scheduling, WebSockets, and MCP server support.

> Cloudflare Agents shipped a breaking rename of its JavaScript backend to CodemodeBackend and codemode: host modules, alongside a major Pi 1.6 harness upgrade adding tool parking/signaling, session seeding from history, a customizable onFetchError hook, manual MCP reconnection after auth failures, and new FLUX.2 image model support.

└──▷ **WHAT SHIPPED · 8 FEATURES** most completely described first ? what's the number?

01 CodemodeBackend rename and codemode: module prefix

Not checked yet

95

BREAKING

The JavaScript backend class is renamed from `WorkerJavaScriptBackend` (imported from `@cloudflare/computer/backends/worker-javascript`) to `CodemodeBackend` (imported from `@cloudflare/computer/backends/codemode`), and host module names change from the `ws:` prefix to `codemode:` — e.g. `ws:tools` becomes `codemode:tools` and `ws:git` becomes `codemode:git` — in backend `modules` config and any prompts or code referencing them. The backend's default id also changes to `codemode`, so code referencing an id like `js` must be updated or pass that `id` explicitly to `CodemodeBackend`. Think now requires `agents` >=0.34.0 and `@cloudflare/computer` ^0.10.0.

Migrate an existing Think workspace to Computer 0.10 by swapping the backend import and renaming host modules.

typescript

```typescript
import { CodemodeBackend } from
"@cloudflare/computer/backends/codemode";

const backend = new CodemodeBackend({
  modules: {
    "codemode:tools": myToolsModule,
    "codemode:git":   myGitModule,
  },
});
```

— Full rename mapping and migration code given `@cloudflare/think@0.24.0`

02 **onFetchError hook for custom error responses** Not checked yet

90

BREAKING

Failed requests no longer send error details to the client by default — HTTP requests now return a generic `Internal Server Error` body and WebSocket upgrades close with code 1011. Implementing onFetchError(request, error) on a Lifecycle host or `Agent` lets you choose the response when a request or WebSocket upgrade fails, e.g. returning detailed errors in development while sending a generic message to production clients.

Return a custom error response in development while sending a generic 500 to production clients when a request handler throws.

typescript

```
import { Agent } from 'agents';

export class MyAgent extends Agent<Env> {
  onFetchError(request: Request, error: unknown): Response {
    if (new URL(request.url).hostname === 'localhost') {
      return new Response(String(error), { status: 500 });
    }
    return new Response('Internal Server Error', { status: 500
});
  }
}
```

— Named hook, before/after behavior, and runnable example `agents@0.34.0`

03 **Pi 1.6 upgrade with tool parking and signaling** Not checked yet

85

BREAKING

harness.signal() delivers signals to parked Pi tools, letting an agent sleep until a signal arrives or a deadline passes instead of staying awake. PiSession.submit() and continue() gain `notify` and `onAbort: "keep"` options, letting a parked tool wake when another session's input settles (sub-agent coordination) and ensuring queued input isn't dropped when a session aborts mid-answer. Underlying pi-durable 1.6.0 changes rename `RunEndTask` to `HookTask`, let a tool's execute() return a park, make `until` optional on sleeping tasks, and source code-mode namespaces from `Extension.namespace`. Pi is updated to `~1.6.0` via `agents/harness/pi` and `agents/models/pi-ai`, and Pi 1.5 is no longer supported. Signals must now be delivered via harness.signal() rather than (await harness.pi()).signal(), and harness.root() must be called before reaching the root through harness.pi().

— Detailed mechanism and renamed APIs, no runnable example `agents@0.34.0`

04 **MCP re-authorization errors and manual reconnect** Not checked yet

85

BREAKING

Catching `MCPAuthorizationRequiredError` from callTool(), readResource(), or getPrompt() now surfaces a re-authorization URL so an app can prompt the user to sign in again. A failed MCP server is no longer retried automatically on each wake — addMcpServer() on a failed server throws its recorded error, and callers must explicitly call reconnectMcpServer(id) to retry.

— Named error type and functions, but no example `agents@0.34.0`

05 **Session seeding from existing history** Not checked yet NEW

75

A Pi session can now be seeded from an existing conversation history by passing Pi's `init` to harness.root(), sessions.create(), or sessions.fork(), letting an agent resume where a prior session left off.

Seed a new Pi session from an existing conversation history so an agent resumes where a prior session left off.

typescript

```typescript
import { PiHarness } from 'agents/harness/pi';

const session = await harness.sessions.create({
  agent: myAgent,
  init: { history: existingHistory }
});
```

— Named APIs with a runnable code example `agents@0.34.0`

06 **FLUX.2 image model support** Not checked yet NEW

75

Run FLUX.2 image models — `@cf/black-forest-labs/flux-2-klein-9b`, `flux-2-klein-4b`, and `flux-2-dev` — via ai.image() and pi-ai's ai.images(), with multipart input and up to four reference images.

— Named models and functions, no example `agents@0.34.0`

07 **workerIdScope prevents duplicate Dynamic Workers** Not checked yet

NEW

65

The new `workerIdScope` option on `ExtensionManager` controls extension Dynamic Worker identity by scoping the worker-id hash to include agent and context labels, stopping extensions that use a host binding from creating duplicate Dynamic Workers on restart.

— Named option and mechanism but no code example `@cloudflare/think@0.24.0`

08 **CDP protocol exports for custom browser sources** Not checked yet

NEW

60

`agents/browser` now exports `normalizeCdpSpec` and `RawCdpProtocol`, letting a custom `BrowserSource` build its spec() from protocol JSON it already has.

— Named exports but minimal context, no example `agents@0.34.0`

└──▷ BREAKING ON UPGRADE

! The JavaScript backend class is renamed from `WorkerJavaScriptBackend` (imported from `@cloudflare/computer/backends/worker-javascript`) to `CodemodeBackend` (imported from `@cloudflare/computer/backends/codemode`); existing imports will break.

! Host module names change from the `ws:` prefix to `codemode:` (e.g. `ws:tools` → `codemode:tools`, `ws:git` → `codemode:git`); any backend `modules` config or code/prompts that reference `ws:*` will break.

! The backend's default id changes to `codemode`; anything that references an id such as `js` must be updated or must pass that `id` explicitly to `CodemodeBackend`.

! Think now requires `agents` >=0.34.0 and `@cloudflare/computer` ^0.10.0; older versions of either dependency are no longer supported.

! Pi 1.5 is no longer supported; once a conversation runs on Pi 1.6, Pi 1.5 cannot run its tool calls or model requests, so a deployment cannot roll back past this release.

! In pi-durable 1.6.0, `RunEndTask` is renamed to `HookTask`; a tool's execute() may now return a park; `until` is optional on sleeping tasks; and code mode namespaces come from `Extension.namespace`.

! Failed requests no longer send error details to the client — HTTP requests get a generic `Internal Server Error` body and WebSocket upgrades close with 1011. Implement `onFetchError` to customise the response.

! A failed MCP server is no longer retried on each wake; addMcpServer() on a failed server throws its recorded error. Callers must now explicitly call reconnectMcpServer(id) to retry.

! Deliver signals with harness.signal() rather than (await harness.pi()).signal(), which does not schedule the harness's wake first.

! Call harness.root() before reaching the root through harness.pi(); until the root is first used, sessions.list() leaves it out and pi.conversation(ROOT_CONVERSATION_ID) returns `undefined`.

WAS THIS USEFUL? 👍 👎

AWS Strands Decider

Feedback ⓘ v0.4.0 NOTES ↗ RANK

Strands Decider is a small, fast decision model, or system one model, for agentic workflows. Pick between options or rate on a scale faster than an LLM, with a calibrated confidence on every decision.

> Strands Decider v0.4.0 adds serverless deployment with image support on AWS Lambda, a torch-free core library, a transformers backend for decision checkpoints, SageMaker DLC deployment, and a concurrency cap for the serve command — while also bumping the minimum Python version to 3.11.

└──▷ **WHAT SHIPPED · 8 FEATURES** most completely described first ? what's the number?

01 **Concurrency cap for serve command** Not checked yet NEW

75

Cap concurrent requests to the `serve` command with the `--max-concurrency` flag to control load under high traffic.

— Names exact flag and command, no numeric default given v0.4.0

THINNER COVERAGE BELOW

02 **Lambda deployment with image support** Not checked yet NEW

52

Deploy Strands Decider on AWS Lambda using the base model's vision tower, enabling image inputs in serverless deployments.

— Names mechanism (vision tower) but no deployment command shown v0.4.0

03 **Torch-free core library** Not checked yet NEW

50

A new torch-free core library defines the decision interface without a PyTorch dependency, reducing installation footprint.

— Explains purpose and benefit, no install instructions v0.4.0

04 **Lambda target for benchmark jobs** Not checked yet NEW

50

Benchmark jobs now support Lambda as a target via the benchmark skill's EC2 runner.

— Names benchmark skill and EC2 runner, no command given v0.4.0

05 **Python 3.11 required** Not checked yet BREAKING

45

Python 3.11 is now required; environments running an earlier Python version will break on upgrade.

– Clear migration note with exact version requirement v0.4.0

06 **Transformers backend for decision checkpoints** Not checked yet

NEW

40

A new transformers backend handles decision checkpoints, with image support included.

– Names backend but lacks usage detail v0.4.0

07 **Opt-in token readout with v2 training recipes** Not checked yet

NEW

40

An opt-in token readout is now available alongside the v2 training recipes and export.

– Mentions feature and recipe version, no flag named v0.4.0

08 **SageMaker DLC deployment** Not checked yet NEW

35

Deploy Strands Decider on Amazon SageMaker using the `strands-decider` DLC.

– Names the DLC but gives no setup steps v0.4.0

└──▷ **BREAKING ON UPGRADE**

! Python 3.11 is now required; environments running an earlier Python version will break on upgrade.

WAS THIS USEFUL? 👍 👎

CopilotKit

Feedback ⓘ v1.90.0 NOTES ↗ RANK

CopilotKit connects AI agents to application interfaces with shared state, tools, and generative UI components.

> CopilotKit's single release in this window adds client-version reporting to the Intelligence gateway, enabling the gateway to tailor behavior per client version.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **Client version sent to Intelligence gateway** Not checked yet NEW

30

CopilotKit now sends the client version to the Intelligence gateway so the gateway can tailor behavior per client version.

— States the mechanism but no config key, flag, or scope detail. v1.90.0

WAS THIS USEFUL? 👍 👎

CopilotKit OpenBot

Feedback ⓘ v1.1.2 NOTES ↗ RANK

OpenBot runs self-hosted AI coworkers that each get their own browser, files, and gated tools, with every action decided and recorded.

> OpenBot ships a Kubernetes agent Helm chart for supervised, signed remote management of self-hosted deployments, while flipping the default for task-thread delegation and tightening several platform details.

└──▷ **WHAT SHIPPED · 6 FEATURES** most completely described first ? what's the number?

01 Kubernetes auto-update agent chart

Not checked yet NEW

95

Install the `openbot-agent` Helm chart from `oci://ghcr.io/copilotkit/charts/openbot-agent` to get automated, supervised Kubernetes upgrades with backup gating and a weekly maintenance window. The agent accepts five signed remote commands — restart, rollback, pause, resume, and support-bundle — and can be paused without uninstalling via `helm upgrade openbot-agent ... --set paused=true`.

Temporarily suspend automated upgrades without removing the agent, for example before a planned manual change.

```
$ helm upgrade openbot-agent
oci://ghcr.io/copilotkit/charts/openbot-agent --set paused=true
```

— Names chart, OCI path, five commands, and a runnable pause command v1.1.2

02 Task threads off by default

Not checked yet BREAKING

85

The `OPENBOT_TASK_THREADS` server setting now controls whether Bots hand work to task threads, and it is off by default in this release, meaning Bots that previously delegated work will instead do it in the conversation. Set `OPENBOT_TASK_THREADS=on` (or `true` or `1`) to restore the previous v1.1.1 behavior.

Re-enable task threads on a server where Bots previously handed work to agents, restoring v1.1.1 behavior.

```env
OPENBOT_TASK_THREADS=on
```

— Exact env var and values given to restore prior behavior v1.1.2

03 **Stricter release validation for kubernetes apply** Not checked yet

IMPROVED 65

`bun scripts/update.ts kubernetes apply` now accepts a published release by checking version, distribution, repository, channel, and image digest rather than commit identity.

— Names the exact command and the five validation criteria v1.1.2

04 **Configurable engine disk location on Windows** Not checked yet

IMPROVED 60

On Windows, OpenBot now places the engine's disk in the install folder (e.g. `F:`) rather than always on `C:`, and offers a `Move OpenBot's storage to F:` action when `C:` is short on space.

— Names the UI action but not a config key or flag v1.1.2

THINNER COVERAGE BELOW

05 **Approved folders persist through server restart** Not checked yet

IMPROVED 55

OpenBot Desktop now retains approved folders from `Settings > Computer` for up to two minutes during a server restart, handing them back when the server recovers.

— Gives UI path and exact retention window v1.1.2

06 **Duplicate delivery method validation for notification routes**

Not checked yet NEW 45

The API now rejects notification routes that list the same delivery method more than once, returning `400 VALIDATION_ERROR`.

— Names error code but not the specific endpoint or route schema v1.1.2

└──▷ **BREAKING ON UPGRADE**

! Task threads are now off by default: Bots that previously handed work to task threads will instead do the work in the conversation. Set `OPENBOT_TASK_THREADS=on` (or `true` or `1`) on the server to restore the previous behavior.

WAS THIS USEFUL? 👍 👎

trigger.dev

Feedback ⓘ v4.8.8 NOTES ↗ RANK

Trigger.dev – build and deploy durable AI agents and workflows

> trigger.dev's v4.8.8 release substantially expands its remote MCP server — adding bulk action management with approval gating, richer run filtering, tag/log/metadata tooling, and new OAuth scopes — while tightening root API key display and reclassifying OIDC throttling errors.

└──▷ **WHAT SHIPPED · 6 FEATURES** most completely described first ? what's the number?

01 **OIDC token throttling now throws RateLimitError** Not checked yet

85

BREAKING

auth.getOIDCToken() now throws `RateLimitError` with code `RATE_LIMIT_EXCEEDED` — instead of `OIDCTokenError` — when OIDC token requests remain throttled after retries, covering both per-run and shared API limits. If you catch these errors with `instanceof OIDCTokenError`, check for `RateLimitError` instead.

– Gives exact error class and migration action. v4.8.8

02 **MCP list_runs filtering and queue precedence** Not checked yet

IMPROVED

70

The MCP `list_runs` tool now filters runs by queue, batch, schedule, and bulk action, and clarifies that `period` takes precedence over `from` and `to`.

– Names exact filters and parameter precedence. v4.8.8

03 **Bulk actions for cancelling and replaying runs** Not checked yet

NEW

60

The remote MCP server can now list, read, create, and abort bulk actions that cancel or replay many runs at once, with every bulk action requiring your approval before it executes.

– Describes approval mechanism but no exact command syntax. v4.8.8

THINNER COVERAGE BELOW

04 **New OAuth scopes for MCP server access** Not checked yet NEW

55

The remote MCP server now offers new OAuth scopes for errors, queues, tasks, batches, waitpoints, environments, deployments, branches, environment variable names, and project settings; existing connections keep their current access.

— Names scope categories without usage instructions. v4.8.8

05 **MCP server run management: tags, logs, metadata, streaming**

Not checked yet NEW 50

The remote MCP server can now list tags used on runs, search and read logs, add tags, update metadata, reset idempotency keys, reschedule delayed runs, and read what a run streamed; run details now include metadata values.

— Lists capabilities but no exact tool names or commands. v4.8.8

06 **Root API keys masked in dashboard** Not checked yet BREAKING 50

The dashboard now shows only the last four characters of root API keys for all projects; existing root keys still work but can no longer be copied from the dashboard. Create a separate API key for each new integration.

— States mitigation but no config or API specifics. v4.8.8

└──▷ **BREAKING ON UPGRADE**

! If you catch auth.getOIDCToken() errors with `instanceof OIDCTokenError` to handle throttling, check for `RateLimitError` instead — throttled requests now throw `RateLimitError` with code `RATE_LIMIT_EXCEEDED`.

! Root API keys can no longer be copied from the dashboard; only the last four characters are shown. Create a separate API key for each new integration.

WAS THIS USEFUL? 👍 👎

openclacky

Feedback ⓘ v1.5.25 NOTES ↗ RANK

openclacky is a token-efficient open-source AI agent framework.

> OpenClacky v1.5.25 adds per-agent tool restrictions, a built-in knowledge base extension, two new hosted models, and a round of extension/skills ecosystem and settings improvements.

└──▷ **WHAT SHIPPED · 6 FEATURES** most completely described first ? what's the number?

01 **Per-agent tool restrictions via `tools_allow`** Not checked yet

NEW 70

Agents can be capped to a specific set of tools using a `tools_allow` list, which restricts which tools are available in a session.

– Names exact config key but no further mechanism or scope detail v1.5.25

THINNER COVERAGE BELOW

02 **Extensions and skills ecosystem improvements** Not checked yet

NEW 55

Extension-contributed tools can now attach to other agents, skills can be published and updated directly from the skills page, and locally built extensions can be browsed and iterated on in a personal creations tab.

– Names three UI surfaces grouped from thin bullets, no mechanism detail v1.5.25

03 **Built-in knowledge base extension** Not checked yet NEW

45

A built-in knowledge base extension answers questions from your own markdown notes.

– Explains behaviour but no setup steps or config named v1.5.25

04 **Channel instance renaming and session titles** Not checked yet

IMPROVED 45

Allows renaming a channel instance from its row menu and titles channel sessions after the instance name.

– Describes UI behaviour with clear entry point, no further mechanism v1.5.25

05 **GPT-6.2 Astra and Sol hosted models** Not checked yet NEW

40

Adds GPT-6.2 Astra and Sol to the available hosted models.

— Names the models but no usage or selection detail v1.5.25

06 **Default agent setting for new sessions** Not checked yet NEW

35

Users can pick a default agent for new sessions in Settings.

— Clear UI path but minimal explanation v1.5.25

WAS THIS USEFUL? 👍 👎

lexmount Moli

Feedback ⓘ v1.1.27 NOTES ↗ RANK

Moli is a best headless browser for AI agents. Lite, Fast, High-Compatibility. Built in Rust

> Moli scoped web security controls to individual browser instances, made DevTools' external renderer automation optional, and added runtime diagnostics along with new native CSSOM and event-handling interfaces.

└──▷ **WHAT SHIPPED · 5 FEATURES** most completely described first ? what's the number?

01 **Instance-scoped web security flag** Not checked yet IMPROVED

70

The `--disable-web-security` flag is now scoped per browser instance rather than applying globally, letting callers disable web security for a single instance without affecting others.

— Names exact flag and describes scope change precisely. v1.1.27

THINNER COVERAGE BELOW

02 **New PageRevealEvent and PageSwapEvent interfaces** Not checked yet

NEW

50

Moli adds the native `PageRevealEvent` constructor and `PageSwapEvent` interface for use in event handling.

— Names both new interfaces, no usage example given. v1.1.27

03 **Optional external renderer automation in DevTools** Not checked yet

IMPROVED

45

External renderer automation in DevTools is now optional, allowing DevTools to run without requiring a connected renderer.

— Explains behavior change but no exact config surface given. v1.1.27

04 **Container condition snapshots via CSSOM** Not checked yet NEW

40

Native container condition snapshots are now exposed through CSSOM, allowing inspection of container condition state.

— Names CSSOM surface but no concrete API detail. v1.1.27

05 **Runtime exception and script load diagnostics** Not checked yet

NEW

30

Moli now publishes exception and script load diagnostics at runtime, surfacing errors that were previously silent.

— States what's surfaced but no mechanism or access method. v1.1.27

WAS THIS USEFUL? 👍 👎

Nous Research Hermes Feedback ⓘ CVE-2026-130164 CVE ↗

Hermes is a terminal-native, self-improving AI agent for autonomous coding and tasks, with persistent memory, agent-created skills, and a multi-platform messaging gateway.

`CVE-2026-130164`: Bot Desktop screen env file from the terminal backend sets arbitrary environment variables for agent-machine processes (CVSS 7.8, affects `>= 0.21.6, < 0.23.2`) Not checked yet

WAS THIS USEFUL? 👍 👎

Mastra

Feedback ⓘ announced · Nov 19, 2026 BLOG ↗ RANK

Mastra is a TypeScript framework for AI applications and agents, providing workflows, agent memory, RAG, tool calling, and evals.

> Mastra shipped a hosted model gateway unifying 500+ models behind a single API key with spend controls and failover, plus observability alerts on agent metrics that notify Slack, PagerDuty and other channels.

└──▷ WHAT SHIPPED · 2 FEATURES most completely described first ? what's the number?

01 **Mastra Gateway for unified model access** Not checked yet

NEW INTEGRATION NEW 95

Setting `MASTRA_GATEWAY_API_KEY` in a project's `.env` file authenticates all model calls through the Mastra Gateway with a single key, and prefixing any model ID with `mastra/` (e.g. `mastra/anthropic/claude-fable-5.2`) in the `model` field of an Agent routes inference through it. The gateway provides access to 500+ language, voice, video, and decision models — including frontier, open-source, and free models — across all major providers, with built-in automatic failover that retries failed or underperforming requests against a secondary provider with no extra configuration. Per-key spend limits can be set at hourly, daily, weekly, monthly, or lifetime granularity, allowed models can be restricted per key, and regional data residency can be enforced per key with a US or EU limit or by routing directly to the dedicated EU endpoint.

Route an agent's inference through the Mastra Gateway so all provider calls go through a single key and bill.

env

```
MASTRA_GATEWAY_API_KEY=your-key
```

Point a Mastra Agent at a specific provider model via the gateway prefix instead of configuring per-provider credentials.

typescript

```
import { Agent } from "@mastra/core/agent";

export const agent = new Agent({
  id: "assistant",
  name: "Assistant",
  instructions: "You are a helpful assistant.",
  model: "mastra/anthropic/claude-fable-5.2"
});
```

— Names env var, model prefix, and runnable code example launch-20261119-373e7913

02 **Observability alerts on agent metrics** Not checked yet

NEW INTEGRATION NEW 75

Mastra Platform supports a new Observability alert type that triggers on agent metrics — error rate, latency, time to first token, total/input/output/reasoning/cached token counts, spend, and eval scores — and sends notifications to Slack, Discord, Microsoft Teams, email, webhooks, PagerDuty, or incident.io. Alerts can be scoped across all matching traffic or broken out per agent, tool, model, or user, previewed with a Backtest chart against recent traffic before creation, started from a built-in template (e.g. Tool failures) or built from scratch with Start blank, and are evaluated every five minutes; alerts can also be created directly from any chart on the Metrics dashboard via its bell icon.

Walk through the Mastra Platform UI to create an observability alert that fires when tool failure rate exceeds your threshold and notifies your Slack channel.

📍 1. Open the Mastra Platform dashboard and choose Alerts from the sidebar menu. 2. Select New alert, then Observability alert. 3. Choose the Tool failures template. 4. Select your project, set a threshold, and choose Slack as the notification destination. 5. Use the Backtest chart to verify the rule against recent traffic, then select Create alert.

— Clear UI walkthrough but no API or config surface launch-20261119-331984c7

NEW INTEGRATION Mastra Connect additional context connectors

announced · Nov 18, 2026 BLOG ↗

Mastra Factory adds app previews, cloud workspaces with Docker/CPU/memory customization, Slack-native workflow control, Signals observability primitive, and three-mode model flexibility.

Bootstrap a new factory with `npx create-factory <factory-name>`, then start it with `npm run dev` and open it at `http://localhost:4111`. Not checked yet

Customizable workflow stages — Triage, Plan, Review, and Done — let teams attach their own skills, tooling, guardrails, and human-approval gates at each step. Not checked yet

Cloud workspaces now support customizable Docker images, memory and CPU allocation, and persistent sandbox environments with stable preview URLs that survive across sessions. Not checked yet

App previews give every work item a live ephemeral environment that persists across sessions, so teams can interact with the real running app before merging. Not checked yet

> Slack-native mode lets teams file work, answer agent questions, and approve plans from a thread; it also supports ambient mode where the factory agent decides when to respond, Signals for observability-driven steering, and opt-in generative UI with rich Slack-native charts. Not checked yet
>
> +2 more

WAS THIS USEFUL? 👍 👎

PydanticAI

Feedback ⓘ CVE-2026-133161 CVE ↗ STABLE

PydanticAI is a framework that builds type-safe Python agents with dependency injection, model integrations, tools, and structured outputs.

`CVE-2026-133161`: Harness `ToolOutputLimits` ANSI stripping can take quadratic time, blocking the event loop (CVSS 7.5, affects `>= 0.4.0, <= 0.69.0`)
Not checked yet

▾ **1 more release in this issue** · 2026-11-04 → 2026-11-18

> CVE-2026-133505 CVE ↗ STABLE
>
> `CVE-2026-133505`: Tool result data shaped like a file is rebuilt as a file download or provider file reference when message history is reloaded (CVSS 5.9, affects `>= 1.50.0, <= 1.107.9`) Not checked yet

WAS THIS USEFUL? 👍 👎

AWS Strands Agents

Feedback ⓘ announced · Nov 18, 2026

BLOG ↗ RANK

Strands Agents is AWS's open-source SDK for production agents in Python and TypeScript, with a model-driven loop, tool integration, and deployment targets.

> Strands Agents shipped a new `strands-mongodb` package that lets agents persist sessions and store searchable long-term memory in MongoDB, with configurable scoping, fact extraction, expiry, and search modes.

└──▷ **WHAT SHIPPED · 2 FEATURES** most completely described first ? what's the number?

01 **MongoDB long-term memory store for agents** Not checked yet NEW

90

Connect MongoDB as a long-term memory backend using `MongoDBMemoryStore` with `MemoryManager` to store and recall facts across separate conversations. Memories can be scoped to a user or project via the `scope` parameter to keep different agents' knowledge isolated, automatically extracted from conversations by setting `extraction=True`, expired after a limited time via `ttl_seconds`, and searched using vector, full-text, or hybrid modes via the `search_type` key in the MongoDB config dict.

Give an agent persistent sessions and cross-conversation memory backed by MongoDB Atlas, with fact extraction enabled and vector search for semantic recall.

```python
from strands import Agent
from strands.memory import MemoryManager
from strands.session import SnapshotSessionManager
from strands_mongodb import MongoDBMemoryStore, MongoDBStorage
import os

config = {
    "connection_string": os.environ["MONGODB_URI"],
    "database": "agents",
    "search_type": "vector",
}

storage = MongoDBStorage(config, collection="sessions")
memories = MongoDBMemoryStore(
    name="project_notes",
    config=config,
    scope="project-123",
    writable=True,
    extraction=True,
)

async with Agent(
    storage=storage,
    session_manager=SnapshotSessionManager(
        session_id="conversation-001",
        save_latest_on="message",
    ),
    memory_manager=MemoryManager(stores=[memories]),
) as agent:
    response = await agent.invoke_async("Remember: integration
tests use a separate database.")
    print(response)
```

— Names all config params with mechanism and a working code sample

launch-20261118-76361ed0

02 **MongoDB session persistence for agents** Not checked yet NEW

75

Connect MongoDB as a session backend using `MongoDBStorage` together with `SnapshotSessionManager` to persist and resume agent conversations across restarts. How often state is saved is controlled via the `save_latest_on` parameter on `SnapshotSessionManager` (e.g. `'message'` saves after every message).

Give an agent persistent sessions and cross-conversation memory backed by MongoDB Atlas, with fact extraction enabled and vector search for semantic recall.

python

```python
from strands import Agent
from strands.memory import MemoryManager
from strands.session import SnapshotSessionManager
from strands_mongodb import MongoDBMemoryStore, MongoDBStorage
import os

config = {
    "connection_string": os.environ["MONGODB_URI"],
    "database": "agents",
    "search_type": "vector",
}

storage = MongoDBStorage(config, collection="sessions")
memories = MongoDBMemoryStore(
    name="project_notes",
    config=config,
    scope="project-123",
    writable=True,
    extraction=True,
)

async with Agent(
    storage=storage,
    session_manager=SnapshotSessionManager(
        session_id="conversation-001",
        save_latest_on="message",
    ),
    memory_manager=MemoryManager(stores=[memories]),
) as agent:
    response = await agent.invoke_async("Remember: integration
tests use a separate database.")
    print(response)
```

— Names exact classes and config key with a runnable example `launch-20261118-76361ed0`

WAS THIS USEFUL? 👍 👎

AutoGPT Feedback ⓘ CVE-2026-133460 CVE ↗ STABLE RANK

AutoGPT automates multi-step workflows with continuous AI agents that users can build and deploy.

`CVE-2026-133460`: Public webhook endpoints buffer large unauthenticated bodies before verification (affects `< 0.11.0`) Not checked yet

▾ **14 more releases in this issue** · 2026-11-17

> CVE-2026-133459 CVE ↗ STABLE
>
> `CVE-2026-133459`: Office preview expands an unbounded embedded ZIP thumbnail (affects `>= 0.6.63, < 0.11.0`) Not checked yet

> CVE-2026-133456 CVE ↗ STABLE
>
> `CVE-2026-133456`: Anyone with an Exa Search Monitor trigger's URL can start its workflow with fake events, even with a secret set (CVSS 7.5, affects `>= 0.11.1, < 0.11.3`) Not checked yet

> CVE-2026-133454 CVE ↗ STABLE
>
> `CVE-2026-133454`: A workflow that runs itself as a sub-workflow fills every executor worker, blocking other users' runs (CVSS 6.5, affects `< 0.11.2`) Not checked yet

> CVE-2026-133451 CVE ↗ STABLE
>
> `CVE-2026-133451`: AutoPilot never cleans up the media files its block runs write to the executor's disk, so a user can fill it (CVSS 4.3, affects `>= 0.6.46, < 0.11.2`) Not checked yet

> CVE-2026-133450 CVE ↗ STABLE
>
> `CVE-2026-133450`: Any member of a Slack workspace, Telegram group or Teams team can link it to their own AutoGPT account (CVSS 4.6, affects `>= 0.6.67, < 0.11.3`) Not checked yet

> CVE-2026-133449 CVE ↗ STABLE
>
> `CVE-2026-133449`: Credentials followed a redirect off-origin when the Location header hid the new host behind a backslash (CVSS 5.3, affects `>= 0.6.1, < 0.11.0`) Not checked yet

CVE-2026-132271 CVE ↗ STABLE

`CVE-2026-132271`: Template rendering was not bounded in size or time (CVSS 6.5, affects `< 0.11.0`) Not checked yet

CVE-2026-132266 CVE ↗ STABLE

`CVE-2026-132266`: A Slack workspace member can make the Slack bot send its bot token to their server and fetch internal URLs (CVSS 7.7, affects `>= 0.6.67, < 0.11.3`) Not checked yet

CVE-2026-132265 CVE ↗ STABLE

`CVE-2026-132265`: Sandbox file extraction has no cumulative byte or file-count limit (affects `>= 0.6.49, < 0.11.0`) Not checked yet

CVE-2026-132260 CVE ↗ STABLE

`CVE-2026-132260`: AutoGPT Platform E2B code-execution blocks leave sandboxes running on the operator's shared credential far past what the block actually charges for (CVSS 6.5, affects `>= 0.4.2, < 0.11.0`) Not checked yet

CVE-2026-132255 CVE ↗ STABLE

`CVE-2026-132255`: LLM and other dynamically-priced blocks run free when a priced input is left at its schema default (CVSS 6.5, affects `>= 0.3.4, < 0.11.0`) Not checked yet

CVE-2026-132254 CVE ↗ STABLE

`CVE-2026-132254`: A member removed from an organization keeps creating workflows in it with an API key made while a member (CVSS 4.3, affects `>= 0.6.67, < 0.11.3`) Not checked yet

CVE-2026-130171 CVE ↗ STABLE

`CVE-2026-130171`: Marketplace and shared agents can redirect a user's saved SMTP or database login to a server the agent's author controls (CVSS 6.8, affects `>= 0.6.5, < 0.11.0`) Not checked yet

CVE-2026-130170 CVE ↗ STABLE

`CVE-2026-130170`: One user can stall other users' workflow runs by leaving executor workers stuck in a video encode or email send (CVSS 5.3, affects `< 0.11.3`) Not checked yet

WAS THIS USEFUL? 👍 👎

◆ MCP Servers & Tooling

Prefect FastMCP

Feedback ⓘ v4.2.4 NOTES ↗ RANK

FastMCP is a Python framework for building MCP servers and clients, handling transport, schema generation, authentication, and server composition.

> FastMCP v4.2.4 adds a warning for mounts that mix versioned and unversioned components under the same name.

└──▷ WHAT SHIPPED · 1 FEATURE most completely described first ? what's the number?

01 **Warning for mixed versioned mounts** Not checked yet IMPROVED

42

FastMCP now warns when a mount mixes versioned and unversioned components under the same name, surfacing ambiguous server configurations at mount time.

— Describes behavior but no exact API or flag name v4.2.4

WAS THIS USEFUL? 👍 👎

canvas-mcp Feedback ⓘ v1.19.0 NOTES ↗ RANK

Canvas LMS MCP server — up to 102 tools and 8 agent skills for students & educators. Works with Claude, Cursor, Codex, and 40+ agents.

> canvas-mcp locks down local installations to read-only by default, requiring explicit opt-in to restore write capabilities like editing, publishing, grading, and messaging.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 Write-allowlist for local installations Not checked yet BREAKING

86

Local installations of canvas-mcp now default to read-only mode, disabling editing, publishing, grading, and messaging tools unless explicitly allowed. Set `ALLOWED_WRITE_TOOLS` to a comma-separated list of tool names (e.g. `edit_page_content`) in your environment, or enter them under Allowed write tools in Claude Desktop; leaving it unset keeps the installation read-only.

— Names exact env var, UI setting, and example tool v1.19.0

└──▷ **BREAKING ON UPGRADE**

! Local installations now start read-only by default; any workflow that previously performed edits, publishing, grading, or messaging without explicit configuration will stop working until the relevant tools are listed in `ALLOWED_WRITE_TOOLS` or the Allowed write tools setting in Claude Desktop.

WAS THIS USEFUL? 👍 👎

gram

Feedback ⓘ server@2.39.0 NOTES ↗ RANK

Gram provides MCP servers and AI-agent tools from APIs with access controls and observability.

> Gram shipped an interactive Access Requests workflow spanning Slack and the dashboard with new management endpoints, pagination for API key listing, and a new rule-annotation value for ungoverned tools.

└──▷ **WHAT SHIPPED · 3 FEATURES** most completely described first ? what's the number?

01 Access Requests workflow with Slack and dashboard approvals

Not checked yet NEW 91

Access requests can now be approved, denied, or withdrawn via three new endpoints — `listAccessRequests`, `decideAccessRequest`, and `withdrawAccessRequest` — issuing temporary grants of 1, 2, or 4 hours or until end of day. Approvals can be done interactively from Slack DMs with a duration picker and per-card outcome updates (requesters notified by DM when approved, denied, or expired), or from the new Access Requests tab under Roles & Permissions in the dashboard, which also supports assigning a role directly instead of a temporary grant.

Fetch the first page of pending access requests as an org admin to review what needs a decision.

```
$ curl -X POST https://<your-gram-server>/api/listAccessRequests
\
  -H 'Authorization: Bearer <token>' \
  -H 'Content-Type: application/json' \
  -d '{}'
```

As an administrator, approve a pending access request for 4 hours using the new endpoint.

```
$ curl -X POST https://api.example.com/decideAccessRequest \
  -H 'Authorization: Bearer <token>' \
  -d '{"requestId": "<request_id>", "decision": "approve",
"duration": "4h"}'
```

— Names endpoints, durations, Slack and UI surfaces, with runnable example

dashboard@0.153.0 server@2.39.0

02 Pagination and search for keys.list Not checked yet IMPROVED

80

`keys.list` now supports `limit` (1–100), `cursor` (populated from the response's `next_cursor`), and a case-insensitive `search` by key name; callers that omit `limit` still receive all keys in one response, preserving prior behavior.

Walk through all API keys for a large organization without timing out by paging 50 at a time and filtering by name.

```
$ curl -X POST https://<your-gram-server>/api/keys.list \
  -H 'Authorization: Bearer <token>' \
  -H 'Content-Type: application/json' \
  -d '{"limit": 50, "search": "prod", "cursor": "<next_cursor>"}'
```

Fetch the first page of API keys whose name contains `prod`, then retrieve the next page using the returned cursor.

```
$ # Page 1
curl -X POST https://api.example.com/keys.list \
  -H 'Authorization: Bearer <token>' \
  -d '{"limit": 50, "search": "prod"}'

# Page 2 — pass the next_cursor from the previous response
curl -X POST https://api.example.com/keys.list \
  -H 'Authorization: Bearer <token>' \
  -d '{"limit": 50, "search": "prod", "cursor": "<next_cursor>"}'
```

— Exact param names and limits with a runnable paging example dashboard@0.153.0 server@2.39.0

THINNER COVERAGE BELOW

03 **Unannotated rule value for tool governance** Not checked yet NEW

45

A new `unannotated` rule annotation value is available in the access, agents, and Platform MCP APIs to classify and govern tools that carry no annotations.

— Names the APIs and value but no usage example or mechanism dashboard@0.153.0

WAS THIS USEFUL? 👍 👎

projectmem

Feedback ⓘ v0.3.6 NOTES ↗ RANK

Projectmem is an open-source coding agents memory. Records issues, attempts, fixes and decisions, then warns your agent before it repeats an approach that already failed.

> projectmem v0.3.6 adds a pre-edit hook that surfaces a file's history before changes, a batch precheck API, and a consolidated session-start call, alongside a breaking change to issue tracking for parallel agents.

└──▷ **WHAT SHIPPED · 4 FEATURES** most completely described first ? what's the number?

01 Pre-edit hook for Claude Code and Codex CLI

Not checked yet NEW

80

Run `pjm hooks install --agent claude-code` or `pjm hooks install --agent codex` to install a hook that automatically surfaces a file's open issues, failed attempts, and decisions before every edit.

Wire up the pre-edit hook once so your Claude Code agent automatically sees a file's past failures and open issues before touching it.

```
$ pjm hooks install --agent claude-code
```

— Exact install command and agents named, mechanism clear. v0.3.6

02 issue_id now required on attempt and fix recording

Not checked yet BREAKING

70

`record_attempt` and `record_fix` now require an `issue_id` parameter so that parallel agent sessions each track their own issues independently; calls omitting it will break existing integrations.

— Clear before/after and migration impact, no code sample. v0.3.6

03 Batch file precheck via precheck_files()

Not checked yet NEW

65

precheck_files(file_paths) checks multiple files at once for past failures and open issues, replacing one-file-at-a-time lookups — useful for CI or multi-file refactors.

Before starting work on a set of files in CI or a multi-file refactor, batch-check them all for known issues in one call.

```python
precheck_files(["src/auth.py", "src/session.py",
"src/models.py"])
```

— Named function signature with a runnable example. v0.3.6

THINNER COVERAGE BELOW

04 **Consolidated start_session() API** Not checked yet NEW

50

start_session() retrieves rules, memory summary, and the current plan in a single call, rather than requiring separate lookups.

— Names the API but gives no usage example. v0.3.6

└──▷ **BREAKING ON UPGRADE**

! `record_attempt` and `record_fix` now require an `issue_id` argument when multiple issues are open — calls that omit it will break existing integrations.

WAS THIS USEFUL? 👍 👎

Google MCP Toolbox for Databases

Feedback ⓘ

v1.18.0 NOTES ↗ RANK

MCP Toolbox for Databases is an open-source MCP server from Google that gives AI agents secure, pooled access to databases through declared tools.

> MCP Toolbox for Databases added server-side query parameter binding for ClickHouse SQL execution, reducing risk of parameter values leaking into raw SQL strings.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **Server-side parameter binding for ClickHouse execute-sql**

Not checked yet IMPROVED 60

The `tool/clickhouse-execute-sql` tool now binds query parameters on the server side, keeping parameter values out of the raw SQL string.

— Names the exact tool and mechanism but no usage example v1.18.0

WAS THIS USEFUL? 👍 👎

opentakeoff

Feedback ⓘ mcp-v0.9.116 NOTES ↗ RANK

opentakeoff measures rooms, symbols, and materials from PDF construction and flooring plans, designed for AI agents to drive via MCP.

> opentakeoff's latest MCP release adds a vector-linework Doors suite and an acoustical ceilings suite, new revision comparison tools, stitched sheet support, in-browser PDF page editing, a 3D plan viewer, and several smaller takeoff and localization improvements.

└──▷ **WHAT SHIPPED · 10 FEATURES** most completely described first ? what's the number?

01 **Doors suite reads door swings from plan linework** Not checked yet

NEW 70

A new Doors suite reads every door swing from plan vector linework using Rust/WASM, matches doors to the schedule and spec `08 71 00`, and adds them to the estimate.

– Names mechanism (Rust/WASM) and matched spec, but no direct command shown mcp-v0.9.116

02 **Plan sheet revision comparison with compare_sheets**

Not checked yet NEW 70

Adds a `compare_sheets` MCP tool to compare plan sheet revisions, including all vector linework, plus a new visual revision overlay and slip-sheeting capability for comparing plan sheet versions.

– Named MCP tool with described comparison behaviour mcp-v0.9.116

THINNER COVERAGE BELOW

03 **PDF page editing tools in browser** Not checked yet NEW

55

New PDF page tools in the browser let users rotate, reorder, split, and merge pages, with the takeoff following its page.

– Names concrete actions and UI location, no exact path given mcp-v0.9.116

04 **3D view for plan takeoffs** Not checked yet NEW 55

A new 3D view from the Spline viewer renders floors, base, stairs, and ceilings at height, using lazy three.js loading.

– Explains rendering mechanism but no navigation or command given mcp-v0.9.116

05 **Acoustical ceilings suite with ceiling_takeoff tool**

Not checked yet NEW 50

A new acoustical ceilings suite adds `ceiling_takeoff` to the suite tools.

– Names the tool but gives no further mechanism or scope mcp-v0.9.116

06 **Roof pitch unit support in roof_takeoff** Not checked yet IMPROVED

50

Support reading `roof_takeoff` pitch as percent, degrees, or ratio.

– Named field and unit options, but minimal mechanism mcp-v0.9.116

07 **Byte-range loading for large plan sets** Not checked yet IMPROVED

50

Large plan sets now load via byte-range loading with shared worker documents and a text scan that finishes.

– Describes mechanism but no limits or commands given mcp-v0.9.116

08 **Stitched sheet support across MCP sessions and export**

Not checked yet NEW

40

Adds stitched sheet support to MCP sessions, tools, and export (`feat(mcp): stitched sheet support`).

– Mentions surfaces affected but no mechanism detail mcp-v0.9.116

09 **Metric units and named ignored arguments in suite tools**

Not checked yet IMPROVED

25

Suite tools now work in metric, and arguments they ignore are named.

– Vague description with no named arguments or specifics mcp-v0.9.116

10 **Russian localization support** Not checked yet NEW

25

Russian localization is now available.

– Bare statement with no further detail mcp-v0.9.116

WAS THIS USEFUL? 👍 👎

Composio

Feedback ⓘ @composio/cli@0.6.3 NOTES ↗ RANK

Composio connects AI agents to external tools and services with managed authentication and integrations.

> Composio's CLI release adds a `--timeout` flag for login polling, tightens stdin handling in both `execute` and `proxy` (both breaking changes), introduces a `release-manifest.json` upgrade contract, and renders control characters visibly across several commands.

└──▷ **WHAT SHIPPED · 5 FEATURES** most completely described first ? what's the number?

01 Stdin timeout in execute command

Not checked yet BREAKING

95

`composio execute` now caps the stdin wait to 1 second: a pipe that delivers nothing within that window exits with code 1 and the message `No input arrived on stdin within 1 second`, instead of silently running the tool with an empty argument set `{}`. Pass arguments with `-d '{...}'` or `-d -` to preserve the old behaviour.

Execute a tool from an agent harness by passing JSON arguments inline, avoiding the 1-second stdin timeout that would otherwise fail the call.

```
$ composio execute GITHUB_CREATE_ISSUE -d '{"owner": "my-org",
"repo": "my-repo", "title": "Automated issue"}'
```

— Exact error message, exit code, migration path and runnable example @composio/cli@0.6.3

02 Timeout flag for login polling

Not checked yet NEW

75

`composio login --poll` now accepts `--timeout <seconds>` to bound how long it waits for authentication; if the wait expires, the CLI prints the authorization URL and the resume command instead of hanging indefinitely.

Poll for login completion in CI with a fixed deadline, so the job fails fast if auth stalls rather than waiting indefinitely.

```
$ composio login --poll --timeout 60
```

— Exact flag and runnable example given, but no detail on resume command format

@composio/cli@0.6.3

03 release-manifest.json upgrade contract

Not checked yet NEW

65

Release archives now include a `release-manifest.json` listing support files; `composio upgrade` installs exactly what the manifest lists, and a release can declare the minimum CLI version allowed to upgrade to it.

— Mechanism and config file named but no usage example given `@composio/cli@0.6.3`

---

04 **Proxy stdin body requires explicit flag** Not checked yet BREAKING

60

`composio proxy` no longer uses piped stdin as a request body automatically without `-d`; any non-GET request sent without `-d` from a non-terminal stdin now prints a warning on stderr that it is sent without a body.

— Clear before/after but no runnable example of the new flow `@composio/cli@0.6.3`

THINNER COVERAGE BELOW

05 **Visible rendering of control characters in output** Not checked yet

IMPROVED

45

Control characters in event output from `composio listen --stream`, `composio dev listen`, `composio execute`, `composio proxy`, and API error messages are now rendered as visible escapes on the terminal.

— Names affected commands but no mechanism detail or example `@composio/cli@0.6.3`

└──▷ **BREAKING ON UPGRADE**

! In `composio proxy`, piped stdin is no longer used as a request body without `-d`. Any non-GET request sent without `-d` from a non-terminal stdin now prints a warning on stderr that it is sent without a body.

! In `composio execute`, a piped stdin that sends nothing within 1 second now exits with code 1 instead of running the tool with `{}`; pass arguments with `-d '{...}'` or `-d -` to preserve the old behaviour.

WAS THIS USEFUL? 👍 👎

frappe_assistant_core

Feedback ⓘ v3.4.0 NOTES ↗

Frappe_assistant_core is an infrastructure that connects LLMs to ERPNext. Frappe Assistant Core works with the Model Context Protocol (MCP) to expose ERPNext functionality to any compatible Language Model

frappe_assistant_core adds in-chat data migration with undo, per-user tool access, an agent workflow builder, open-page awareness, and answer feedback.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v3.4.0
https://github.com/buildswithpaul/Frappe_Assistant_Core.git
# already have the repo? check out this version:
$ git checkout v3.4.0
```

Restrict individual tools to specific users (in addition to roles) using the new tool user access DocTypes added by `bench migrate`. Not checked yet

Enable plugins to declare tool packs so each tool loads only when needed, with a per-user opt-out switch. Not checked yet

Run data migrations from chat by attaching a file — the assistant resolves open decisions, creates missing masters, and imports data. Not checked yet

Review and undo assistant changes from chat — each answer lists what it changed, and changes that activate together share one go-live card. Not checked yet

Build agent workflows with a decision step in the builder, a setup panel, and an Activate switch. Not checked yet

+3 more

NEW INTEGRATION Tally and Zoho Books data migration presets

└──▷ BREAKING ON UPGRADE

! Run `bench migrate` — it adds the import session, page DocType, and tool user access DocTypes required by this release.

! `assistant_runtime_sdk==1.16.0` is now required; run `bench setup requirements` to install it.

WAS THIS USEFUL? 👍 👎

Supabase MCP Server

Feedback ⓘ mcp-server-supabase-v0.19.0

NOTES ↗ RANK

Supabase MCP Server is a connect Supabase to your AI assistants.

> Supabase MCP Server now lets AI assistant apps create Supabase projects and branches through built-in UI flows, without leaving the app.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **UI flows for project and branch creation** Not checked yet NEW

40

MCP apps now include UI flows for creating a project and creating a branch without leaving the app.

— Describes the two flows but no exact UI path or command mcp-server-supabase-v0.19.0

WAS THIS USEFUL? 👍 👎

reshapr

Feedback ⓘ 1.2.1 NOTES ↗ RANK

Reshapr is the open source runtime that turns your APIs AI-Native, without rebuilding them

> Reshapr added service account management commands to its CLI.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **Service account management CLI commands** Not checked yet NEW

25

Reshapr's CLI now includes new commands for managing service accounts.

— No command names, flags, or usage shown. 1.2.1

WAS THIS USEFUL? 👍 👎

fablecut

Feedback ⓘ v1.20.0 NOTES ↗ RANK

Fablecut is a zero-dependency browser video editor that AI agents can drive — JSON timeline, MCP + REST, live-reloading UI

> Fablecut introduces Workflows, letting agents save, list, and replay named sequences of editing steps across projects via a new tool and REST endpoint.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **Workflows for reusable editing routines** Not checked yet NEW

85

Named sequences of plain-language steps and tool calls can be saved as Workflows — such as a podcast clean-up or weekly reel — stored in `library/workflows/` and replayed by an agent on future projects. Workflows are accessible to agents via `fablecut_workflows` and over REST at `/api/workflows`.

Trigger an agent to list saved workflows and apply the podcast clean-up routine to new footage without hard-coding each step

```
$ fablecut_workflows
```

Query available workflows from an external system or CI pipeline via the REST API

```
$ curl http://localhost:<port>/api/workflows
```

— Names storage path, tool, and endpoint with usage examples. v1.20.0

WAS THIS USEFUL? 👍 👎

◆ Browser & Computer Use

browser-use

Feedback ⓘ 1.0.7 NOTES ↗ RANK

The browser-use Python library automates web browsers for AI agents.

> browser-use 1.0.7 addresses MCP client timeouts on long-running browser automation calls.

└──▷ WHAT SHIPPED · 1 FEATURE most completely described first ? what's the number?

01 **Partial output streaming for browser_exec** Not checked yet

IMPROVED 70

Calling `browser_exec` again with empty code retrieves the remaining output after a long-running cell exceeds 50 seconds, so Cursor and other MCP clients with a 60-second limit receive a result instead of timing out.

– Names exact mechanism and timing but lacks broader usage detail 1.0.7

WAS THIS USEFUL? 👍 👎

◆ Agent Memory

claude-mem

Feedback ⓘ v16.0.2 NOTES ↗ RANK

claude-mem gives coding agents persistent memory by capturing session activity, compressing it with a model, and injecting relevant context into later sessions.

> claude-mem's v16.0.2 release adds a Codex-based fallback for quota-limited observers and makes cache pruning fully traceable from its logs.

└──▷ **WHAT SHIPPED · 2 FEATURES** most completely described first ? what's the number?

01 **Traceable cache prune logging** Not checked yet IMPROVED

57

Cache prune operations now log each removed version along with its reason (orphaned or superseded), plus the `.in_use` records it rejected and why, so any removal can be traced from the log alone.

— Names log fields but no command to trigger or view them v16.0.2

02 **Codex quota fallback for observer** Not checked yet NEW

56

When Claude hits its quota guard, the observer can now fall back to the local Codex CLI signed in with ChatGPT, then return to Claude through the existing recovery path.

— Describes mechanism but no user-facing command or flag given v16.0.2

WAS THIS USEFUL? 👍 👎

◆ Agent Sandboxes

Brig

Feedback ⓘ v0.6.0 NOTES ↗ RANK

Brig runs coding agents in a microVM sandbox, made up of the brig CLI and the brigd session daemon.

> Brig v0.6.0 adds Brigfile-driven target workflows with approval and config digesting, a new `brig report` diagnostic command, enforced microVM isolation via `--isolation microvm`, and a `github-api` preset with GraphQL allow-list enforcement, alongside a breaking change to how presets rewrite git upstreams.

└──▷ **WHAT SHIPPED · 11 FEATURES** most completely described first ? what's the number?

01 **Git rewrites now follow preset's upstream** Not checked yet

95

BREAKING

A preset's git rewrites now follow its upstream: a route that moves `github-token` to another `upstream` rewrites git for that host only, no longer for `github.com`. Policies that relied on `github.com` rewrites (such as a mirror) must write `guest.git.insteadOf` explicitly; see A preset on another upstream in `docs/migration.md`.

— Exact migration path and config key given v0.6.0

02 **Brigfile-driven target workflows** Not checked yet NEW

88

Adds Brigfile target management commands `brig target ls`, `brig config`, and `brig init` to initialise and inspect configuration; `brig plan --target` to show a target's origins, routes, args and any conflicting routes; approval is now required before booting any Brigfile target's sandbox, with approvals recorded and a digest of what a target boots to detect config changes; and new Brigfile-aware session lifecycle commands to stop, remove and list the sessions of Brigfile targets.

— Several named CLI commands with described behavior, no runnable example v0.6.0

03 **Enforced microVM isolation for runs** Not checked yet NEW

85

`brig run` supports `--isolation microvm` to require microVM-level isolation for a run; Brig refuses to start if the host cannot provide that boundary, preventing accidental container-only execution.

Enforce microVM isolation for a run — Brig refuses to start if the host cannot provide that boundary, preventing accidental container-only execution.

```
$ brig run --isolation microvm claude ~/code/demo
```

— Flag, behavior and runnable command given v0.6.0

04 **GitHub API preset with GraphQL enforcement** Not checked yet NEW

85

A new `github-api` preset understands the GitHub CLI `gh` token scheme, maps Enterprise paths onto `api.github.com`, and enforces `request.allow.graphql` to parse and enforce GitHub GraphQL requests against an allow-list, limiting mutations by repository.

— Mechanism and config key named, no example v0.6.0

05 **Diagnostic report command for sandboxes** Not checked yet NEW

80

New `brig report <ref>` command collects and renders a diagnostic report for a sandbox, useful to attach to a bug report or investigate a silent guest.

Collect a diagnostic report for a sandbox to attach to a bug report or investigate a silent guest.

```
$ brig report claude
```

▶ View command output v0.6.0

— Full command with example and use case v0.6.0

06 **Host route mode under brig.internal** Not checked yet NEW

70

A new `host` route mode lets an endpoint route be reached by name under `brig.internal` subdomains, using the `{route.host}` placeholder in policy.

— Named placeholder and domain, no full example v0.6.0

07 **Policy ARN scoping for AWS secrets** Not checked yet NEW

65

`brig secret` now accepts `--policy-arns` to narrow an AWS role session's permissions at run time.

— Flag named but no usage example v0.6.0

08 **Hull sandbox support on Linux** Not checked yet NEW

65

Brig now defaults to hull when it is installed on Linux, enforcing egress policy on hull, routing and publishing on hull, and refusing workloads that hull cannot yet run.

— Behavior described, no flag or command named v0.6.0

THINNER COVERAGE BELOW

09 **Policy document input from stdin** Not checked yet IMPROVED

45

`brig policy create` can now read its policy document from stdin.

— Single command behavior, no example v0.6.0

10 **Shape-based credential masking in redact output** Not checked yet

IMPROVED

45

Redact output now masks credentials by shape, flagging strings that still look random after masking.

— Thin description, no config surface named v0.6.0

11 **Blocked Google identities in GCP routes** Not checked yet NEW

40

Brig now blocks Google identities that the host passes into a GCP route's guest.

— Brief description, no mechanism detail v0.6.0

└──▷ **BREAKING ON UPGRADE**

! A preset's git rewrites now follow its upstream: a route that moves `github-token` to another `upstream` rewrites git for that host only, no longer for `github.com`. Policies that relied on `github.com` rewrites (such as a mirror) must write `guest.git.insteadOf` explicitly. See A preset on another upstream in `docs/migration.md`.

WAS THIS USEFUL?  

◆ Applied AI Agents

Harvey

Feedback ⓘ announced · Nov 20, 2026 BLOG ↗ RANK

Harvey is an AI platform for legal and professional services teams.

> Harvey shipped information governance controls in Command Center — self-serve retention policies, early-access legal holds, and in-app audit log search and export — alongside a major Review Tables upgrade adding agent-driven table creation from templates, multi-source document pulls, sample-document prompt testing, and a new Sync feature for incremental updates.

└──▷ **WHAT SHIPPED · 7 FEATURES** most completely described first ? what's the number?

01 **Sync for Review Tables** Not checked yet NEW 85

Sync for Review Tables lets teams connect a table to a Vault or folder so Harvey flags newly arrived documents and, when the Sync button is clicked, processes only rows for those new documents while verified or locked answers remain unchanged; hovering over the Sync button shows when the table was last synced.

Use Sync to bring a Review Table up to date after new documents arrive in a Vault, without re-running analysis on already-verified rows.

> 📍 In your Review Table, go to the Sync button and click it; Harvey processes only rows for new documents while verified or locked answers remain unchanged. Hover over the Sync button to see when the table was last synced.

— Exact mechanism, UI control, and incremental-processing behaviour described.

launch-20261118-ae55ae76

02 **In-app audit log search and export** Not checked yet IMPROVED 78

Command Center now lets admins view, filter, search, and export audit logs directly in-app — for example filtering by person, action type, and date range before exporting the result as evidence — in addition to the existing API access that pipes audit data into a SIEM such as Splunk.

Audit who changed access or settings on a specific matter by filtering the in-app audit log by person, action, and date, then exporting the result as evidence.

> 📍 In Command Center, open Audit Logs, apply filters for the relevant person, action type, and date range, then export the filtered log.

— Concrete filter/export steps and a named SIEM integration. launch-20261120-1e16d05a

---

03 **Self-serve retention policies in Command Center** Not checked yet

NEW

65

Workspace admins can now set data retention periods directly in Command Center at the workspace, vault, or matter level without submitting a request to Harvey, including overriding the workspace default with a custom period for an individual matter — for example a 6-year retention set on a single client matter to satisfy a contract records clause.

Restrict a single client matter to a six-year retention period to satisfy a contract records clause, while keeping the workspace default shorter.

> 📍 In Command Center, navigate to the matter, open its retention settings, and set a custom retention period of 6 years — overriding the workspace-level policy for that matter only.

— Names the three levels and override behaviour with a worked example.

launch-20261120-1e16d05a

---

04 **Prompt testing on sample documents before full run**

Not checked yet NEW

65

Teams can now test a Review Table prompt on ten documents before running it across the full document set, making it faster to iterate and validate questions prior to a costly full run.

Before running a prompt across thousands of documents, validate it quickly on a small sample to avoid costly full runs with poorly worded questions.

> 📍 In your Review Table setup, select the option to test your prompt on ten documents, review and refine the output, then run on the full document set.

— Specific sample size and a clear test-then-run workflow. launch-20261118-ae55ae76

THINNER COVERAGE BELOW

---

05 **Legal holds for custodians (Early Access)** Not checked yet NEW

52

Legal holds can now be placed on individual custodians directly in Command Center, available in Early Access from December 2026, suspending routine deletion of their Harvey data and targeting preservation to specific people rather than the entire workspace.

— Clear scope and timing, but feature is not yet generally usable.

launch-20261120-1e16d05a

06 **Agent-created Review Tables from firm templates** Not checked yet

NEW

45

Agents in a Space can now create Review Tables directly from a firm's existing templates, including generating several tables at once for a single matter.

— States the capability and scope but gives no concrete steps. launch-20261118-ae55ae76

07 **Multi-source document sourcing for Review Tables** Not checked yet

IMPROVED

45

Review Tables now accept multiple Vaults and external data rooms as sources, letting teams pull together documents from more than one location into a single review.

— Describes the capability but no limits or steps given. launch-20261118-ae55ae76

NEW INTEGRATION Two-way sync with iManage

NEW INTEGRATION DeepJudge integration

└──▷ ALSO FROM THESE RELEASES

announced · Nov 20, 2026 BLOG ↗

Harvey adds two-way iManage sync in early access, letting lawyers save redlines, versions, and drafts back to iManage with version history and multi-document approval.

Save redlines, new versions, and first and final drafts directly from Harvey back to iManage, so iManage remains the system of record without manual download and re-upload. Not checked yet

Save multiple documents in a single approval, with Harvey filing each back to iManage as its next version and updating the synced Vault and iManage together. Not checked yet

Harvey checks check-out status and write access before saving, so a colleague's open document is not overwritten during a multi-document save. Not checked yet

announced · Nov 19, 2026 BLOG ↗

Harvey adds a DeepJudge integration in open early access, letting lawyers draw on firm prior work product inside Harvey drafting and research workflows.

Harvey now adds a DeepJudge integration that lets lawyers draw on their firm's prior work product — precedents, negotiated positions, and matter experience —

> directly while drafting and researching in Harvey. Not checked yet
>
> DeepJudge Collections now sync with Harvey Vault, keeping a firm's curated work product current and available inside Harvey. Not checked yet

---

WAS THIS USEFUL? 👍 👎

◆ Web Search & Scraping for Agents

Exa

Feedback ⓘ announced · Nov 17, 2026 BLOG ↗ RANK

Exa is a web search API built for AI agents, with search, content retrieval and research endpoints.

> Exa opened Websets V2 API access to all users under a new usage-based pricing plan, while improving match quality and speed over V1.

└──▷ WHAT SHIPPED · 3 FEATURES most completely described first ? what's the number?

01 **Legacy plan pricing retained unless manually switched**

Not checked yet BREAKING 70

Existing users on legacy Starter or Pro plans keep their old per-match pricing (fixed 10 credits, $0.045/match on Pro) until they manually switch to usage-based pricing on the billing page; staying on the legacy plan means no access to the lower pricing.

– Exact pricing figures and clear migration action named launch-20261117-c392faa1

THINNER COVERAGE BELOW

02 **Open API access with usage-based pricing** Not checked yet NEW 45

Websets API access is now open for all users on the usage-based plan, requiring no monthly subscription, replacing the previous subscription-gated access model.

– Names the plan change but no pricing mechanics or endpoint launch-20261117-c392faa1

03 **Websets V2 match quality and speed improvements** Not checked yet

IMPROVED 45

Websets V2 now finds 20% more matches on average and runs 4x faster than V1, with the launch blog also citing 2x lower cost per match.

– Concrete performance numbers but no usage mechanism launch-20261117-c392faa1

└──▷ BREAKING ON UPGRADE

! Existing users on legacy Starter or Pro plans retain their old per-match pricing (fixed 10 credits, $0.045/match on Pro) until they manually switch to usage-based pricing on the billing page; staying on the legacy plan means no access to the lower pricing.

WAS THIS USEFUL? 👍 👎

DEPLOY

◆ AI / LLM Gateways

TrueFoundry

Feedback ⓘ announced · Nov 21, 2026 BLOG ↗

TrueFoundry deploys and secures LLMs and AI agents through an enterprise gateway.

TrueFoundry AI Gateway adds Oodle observability export — trace and metrics export via gateway settings, plus Oodle MCP server access through the MCP Gateway.

NEW INTEGRATION Enable trace export from the AI Gateway settings to send every LLM call to Oodle as an OpenTelemetry trace carrying model, latency, cost, guardrail results, and prompt/response data — no application code changes required. Not checked yet

Enable metrics export from the AI Gateway settings to send gateway metrics — including requests, HTTP status codes, cost, latency, time to first token, inter-token latency, budget and rate-limit checks, and MCP request counts — to Oodle alongside traces. Not checked yet

Access the Oodle MCP server through the MCP Gateway so agents can query Oodle logs, metrics, traces, and alerts under the gateway's access controls and per-user OAuth sign-in. Not checked yet

WAS THIS USEFUL? 👍 👎

BerriAI LiteLLM

Feedback ⓘ v1.108.14 NOTES ↗ RANK

LiteLLM is an AI gateway that calls over 100 model APIs in OpenAI format with cost tracking, guardrails, load balancing, and logging.

> LiteLLM expanded the content types accepted by its decisions API.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **input_json content parts on decisions endpoint** Not checked yet

NEW 60

LiteLLM now supports OpenAI `input_json` content parts on the `/v1/decisions` endpoint, expanding the content types that decisions requests can carry.

— Names exact endpoint and content type but no further mechanism. v1.108.14

WAS THIS USEFUL? 👍 👎

◆ ML Platforms & SDKs

Hugging Face

Feedback ⓘ New Home Feed NOTES ↗ RANK

Hugging Face is an open platform for building, sharing, and deploying machine-learning models, datasets, and applications.

> Hugging Face introduced a new Home feed on the logged-in homepage, blending personal activity, followed accounts, and trending Hub content.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **Home feed on logged-in homepage** Not checked yet NEW 55

The logged-in homepage now defaults to a Home feed that combines your own activity, activity from people and organizations you follow, trending model releases, and trending contributors across the Hub, with each release grouped into a single card. The previous Following feed remains accessible via a dropdown at the top of the feed.

View your personalized Home feed combining followed-user activity and trending Hub releases after logging in.

📍 1. Log in to huggingface.co. 2. The homepage now defaults to the Home feed. 3. To switch back to the previous view, open the dropdown at the top of the feed and select Following .

— Explains behavior and exact UI path to switch views. New Home Feed

WAS THIS USEFUL? 👍 👎

AWS Amazon Bedrock AgentCore

Feedback ⓘ

Gateway: Stream responses through response interceptors (2026-11-20) BLOG ↗

RANK

Amazon Bedrock AgentCore provides managed agent runtimes, memory, identity, gateways, observability, and evaluation services.

> Bedrock AgentCore introduced a new Evaluations capability PUBLIC PREVIEW for running managed dataset evaluations against Runtime endpoints and generating scenario datasets automatically from agent traces, alongside streaming support for Gateway response interceptors.

└──▷ **WHAT SHIPPED · 4 FEATURES** most completely described first ? what's the number?

01 **Managed dataset runs in AgentCore Evaluations** Not checked yet

NEW SUPPORTED PLATFORM NEW 65

New AgentCore Evaluations feature PUBLIC PREVIEW runs a published dataset against an AgentCore Runtime endpoint, executing every scenario and evaluating sessions in a single managed batch job.

— Describes workflow but no exact API or command Evaluations: Run a dataset evaluation w…

02 **Dataset generation and mixed scenario datasets in Evaluations**

Not checked yet NEW 65

Evaluations datasets now support both fixed scenarios and dynamic scenarios in a single dataset, removing the previous requirement to declare a schema type. List operations on datasets now accept filters by the fields an example has, letting you query across heterogeneous datasets, and dataset generation can now create evaluation scenarios automatically from an agent's traces.

— Names concrete capabilities but no exact endpoint or flag

Evaluations: Dataset generation and sce…

THINNER COVERAGE BELOW

03 **Streaming support for Gateway response interceptors**

Not checked yet IMPROVED 50

Response interceptors on HTTP and inference targets now process streaming responses in batches of events, delivering model output to clients as it is generated rather than waiting for the full response to finish.

— Explains mechanism but no exact config/flag named

Gateway: Stream responses through respo…

04 **Schema type removed from Evaluations datasets** Not checked yet

50

BREAKING

Datasets no longer declare a schema type field; existing datasets or integrations that set or read a schema type field may need to be updated.

— Clear migration note but lacks field/API detail

Evaluations: Dataset generation and sce…

└──▷ **BREAKING ON UPGRADE**

! Datasets no longer declare a schema type — existing datasets or integrations that set or read a schema type field may need to be updated.

> Runtime: Instances supports Amazon EC2 G7 instances (2026-11-20) BLOG ↗
>
> LATEST STABLE
>
> Amazon Bedrock AgentCore Runtime now supports Amazon EC2 G7 instances, adding GPU compute for agents doing 3D rendering or model inference.

WAS THIS USEFUL? 👍 👎

Anthropic TypeScript SDK Feedback ⓘ sdk-v1.12.0

NOTES ↗ RANK

The Anthropic TypeScript SDK provides typed JavaScript and TypeScript clients for Claude APIs, streaming, and tool use.

> The Anthropic TypeScript SDK adds a new API error code for oversized image inputs.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **image_dimensions_too_large error code** Not checked yet NEW

50

The SDK adds the `image_dimensions_too_large` error code to the API error surface, letting applications catch and handle oversized image inputs explicitly.

— Names the exact error code but no handling example. sdk-v1.12.0

WAS THIS USEFUL? 👍 👎

Anthropic Python SDK

Feedback ⓘ v2.12.0 NOTES ↗ RANK

The Anthropic Python SDK provides typed Python clients for Claude APIs, streaming responses, and tool use.

> The Anthropic Python SDK added a new API error code for oversized image rejections.

└──▷ **WHAT SHIPPED · 1 FEATURE** most completely described first ? what's the number?

01 **New image dimensions error code** Not checked yet NEW 50

Adds the `image_dimensions_too_large` error code to the API, letting callers detect and handle oversized-image rejections explicitly.

— Names exact error code but no handling example or limit given. v2.12.0

WAS THIS USEFUL? 👍 👎

skillhub

Feedback ⓘ v0.2.28 NOTES ↗ RANK

Skillhub is a self-hosted, open-source agent skill registry for enterprises. Publish & version skill packages, govern with RBAC and audit logs, deploy on-premise with Docker or Kubernetes.

> SkillHub v0.2.28 tightens access control with granular API token permission scopes and adds GitHub repository snapshot downloads, a batch review workbench for GitHub-synced skills, opt-in team-only automatic publication, and source filtering in the Web market — alongside two breaking changes around token-based auto-publish and Suite Bundle authorization provenance.

└──▷ **WHAT SHIPPED · 7 FEATURES** most completely described first ? what's the number?

01 Granular API token permission scopes

Not checked yet NEW

88

API tokens now separate skill read, submit, private-publication (`skill:publish-private`), management, and deletion capabilities, with web presets making grants explicit. Token creation, rotation, and renewal cannot exceed the calling token's capabilities or expiry, and stored legacy scopes are preserved and marked as compatible.

Issue a new API token with only read and submit capabilities — without private-publication or deletion — to enforce least-privilege access for a CI integration.

> 📍 Steps: In the token creation UI, select the read and submit capability presets. Do not select `skill:publish-private` , management, or deletion. Save the token; it cannot exceed the calling token's capabilities or expiry.

— Names exact capability flags and constraints, gives setup steps. v0.2.28

02 Batch review workbench for GitHub-synced skills

Not checked yet NEW

75

Batch-select up to 50 skill-version review tasks across pages and approve or reject them with per-item results; permissions, current scan results, and resource state are rechecked during execution and each item commits independently. GitHub-synchronized skill versions are reviewed grouped by immutable acquisition snapshot in a dedicated workbench, with independently paged members, explicit batch selection, and preserved return context.

— Mechanism and limits are detailed but no exact UI path given. v0.2.28

03 Opt-in team-only automatic publication

Not checked yet NEW

68

TEAM namespace owners and administrators can grant an ordinary member permission to publish eligible team-only submissions automatically after a scan returns SAFE; the feature is off by default and does not widen visibility or bypass scanning.

Grant a TEAM member automatic publication rights so their scan-SAFE submissions publish without manual review, while keeping the feature scoped to team-only visibility.

📍 Steps: As a TEAM namespace owner or administrator, open the team settings and locate the Automatic publication grant panel. Select the eligible member and enable the opt-in grant. Confirm that the feature is enabled only after the backend rollout is complete and no mixed-version backends are running.

— Clear grant steps given but no config key named. v0.2.28

---

04 **GitHub repository snapshot downloads** Not checked yet NEW

67

Download GitHub repository snapshots as ZIP archives containing exact-version skill packages and a checksum manifest, with access checked at download time.

Download a point-in-time ZIP snapshot of a GitHub repository's skills — including exact-version packages and a checksum manifest — for offline audit or air-gapped deployment.

📍 Steps: In the Web market, switch to the GitHub source filter and browse to the target repository. Open the snapshot history, select the desired acquisition snapshot, and click Download. Access is verified at download time and the archive contains exact-version skill packages and a checksum manifest.

— Describes format and access check with a usage walkthrough. v0.2.28

---

05 **SUPER_ADMIN auto-publish on upload removed** Not checked yet

65

BREAKING

Uploading a private skill with a `SUPER_ADMIN` token no longer publishes it automatically; integrations relying on that shortcut must explicitly confirm publication with a `skill:publish-private` or compatible legacy `skill:publish` token.

— Names exact token scopes needed but no migration command. v0.2.28

THINNER COVERAGE BELOW

06 **Source filtering in Web market** Not checked yet IMPROVED

52

Filter skill discovery by platform, GitHub, or all sources in the Web market; automatic homepage popular and latest lists now select platform content, while API searches without a source filter and existing CLI searches retain all-source behavior.

– Describes behavior change but no exact UI path or flag. v0.2.28

07 **Suite Bundle operation authorization provenance required**

Not checked yet

52

BREAKING

Pre-upgrade Suite Bundle operations without recorded authorization provenance cannot resume or retry using a worker's or another user's broader permissions; affected operations must be cancelled and recreated as directed by the UI.

– Gives remediation step but lacks API or command specifics. v0.2.28

└──▷ **BREAKING ON UPGRADE**

! Uploading a private skill with a `SUPER_ADMIN` token no longer publishes it automatically; integrations relying on that shortcut must explicitly confirm publication with a `skill:publish-private` or compatible legacy `skill:publish` token.

! Pre-upgrade Suite Bundle operations without recorded authorization provenance cannot resume or retry using a worker's or another user's broader permissions; cancel and recreate affected operations as directed by the UI.

WAS THIS USEFUL?  

siliconscope

Feedback ⓘ v4.9.1 NOTES ↗ LATEST STABLE

Siliconscope sudoless Apple Silicon system monitor (native SwiftUI GUI) with ANE / Media Engine / memory-bandwidth tracking

SiliconScope adds `sscope-cli --chip-report`, a new CLI command that runs five timed load profiles and saves a sanitised diagnostic zip to the Desktop.

└──▷ GET THIS VERSION ▾

Homebrew

```
$ brew install --cask siliconscope
# already installed? upgrade to this version:
$ brew upgrade --cask siliconscope
```

git

```
$ git clone --branch v4.9.1 https://github.com/kennss/SiliconScope.git
# already have the repo? check out this version:
$ git checkout v4.9.1
```

Homebrew lists this version · checked 21 Nov Wrong command?

Run `sscope-cli --chip-report` to measure a Mac's sensors, power, and bandwidth across five load profiles (rest, E-cores only, all cores, GPU only, GPU memory streaming) and save a sanitised diagnostic zip to the Desktop. Not checked yet

UPDATED SUPPORTED PLATFORM M3 Max, M3 Ultra, and A18 memory-traffic display

NEW SUPPORTED PLATFORM M3 Ultra on macOS 27

▾ 1 more release in this issue · 2026-11-20

v4.9.0 NOTES ↗ STABLE

SiliconScope adds dense bar charts for real-time metrics, hideable dashboard cards that pause measurement when hidden, and a BY APP process-table grouping.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v4.9.0
https://github.com/kennss/SiliconScope.git
# already have the repo? check out this version:
$ git checkout v4.9.0
```

Every real-time chart (CPU, GPU, Neural Engine, Media, memory bandwidth, network, disk) is now dense bars, one per sample, with upload/download and disk write/read sharing one panel. Not checked yet

Hide individual dashboard cards via Settings → Dashboard or right-clicking a card's title; hidden cards also pause measurement unless another view still needs the same reading. Not checked yet

The BY APP view folds the process table by app, with the dropdowns' top lists always grouped the same way. Not checked yet

Memory bandwidth idle readings below histogram resolution now display as `< 128 GB/s` instead of a raw number. Not checked yet

---

WAS THIS USEFUL? 👍 👎

Databricks AI

Feedback ⓘ announced · Nov 19, 2026 BLOG ↗

RANK

Databricks AI provides managed model serving, agent development, evaluation, and retrieval over data in the Databricks platform.

> Databricks AI shipped two new Genie-branded agents in this window: ZeroOps, an autonomous pipeline triage and fix agent now in public beta, and Genie Migrate for Serverless, which automates classic-to-serverless workload migration via CLI or in-session slash command.

└──▷ **WHAT SHIPPED · 2 FEATURES** most completely described first ? what's the number?

01 **Genie Migrate for Serverless** Not checked yet NEW 92

Genie Migrate for Serverless automates classic-to-serverless workload migration, available after enabling the beta in workspace preview settings. Run bulk migrations from the command line with the `compute-migrations` command group in Databricks CLI 1.24.0 or above, building serverless working copies of multiple jobs without an open Genie session; or migrate a single job or notebook interactively by running `/compute migrate` inside Genie, which scans eligibility, proposes changes as a labeled diff, and applies them reversibly with a reviewable diff and 30-day rollback.

Migrate a single job or notebook interactively after enabling the beta in workspace preview settings — the agent scans eligibility and shows a reviewable diff before applying anything.

📍 `/compute` migrate

— Names exact CLI version, command group, and slash command. launch-20261119-9c8460c9

02 **Genie ZeroOps autonomous pipeline agent** Not checked yet NEW 74

Genie ZeroOps is now in public beta and can be turned on for a workspace from the workspace Previews page, after which a banner appears at the top of the Pipelines page. It adds a shared inbox so a whole team sees the same pipeline incidents, custom agent instructions to embed business rules (flaky sources, preferred fixes, naming conventions) into recommendations, one-click PR creation with pre-populated code fixes and details for Git-backed jobs, and notifications with issue summaries sent to Slack, Teams, or PagerDuty (with custom instructions to filter low-priority alerts). Access is governed by Unity Catalog permissions, PII is redacted before any notification leaves Databricks, and it detects silent data quality issues — late-arriving table commits (freshness) and below-expected row counts (completeness) — via Data Quality Monitoring, in addition to hard job failures.

Have a workspace admin activate Genie ZeroOps so the team can start receiving autonomous pipeline triage in the current workspace.

> 📍 In the Databricks workspace UI, go to Settings › Previews and enable the ZeroOps feature. Once enabled, find the ZeroOps banner at the top of the Pipelines page.

— Names integrations and governance but only a UI path, no API/flag.

launch-20261119-dfda78f2

NEW SUPPORTED PLATFORM PostgreSQL 19 on Lakebase

announced · Nov 19, 2026 BLOG ↗

Lakebase adds PostgreSQL 19 on day-one GA, with REPACK CONCURRENTLY, query plan advice extensions, WAIT FOR LSN, and logical replication upgrades.

Reclaim table space without blocking reads or writes using the new REPACK (CONCURRENTLY) command, which rebuilds tables online and optionally clusters rows via `USING INDEX`; requires free capacity under `max_repack_replication_slots` and a non-deferrable primary key or index-based replica identity. Not checked yet

Stabilize unstable query plans without touching application SQL using new `pg_plan_advice` and `pg_stash_advice` extensions; stash advice by query ID and activate it per-database so PostgreSQL applies it automatically on matching queries. Not checked yet

Block a standby until it has replayed a specific WAL position to implement read-your-writes on replicas using new `WAIT FOR LSN`. Not checked yet

Get a clean get-or-create upsert pattern that returns both new and existing rows without phantom dead tuples using new `INSERT ... ON CONFLICT DO SELECT`. Not checked yet

Export query results directly as JSON from `COPY TO`, using `FORCE_ARRAY` to wrap output in a single JSON array instead of one object per line. Not checked yet

+9 more

WAS THIS USEFUL? 👍 👎

Snowflake Cortex AI

Feedback ⓘ announced · Nov 19, 2026

BLOG ↗ RANK

Snowflake Cortex AI provides managed language-model functions, search, and agents that work with data governed in Snowflake.

> Snowflake Cortex AI shipped GA functions for scoring and auto-tuning AI function quality and cost (`AI_FUNCTION_EVALUATION`, `AI_FUNCTION_OPTIMIZATION`) and added natural-language video search and analysis built on TwelveLabs' Marengo and Pegasus models. Alongside these, the same window brought deep Apache Iceberg support for Spark via Snowpark Connect and serverless compute plus new project-creation paths for Snowflake App Runtime.

└──▷ **WHAT SHIPPED · 5 FEATURES** most completely described first ? what's the number?

01 **AI function development and optimization tooling** Not checked yet

NEW 95

Snowflake adds generally available tooling for building and tuning AI functions in SQL: `AI_FUNCTION_EVALUATION` measures AI function output quality, token usage, and cost by testing functions against labeled examples with scoring methods including `llm_judge` and `exact_match`, startable via `/cortex-ai-function-studio` in CoCo or `AI > AI Functions > Evaluation | Optimization` in the Snowflake UI; `AI_FUNCTION_OPTIMIZATION` automates prompt-model search using agentic algorithms including Snowflake's CoCoEvolve and GEPA against provided evaluation data, scoring criteria, and budget, configurable through a guided Snowsight workflow, via CoCo, or directly in SQL against an existing function such as `RECOMMEND_NEXT_ACTION`; and reusable custom AI functions can now be defined in SQL to package instructions and model choice, so shared logic can be called across applications through Snowflake's existing access controls.

Launch an evaluation or optimization experiment interactively from Snowflake CoCo without writing SQL.

📍 1. In Snowflake, open CoCo. 2. Enter `/cortex-ai-function-studio`. 3. Select `Evaluate` to set up an evaluation using your function and labeled data.

— Exact scoring methods and entry command given, directly runnable Names algorithms and an example function but no runnable syntax Describes the capability but no syntax or example shown

launch-20261118-912aa133

02 **App Runtime platform updates** Not checked yet NEW 90

App Runtime adds new ways to start projects, including browser-based Next.js apps via `+ Add new → Next.js App` in Snowsight Workspaces with CoCo assistance, and CLI bootstrapping with `snow app setup` and `snow app deploy`, or `$snowflake-apps` in CoCo CLI and `/snowflake-apps` in CoCo Desktop; serverless compute supporting a Python runtime for FastAPI, Flask, or LangGraph apps with dependencies installed via `pip` or `uv`, running on Python 3.13, with automatic suspend when idle and resume in seconds; a new App Runtime admin page listing every app with filters for status and compute type without requiring a query; and custom URL prefixes for serverless App Runtime apps so an app's address stays stable across redeploys.

Deploy a new App Runtime project from the terminal when you prefer working outside the browser.

```
$ snow app setup
snow app deploy
```

Build and ship a Next.js internal tool entirely in the browser without setting up a local environment.

> 📍 In Snowsight, open Workspaces, then select `+ Add new` → `Next.js App`. Use CoCo to describe the app, preview it live, and deploy from the same screen.

– Exact commands and UI path given, runnable today Names frameworks and version but no deployment command Describes the page but not where to find it Bare statement with no configuration detail

launch-20261119-eaf76a7a

---

03 **New video AI capabilities** Not checked yet NEW 85

Snowflake adds new video AI capabilities: `AI_MULTI_EMBED` generates multimodal video embeddings indexing visual, audio, and transcription signals as 512-dimensional vectors stored directly in Snowflake tables, using the Marengo 3.0 multimodal encoder with transcription embeddings supporting 36 languages, processing video from internal Snowflake stages or external stages (S3, GCS, Azure) and enabling natural-language search of video libraries via Cortex Search; a generally available Cortex AI Function backed by the Pegasus 1.5 video-language model analyzes video and returns free-text answers or structured metadata; and Cortex Agents can orchestrate multi-vector video search across visual, audio, and transcription modalities using tool calling and custom ranking strategies.

– Names function, encoder, vector size and stages but no call syntax Names the model but not the function to call Capability named but mechanism of ranking not detailed

launch-20261118-44a1a402

---

04 **Expanded Iceberg support in Snowpark Connect** Not checked yet

NEW 82

Snowpark Connect for Apache Spark gains a suite of new Iceberg capabilities: write-audit-publish workflows using Iceberg branching PRIVATE PREVIEW that land each Spark batch on a named branch before publishing to main via a single procedure call with no data rewrite and no impact to readers of main; `DataFrameWriterV2` APIs PUBLIC PREVIEW for writing to Iceberg tables with minimal code changes to existing PySpark, Java, and Scala jobs; flexible partitioning PRIVATE PREVIEW supporting expression and bucket partition transforms and partition evolution without a full table rewrite; Time Travel and tags PUBLIC PREVIEW letting Spark jobs query any past Iceberg snapshot, roll back to a known-good state, and use named tags as stable, shareable checkpoints across engines; incremental reads that let Spark pipelines process only the changes between Iceberg snapshots instead of scanning the full table; and support for externally managed Iceberg catalogs including Apache Polaris, AWS Glue, and Unity Catalog through catalog-linked databases alongside Snowflake-managed tables, with new Iceberg SQL extensions and DataFrame options letting existing Spark code run unchanged against both.

— Mechanism is clear but no exact command or procedure name given Names the API but gives no code sample Mechanism named but no invocation shown Behaviour explained but no commands given Clear benefit but no API or flag named Names every supported catalog but no setup steps

launch-20261119-2462f727

THINNER COVERAGE BELOW

05 **Serverless compute to become default for App Runtime**

Not checked yet

35

BREAKING

When serverless compute becomes generally available, Snowflake plans to make it the default compute model for App Runtime and end the managed compute discount; timing has not yet been announced.

— Future change flagged but no date or migration step launch-20261119-eaf76a7a

└──▷ BREAKING ON UPGRADE

! When serverless compute becomes generally available, Snowflake plans to make it the default compute model for App Runtime and end the managed compute discount; timing has not yet been announced.

WAS THIS USEFUL? 👍 👎

Temporal

Feedback ⓘ announced · Nov 18, 2026 BLOG ↗ RANK

Temporal is a durable execution platform for reliable workflows, including long-running AI agents.

> Temporal Cloud launched on Microsoft Azure in Public Preview, adding Azure as a production-ready compute target alongside AWS and Google Cloud, with private connectivity and cross-cloud redundancy options.

└──▷ **WHAT SHIPPED · 3 FEATURES** most completely described first ? what's the number?

01 Temporal Cloud on Azure (Public Preview)

Not checked yet NEW

55

Temporal Cloud now runs on Azure in Public Preview, supporting production workloads in Central US and Australia East. Existing Temporal Cloud customers can add Azure as a compute target alongside their existing AWS or Google Cloud deployments.

— Names regions and preview status but no setup steps launch-20261118-6c18a6b2

02 Cross-cloud replication and failover

Not checked yet NEW

45

Applications needing redundancy across cloud providers can configure cross-cloud replication and failover, available for supported region pairs within the same geography.

— Describes capability and scope but no config keys launch-20261118-6c18a6b2

03 Private Link connectivity for Azure Workers

Not checked yet NEW

40

Workers can connect to Temporal Cloud privately through Azure Private Link.

— Names the mechanism but no config detail launch-20261118-6c18a6b2

WAS THIS USEFUL? 👍 👎

◆ Local LLM Runtimes

llamastash

Feedback ⓘ v0.12.0 NOTES ↗ RANK

Llamastash is a zero-overhead, terminal-native local-LLM orchestrator. Run any engine, any model via the same proxy.

> llamastash v0.12.0 hardens its loopback proxy against DNS-rebinding and cross-origin attacks and tightens fallback routing so embeddings, rerank and mismatched chat requests fail explicitly instead of silently substituting another model's output.

└──▷ WHAT SHIPPED · 5 FEATURES most completely described first ? what's the number?

01 Fallback routing safety controls

Not checked yet IMPROVED

90

Fallback routing for proxy requests can now be disabled globally by setting `proxy.fallback: false` in `config.yaml`. Fallback for `/v1/embeddings` and `/v1/rerank` requests is now disabled entirely by default, so a failed launch returns `503 launch_failed` instead of substituting another model's vectors or scores. Chat and completion fallback candidates are now limited to Ready launches that serve the same mode as the request, with a mismatched pick returning a `family_mismatch` reason instead of `launch_failed`. Fallback responses now stamp `x-llamastash-served-by` and `x-llamastash-fallback-reason` headers on the reply.

Turn off fallback routing globally so a failed model launch always returns an explicit error rather than silently substituting another model.

yaml

```yaml
proxy:
  fallback: false
```

— Config key, endpoints, status codes and headers are all actionable v0.12.0

02 DNS-rebinding and cross-origin protections in loopback proxy

Not checked yet NEW

85

llamastash's loopback proxy now rejects requests whose `Host` header is not `localhost`, a `*.localhost` name, or a loopback IP, returning `403 forbidden_host` with the refused host named in the message, blocking DNS-rebinding attacks. It also rejects requests whose `Origin` header is not the proxy's own origin or a recognised app scheme (`chrome-extension://`, `moz-extension://`, `safari-web-extension://`, `vscode-webview://`, `tauri://`), returning `403 forbidden_origin` including for `null` origins from sandboxed iframes.

– Exact headers, schemes and status codes given, no reproduction steps v0.12.0

03 **Vulkan engine build install option** Not checked yet NEW

65

The Vulkan engine build can be installed on Linux AMD systems with `llamastash init server --install gh-releases:vulkan`.

Install the Vulkan engine build on a Linux AMD machine where the default would otherwise pull the ROCm build.

```
$ llamastash init server --install gh-releases:vulkan
```

– Exact command given, but no explanation of what changes v0.12.0

THINNER COVERAGE BELOW

04 **Rust 1.97 required to build from source** Not checked yet BREAKING

50

Building llamastash from source via `cargo install` now requires Rust 1.97 or newer, up from 1.95.

– Exact version numbers but no migration guidance v0.12.0

05 **Hugging Face Xet-backed file downloads** Not checked yet IMPROVED

45

`pull` and the TUI pull dialog now download Hugging Face Xet-backed files via hf-hub 1.0.

– States the change but no usage detail v0.12.0

└──▷ **BREAKING ON UPGRADE**

! Building from source (`cargo install`) now requires Rust 1.97 or newer, up from 1.95.

WAS THIS USEFUL? 👍 👎

LocalAI

Feedback ⓘ v5.0.1 NOTES ↗ CODE ↗ RANK

LocalAI runs language models locally through self-hosted, OpenAI-compatible APIs.

> LocalAI v5.0.1 brings native Windows support for the server and llama-cpp backend, alongside a macOS Apple Silicon sandbox, a glibc-free Linux AppImage, Bangla UI localization, ROCm 10.1 image updates, and desktop inbox notifications.

└──▷ **WHAT SHIPPED · 6 FEATURES** most completely described first ? what's the number?

01 **Native Windows support for server and llama-cpp backend**

Not checked yet NEW 80

LocalAI now runs natively on Windows via a new `windows/amd64` llama-cpp backend, with OS-aware default directories resolved through system.ResolveDefaultDirs() — no WSL or container required.

Run LocalAI server natively on Windows with the llama-cpp backend — no WSL or container needed.

```
$ local-ai run gemma-4-e2b-it-qat-q4_0
```

— Names backend, resolver function, and a runnable command v5.0.1

THINNER COVERAGE BELOW

02 **Bangla language support in UI** Not checked yet NEW 50

Added Bangla (`bn`) localization to the React UI and built-in add-ons, selectable from the UI language switcher.

— Names locale code and UI path to enable it v5.0.1

03 **Sandbox add-on for macOS Apple Silicon** Not checked yet NEW 35

Enabled the Sandbox add-on `darwin/arm64` package for macOS Apple Silicon hosts.

— States what shipped but no usage or mechanism v5.0.1

04 **glibc-free Linux AppImage** Not checked yet IMPROVED 35

The x64 AppImage can now start on Linux hosts without glibc.

— Names the binary and constraint but no further detail v5.0.1

05 **ROCm 10.1 hipblas and privacy-filter images** Not checked yet NEW

25

Added ROCm 10.1 hipblas images and a new ROCm privacy-filter image.

— Named only, no mechanism or usage described v5.0.1

06 **Desktop inbox notifications** Not checked yet NEW 10

Added desktop inbox notifications.

— Bare mention with no detail on scope or trigger v5.0.1

WAS THIS USEFUL? 👍 👎

Ollama

Feedback ⓘ v0.40.17 NOTES ↗ CODE ↗ RANK

Ollama downloads, serves, and chats with local language models through a CLI and REST API.

> Ollama's v0.40.17 release adds local support for the new Inkling 2 Mini model, improves three editor/agent launch integrations (Claude Code, OpenCode, Pi), refines API tool-call and thinking-control behavior, and ships several CLI quality-of-life fixes.

└──▷ **WHAT SHIPPED · 8 FEATURES** most completely described first ? what's the number?

01 **Inkling 2 Mini model support** Not checked yet NEW 85

Run the new Inkling 2 Mini model (26B parameters, 5.4B active) locally on Apple silicon with `ollama run inkling-2-mini`; it handles text, images, tool calls, and thinking at around 20 GB.

Run the new Inkling 2 Mini model locally on Apple silicon for coding tasks that benefit from its thinking and vision capabilities.

```
$ ollama run inkling-2-mini
```

— Names exact command, parameter count, and memory footprint. v0.40.17

02 **ChatGPT profile restore confirmation** Not checked yet NEW 65

`ollama launch chatgpt --restore` now prompts for confirmation before overwriting existing settings when restoring a ChatGPT profile.

Restore a ChatGPT profile without accidentally overwriting existing settings — Ollama will prompt for confirmation before making any changes.

```
$ ollama launch chatgpt --restore
```

— Exact flag and command shown via usage example. v0.40.17

03 **Deferred MCP tool loading for Claude Code** Not checked yet IMPROVED 60

`ollama launch claude` now supports Claude Code's deferred tool loading, so MCP tools are loaded on demand instead of all at once.

— Names command and mechanism but no runnable example. v0.40.17

THINNER COVERAGE BELOW

04 **OpenCode thinking options from model metadata** Not checked yet

IMPROVED 50

`ollama launch opencode` now sets thinking options from the model's metadata, so supported levels and defaults are correct for each model.

— Names command but no concrete values or example. v0.40.17

05 **Pi launch sign-in refresh and prompt passthrough** Not checked yet

IMPROVED 50

`ollama launch pi` now refreshes Pi's model list after sign-in and passes along prompt arguments.

— Names command and two behavior changes, no example. v0.40.17

06 **API tool definitions preserve unmodeled JSON schema keywords**

Not checked yet IMPROVED 45

Tool definitions passed through the API now preserve JSON schema keywords Ollama does not model instead of dropping them.

— Describes behavior change but no endpoint or example given. v0.40.17

07 **Disabling thinking maps to lowest reasoning level** Not checked yet

IMPROVED 40

Disabling thinking now maps to the model's lowest supported reasoning level instead of being ignored.

— Clear before/after but no named flag or example. v0.40.17

08 **Help output for argument-less non-interactive shell runs**

Not checked yet IMPROVED 40

Running `ollama` with no arguments in a non-interactive shell now prints help instead of erroring.

— Exact CLI behavior named, minor scope. v0.40.17

WAS THIS USEFUL?  

◆ Inference Serving & Optimization

Experiential Labs Experiential

Feedback ⓘ

v0.7.260 NOTES ↗ RANK

Experiential is the open source, zero markup gateway for BYOK, self-hosted and 1000+ marketplace models. It learns from your traffic to cut costs, recommend better models, and train a specialized model you own.

> Experiential added Claude plan usage across all supported harnesses and image-edit capability for several image models, alongside a cluster of Gemini routing and reliability fixes in the gateway.

└──▷ **WHAT SHIPPED · 3 FEATURES** most completely described first ? what's the number?

01 Claude plan usage across all harnesses

Not checked yet NEW

75

Claude plan usage is now enabled across all supported harnesses — OpenCode, Pi, Hermes, Codex, Copilot, Cursor, and Cline — by prepending Claude Code's billing-header block to system prompts when it is missing; real Claude Code, API-key, and house rungs pass through unchanged.

— Names mechanism and every harness, but no exact command v0.7.260

02 Gemini routing and reliability fixes

Not checked yet IMPROVED

70

Gemini continuations now route to their signature's issuer, and a repeatedly broken tool call now returns a 400 error (gateway change in `exp-gateway-native` 0.3.204). Gemini output blocks surfaced via OpenRouter are now classified and filed as refusals, and a malformed Gemini function call is retried using VALIDATED calling mode.

— Names gateway version and exact retry mode, but fixes are automatic v0.7.258 v0.7.260

THINNER COVERAGE BELOW

03 Image-edit support for GPT Image, Azure, xAI Grok Imagine

Not checked yet NEW

45

Adds image-edit capability on the GPT Image (OpenAI and Azure) and xAI Grok Imagine rungs.

— Names the models but no usage detail or limits v0.7.258

WAS THIS USEFUL?  

Etalab OpenGateLLM

Feedback ⓘ 0.9.3 NOTES ↗ RANK

OpenGateLLM is an open-source, unified interface for running and managing self-hosted LLMs.

> OpenGateLLM 0.9.3 introduces per-provider concurrency limits and admission retries, and ships a cluster of breaking changes around deployment migrations, auth secrets and environment-variable config loading, alongside smaller admin and playground improvements.

└──▷ **WHAT SHIPPED · 11 FEATURES** most completely described first ? what's the number?

01 Per-provider concurrency limits and admission retries

Not checked yet

90

BREAKING

Admins can cap how many concurrent requests a provider accepts and configure router retry behavior via `qos_limit` and `qos_retries_before_reject` fields on `PATCH` requests to `/v1/admin/providers/{provider_id}` and router endpoints; requests that exceed the limit return `503` with a `Retry-After` header instead of failing outright. `PATCH` bodies that omit these fields now receive `422` — pass `null` explicitly to keep a provider or router unlimited.

Cap how many concurrent requests a provider accepts and how many times the router retries before returning 503, to protect an inference engine under sustained load.

```
$ PATCH /v1/admin/providers/{provider_id}
Content-Type: application/json

{
  "qos_limit": 20,
  "qos_retries_before_reject": 3
}
```

— Exact fields, endpoint, status codes and a runnable PATCH example 0.9.3

02 Manual database migration required before deploy

Not checked yet

85

BREAKING

Containers no longer run database migrations automatically on startup; operators must run `python -m alembic -c api/alembic.ini upgrade head` once before deploying and update Compose deployments to add the new `migrate` service and its dependency.

— Exact migration command and Compose change given 0.9.3

03 **Build-time placeholder variables for playground image**

Not checked yet NEW 83

The playground Docker build now accepts a `--build-arg BUILD_PLACEHOLDER_VARIABLES="NAME=placeholder"` argument so build-time placeholder values can be supplied separately from real runtime secrets, keeping secrets out of the image layer.

Build the playground image with placeholder values so real secrets are never baked into the image layer.

```
$ docker build --build-arg
BUILD_PLACEHOLDER_VARIABLES="OPENAI_API_KEY=placeholder,DB_PASSWO
RD=placeholder" -t opengatellm-playground .
```

— Named build arg plus a runnable docker build example 0.9.3

04 **Explicit auth_secret_key required, auth_master_key removed**

Not checked yet BREAKING 80

`settings.auth_secret_key` must now be set explicitly; the deprecated `auth_master_key` fallback has been removed. Reuse the previous key's value to preserve existing API keys, since changing the signing secret invalidates them.

— Named config key with clear migration guidance 0.9.3

05 **Environment-variable and secrets config loading hardening**

Not checked yet BREAKING 75

Environment variables are now resolved after YAML parsing: missing or empty variables without defaults, and invalid `${...}` names, now prevent API and playground startup; `${VAR:-}` produces `""` instead of `null`, and variables can no longer inject YAML structures. Secrets previously truncated by characters such as `#` or `:` are now read in full, which may change the effective signing secret and require recreating API keys.

— Multiple named parsing behaviors but no runnable example 0.9.3

THINNER COVERAGE BELOW

06 **Client-disconnect cancellation for model-forward requests**

Not checked yet NEW 55

When a client disconnects mid-request, OpenGateLLM now cancels the in-flight model-forward request and records it with status `499` instead of letting it run to completion.

— Named status code but no config or flag to invoke 0.9.3

07 **me/info endpoint removal and expired-user rejection**

Not checked yet

50

BREAKING

The `me/info` endpoints have been removed, and expired users are now rejected at authentication time rather than later in the request flow.

— Named endpoint removed, but no migration guidance 0.9.3

08 **Alphabetical sorting of model listing endpoint** Not checked yet

IMPROVED

45

Results from `GET /v1/models` are now sorted alphabetically by model name.

— Named endpoint but a passive behavior change 0.9.3

09 **basic_auth removed from admin providers response** Not checked yet

45

BREAKING

`basic_auth` is no longer included in the response body of `POST /v1/admin/providers`.

— Named endpoint and field, minimal explanation 0.9.3

10 **Admin search filtering by role and organization** Not checked yet

NEW

40

The playground admin search bar can now filter users by role and by organization.

— Clear UI path but no further mechanism detail 0.9.3

11 **Per-model rate-limit usage visibility** Not checked yet NEW

30

Users can now see how much of each allowed model's rate limit they have consumed, with per-model usage data surfaced directly in API responses.

— No field name or endpoint given for the usage data 0.9.3

└──▷ **BREAKING ON UPGRADE**

! Run `python -m alembic -c api/alembic.ini upgrade head` once before deploying; containers no longer migrate on startup — update Compose deployments with the new `migrate` service and its dependency.

! Set `settings.auth_secret_key` explicitly; the deprecated `auth_master_key` fallback is removed — reuse its previous value to preserve existing API keys, as changing the signing secret invalidates them.

! Environment variables are now resolved after YAML parsing: missing or empty variables without defaults, and invalid `${...}` names, prevent API and playground startup; `${VAR:-}` produces `""` not `null`, and variables can no longer inject YAML structures.

! Admin provider and router `PATCH` bodies must include the new `qos_limit` / `qos_retries_before_reject` fields or receive `422`; pass `null` to preserve unlimited admission.

! `basic_auth` is removed from the `POST /v1/admin/providers` response body.

! Secrets previously truncated by YAML characters such as `#` or `:` are now read in full, which may change the effective signing secret and require recreating API keys.

WAS THIS USEFUL? 👍 👎

◆ Inference Providers & GPU Clouds

Anyscale

Feedback ⓘ Version 0.26.118 release notes NOTES ↗ RANK

Anyscale provides a managed Ray platform for distributed AI development, training, batch processing, and model serving.

> Anyscale's 0.26.118 release substantially expands CLI and SDK observability, adding dedicated commands for events, logs, metrics, node/GPU inspection, task/actor aggregation, and scheduler introspection, plus support for injecting custom HTTP headers and a breaking change to CLI exit codes.

└──▷ WHAT SHIPPED · 8 FEATURES most completely described first ? what's the number?

01 **Scheduler queue and consumer inspection** Not checked yet NEW

85

Inspect the scheduler with `anyscale scheduler queue list`, `anyscale scheduler queue get`, and `anyscale scheduler queue workloads` for live queue state, `anyscale scheduler overview` for an organisation-wide summary, and `anyscale scheduler top-consumers` to list the clusters holding the most nodes, or via anyscale.scheduler.list_queues(), anyscale.scheduler.get_queue(), anyscale.scheduler.list_queue_workloads(), anyscale.scheduler.get_overview(), and anyscale.scheduler.list_top_consumers() in the SDK.

— Describes each command's purpose and scope clearly Version 0.26.118 release notes

02 **Custom HTTP header injection for CLI and SDK** Not checked yet

NEW

84

Inject custom HTTP headers into every CLI or SDK request (e.g. for a proxy token) using the `ANYSCALE_EXTRA_HEADERS` environment variable (JSON object of header names to values), the `--header 'Name: Value'` option before any command, or Anyscale(headers=...) in the SDK; later definitions override earlier ones for the same header name.

Pass a proxy authentication header to every Anyscale CLI request without exposing it in shell history.

```
$ export ANYSCALE_EXTRA_HEADERS='{"X-Proxy-Token": "my-secret-token"}'
anyscale job list
```

— Names env var, flag, SDK call, and override rule, plus example

Version 0.26.118 release notes

03 **Log reading and LogQL search** Not checked yet NEW 80

Read log files for services, workspaces, and clusters with `anyscale service logs`, `anyscale workspace_v2 logs`, and `anyscale cluster logs`, and search logs with LogQL using `anyscale logs query`, or via `anyscale.logs` in the SDK.

– Names commands and the LogQL query mechanism Version 0.26.118 release notes

04 **Metrics listing and PromQL querying** Not checked yet NEW 80

List metric names for a resource with `anyscale metrics names`, and evaluate PromQL queries at a point in time or over a window with `anyscale metrics query` and `anyscale metrics query-range`, or via `anyscale.metrics` in the SDK.

– Names commands and PromQL mechanism, no example given Version 0.26.118 release notes

05 **Task and actor aggregation across clusters** Not checked yet NEW 77

Group cluster tasks and actors with `anyscale task aggregate` and `anyscale actor aggregate`; pass `--no-start-cluster` to fail instead of auto-starting the system cluster, or use `anyscale.task` and `anyscale.actor` in the SDK.

– Names commands and a behavioural flag with clear effect Version 0.26.118 release notes

06 **Event listing for jobs, services, workspaces, clusters** Not checked yet NEW 75

List events for jobs, services, workspaces, and clusters with `anyscale job events`, `anyscale service events`, `anyscale workspace_v2 events`, and `anyscale cluster events`, or via `anyscale.events` in the SDK.

– Names every command but gives no usage example Version 0.26.118 release notes

07 **Structured CLI exit codes for image build** Not checked yet 75

BREAKING

`anyscale image build` now exits with structured exit codes (1 for credential issues, 2 for server/rate-limit errors, 3 for timeout, 4 for permission failures) instead of always exiting 1; set `ANYSCALE_LEGACY_EXIT_CODES=1` to restore the previous behaviour.

– Exact codes and rollback env var given, no example Version 0.26.118 release notes

08 **Node and GPU inspection commands** Not checked yet NEW 72

Show cluster nodes and GPU details with `anyscale node list`, `anyscale node get`, `anyscale node gpus`, and `anyscale gpu get` (by UUID), or via `anyscale.node` and `anyscale.gpus` in the SDK.

— Four named commands but no example of output

Version 0.26.118 release notes

└──▷ BREAKING ON UPGRADE

! `anyscale image build` now exits with structured exit codes (1 for credential issues, 2 for server/rate-limit errors, 3 for timeout, 4 for permission failures) instead of always exiting 1; set `ANYSCALE_LEGACY_EXIT_CODES=1` to restore the previous behaviour.

WAS THIS USEFUL? 👍 👎

Microsoft Foundry

Feedback ⓘ announced · Nov 17, 2026

BLOG ↗ RANK

Microsoft Foundry provides managed AI models, agent development, evaluation, observability, and deployment services on Azure.

> Microsoft Foundry deepened its Oracle AI Database integration, adding governed Managed MCP server connections and A2A multi-agent protocol support for agents, alongside a new Oracle Excel plug-in and a limited-preview Azure Workload Center for Oracle.

└──▷ **WHAT SHIPPED · 4 FEATURES** most completely described first ? what's the number?

01 **Managed MCP servers for Oracle AI Database agents** Not checked yet

NEW INTEGRATION NEW 65

Agents built in Microsoft Foundry can connect to Oracle AI Database using Managed Model Context Protocol (`MCP`) servers, with access governed through approved SQL and PL/SQL operations, tools, and predefined reports.

– Names governance mechanism but no setup command or config key launch-20261117-f48ec83f

THINNER COVERAGE BELOW

02 **A2A protocol for multi-agent Oracle workflows** Not checked yet

NEW 55

Microsoft Foundry now supports building multi-agent workflows using the A2A protocol, enabling coordinated agent activity across Oracle AI Database running on Oracle Cloud Infrastructure or as Oracle AI Database@Azure.

– Explains scope across infra but no concrete steps launch-20261117-f48ec83f

03 **Azure Workload Center for Oracle (limited preview)**

Not checked yet NEW 55

Adds a Workload Virtual Instance (`WVI`) construct that brings each workload's architecture, resources, and dependencies into a single view, integrated with Azure Monitor Health Models and the Microsoft Azure Copilot Run Oracle agent.

– Names construct and integrations but preview access unclear launch-20261117-f48ec83f

04 **Oracle AI Database plug-in for Excel** Not checked yet NEW

45

A new Oracle AI Database plug-in for Excel lets users pull Oracle Database data directly into worksheets and reason over it with Copilot without exporting or preparing data manually.

.

— Describes benefit, lacks install or activation detail launch-20261117-f48ec83f

NEW INTEGRATION Insights observability for LangGraph, LangChain, and CrewAI agents

NEW INTEGRATION OpenAI and Anthropic plugins for Work IQ

NEW INTEGRATION Fabric IQ connectors for ChatGPT and Claude

NEW INTEGRATION Fabric IQ ontologies in GitHub Copilot

NEW INTEGRATION SharePoint (indexed) knowledge source with permissions support

NEW INTEGRATION Retrieve API for Microsoft 365 and federated connector data

NEW INTEGRATION Business Applications in Work IQ adding CRM and ERP records

NEW SUPPORTED PLATFORM Web IQ Places and Weather sources

announced · Nov 17, 2026 BLOG ↗

Microsoft Foundry adds a continuous improvement loop with new preview and GA capabilities including agent optimizer, skill optimization, cost optimization, annotations on traces, golden datasets, Programmed Training, Insights, hosted agents, Microsoft Copilot harness, and voice agents.

Enable agents to learn from production traces and expert feedback using annotations on traces (now in preview) so reviewers can mark real conversations and feed results back into evaluations. Not checked yet

Build evaluation datasets from annotated real-usage traces (now in preview) instead of hypothetical test cases, and lock in a golden dataset (now in preview) as the benchmark each new agent version must beat. Not checked yet

Add safety criteria alongside quality criteria to rubric evaluators (now in preview) to score agent responses against business-defined checklists. Not checked yet

Send a recurring production issue directly to agent optimizer (now in preview) to generate candidate fixes tested against your evaluations in one step. Not checked yet

Use skill optimization (now in preview) to have agent optimizer review past runs and expert feedback, then generate versioned new or improved skills that can be rolled back or reused across agents. Not checked yet

+6 more

announced · Nov 17, 2026 BLOG ↗

Microsoft Foundry adds knowledge graphs and a new retrieval model for Foundry IQ, Programmed Training for custom fine-tuning loops, Agent Optimizer for automated

agent improvement, and usage-based Foundry IQ Serverless pricing.

Knowledge graphs in Foundry IQ, now in preview, let agents reason across relationships in organizational knowledge spanning unstructured data and applications. Not checked yet

Microsoft Retriever 1, now in preview, is a new retrieval model that improves results for agentic grounding. Not checked yet

Enable Programmed Training in public preview to write custom training loops — including supervised fine-tuning and reinforcement learning — for supported open-weight models while Foundry manages the compute. Not checked yet

Agent Optimizer in Foundry now evaluates an agent against team-defined success criteria, generates improved configurations ranked by quality, latency, and cost per successful outcome, and surfaces the best one for deployment. Not checked yet

Foundry IQ Serverless, generally available in December 2026, adds usage-based pricing with no capacity to provision or manage and includes free monthly allowances. Not checked yet

+11 more

WAS THIS USEFUL? 👍 👎

AWS Amazon SageMaker AI

Feedback ⓘ November 17, 2026

NOTES ↗ RANK

Amazon SageMaker AI provides managed tools and infrastructure for building, training, fine-tuning, and deploying machine-learning models.

> Amazon SageMaker AI shipped new IAM managed policies for HyperPod global inference gateways and raised pod density on HyperPod EKS clusters via dynamic ENI attachment.

└──▷ **WHAT SHIPPED · 2 FEATURES** most completely described first ? what's the number?

01 IAM policies for HyperPod global inference gateway

Not checked yet NEW 80

Three new AWS managed IAM policies — `AmazonSageMakerHyperPodInferenceGatewayGlobalRoutingAccess`, `AmazonSageMakerHyperPodInferenceGatewayMetricsPublisherAccess`, and `AmazonSageMakerGlobalInferenceGatewayExecutionAccess` — can be attached to roles used by SageMaker HyperPod inference gateways that join a global inference gateway.

Attach all three managed policies to an IAM role used by a HyperPod inference gateway joining a global inference gateway.

```
$ aws iam attach-role-policy --role-name <HyperPodGatewayRole> --policy-arn
arn:aws:iam::aws:policy/AmazonSageMakerHyperPodInferenceGatewayGlobalRoutingAccess
aws iam attach-role-policy --role-name <HyperPodGatewayRole> --policy-arn
arn:aws:iam::aws:policy/AmazonSageMakerHyperPodInferenceGatewayMetricsPublisherAccess
aws iam attach-role-policy --role-name <HyperPodGatewayRole> --policy-arn
arn:aws:iam::aws:policy/AmazonSageMakerGlobalInferenceGatewayExecutionAccess
```

— Names all three policies and gives attach command November 17, 2026

THINNER COVERAGE BELOW

02 Dynamic ENI attachment for HyperPod EKS clusters

Not checked yet

IMPROVED 45

SageMaker HyperPod clusters orchestrated by Amazon EKS now support dynamic elastic network interface (ENI) attachment, raising the pods-per-node limit to the instance type's maximum.

— Explains mechanism and effect but no config steps given November 11, 2026

WAS THIS USEFUL? 👍 👎

DATA

◆ Vector DBs & RAG

liliang-cn CortexDB

Feedback ⓘ v2.160.0 NOTES ↗ RANK

CortexDB is a local-first AI agent memory with a knowledge graph, ontology constraints, and traceable reasoning—all in one SQLite file.

> CortexDB trimmed graph query responses by dropping node embeddings by default and added time-aware memory recall, including displacement tracking and point-in-time queries.

└──▷ WHAT SHIPPED · 2 FEATURES most completely described first ? what's the number?

01 **Displacement tracking and point-in-time memory recall**

Not checked yet NEW 55

Memory recall now returns what a hit replaced, so callers can see which prior memory was displaced by the current one. It also accepts a past-instant parameter, letting callers query what the memory held at any earlier point in time.

— Explains mechanism but no concrete parameter name or API given v2.159.0

02 **Default embedding omission in graph tools** Not checked yet

IMPROVED 50

CortexDB graph tools now omit node embeddings from answers by default, reducing response payload size; embeddings must be requested explicitly to be returned.

— Describes behavior change but no flag name given v2.160.0

WAS THIS USEFUL? 👍 👎

dosco GraphJin

Feedback ⓘ v3.21.22 NOTES ↗ RANK

GraphJin is a governed graph for AI agents — GraphQL + MCP over your databases, files, APIs, and code.

> GraphJin renamed its CLI to `deeporg` (keeping `graphjin` as a compatible alias) and added a BigQuery integration, while significantly expanding its agent platform with a write-approval workflow for data changes, an inbox for notifications, OpenAPI docs for agent endpoints, answer charting, service identity for org agents, an evaluation system for the proactive agent, and catalog improvements.

└──▷ **WHAT SHIPPED · 11 FEATURES** most completely described first ? what's the number?

01 **Write-approval workflow for agent data changes** Not checked yet

NEW

65

Agent data writes can now be held for user approval or rejection before they execute via the console chat, and routines running on the data agent can propose data changes for their owner's approval, with the lead's reply stating the change made.

– Describes approval flow but no API or flag named. v3.21.21

02 **BigQuery integration for GraphJin serve** Not checked yet NEW

60

GraphJin serve can now connect to BigQuery via the new feat(bigquery) integration, which is now testable in CI against the BigQuery driver.

– Names integration tag but no config or flag details. v3.21.21

THINNER COVERAGE BELOW

03 **CLI renamed to `deeporg`** Not checked yet IMPROVED

55

The primary CLI command is renamed to `deeporg`, with `graphjin` kept as a backward-compatible alias.

– Names both commands but no further mechanism given. v3.21.22

04 **Evaluation system for proactive agent** Not checked yet NEW

55

Adds routine eval, smoke suites, and docs for the proactive agent, including an eval system that can mine tests from questions people have already asked, and a public eval suite `2028.5` that scores saved metrics on the agent's query.

– Names eval suite version but no run command given. v3.21.21

05 **Answer charting in console** Not checked yet NEW 45

Agent answer charts are now drawn in the console from engine-written display hints.

— Mechanism named but no UI path or config. v3.21.21

06 **Catalog table cards and saved-query handling** Not checked yet 45

IMPROVED

Table cards in the catalog now teach names, joins, and a database total, and the data agent now reads saved queries rather than running them, excluding saved queries from its catalog.

— Describes behavior change but no config or API named. v3.21.21

07 **OpenAPI description for agent endpoints** Not checked yet NEW 40

Adds an OpenAPI description for the agent endpoints, making the agent API surface machine-readable and explorable.

— No endpoint paths or spec location given. v3.21.21

08 **Inbox for agent notifications** Not checked yet NEW 35

Adds an inbox that GraphJin posts notifications to when the user is away.

— No UI path or API detail provided. v3.21.21

09 **Agent continuations, autonomy levels and budgets** Not checked yet 35

NEW

Agents now support continuations, follow-up questions, autonomy levels, and budgets.

— List of capabilities without mechanism or config. v3.21.21

10 **Service identity for organization agents** Not checked yet NEW 30

Adds service identity for organization agents, with delegation recorded.

— Brief mention, no config or API named. v3.21.21

11 **Expected-outcome ledger for changes** Not checked yet NEW 30

Changes now carry their expected outcome, checked later on a ledger.

— No mechanism for how the ledger is accessed or checked. v3.21.21

WAS THIS USEFUL?  

stashbase

Feedback ⓘ v2.18.13 NOTES ↗ RANK

stashbase is an open-source, local-first writing workspace where you write with Claude Code and Codex, drawing on your own sources.

> stashbase v2.18.13 adds local voice typing to Chat, context menus for document tabs, and link-preserving file renaming, while removing the persona selection feature.

└──▷ **WHAT SHIPPED · 5 FEATURES** most completely described first ? what's the number?

01 **Local voice typing in Chat** Not checked yet NEW 65

stashbase adds local voice typing to Chat with automatic model preparation on first use, offline reuse after preparation, and handling for long recordings, cancellation, and retries.

Use local voice typing in Chat to dictate a query without an internet connection — the model is prepared once and reused offline from then on.

📍 In Chat, click the microphone icon to start voice typing. stashbase will prepare the local model on first use; subsequent recordings work offline.

– Mechanism and offline behavior explained, with usage steps given v2.18.13

THINNER COVERAGE BELOW

02 **Link-preserving file renaming** Not checked yet IMPROVED 45

Renaming files now preserves Markdown, HTML, and Wiki links, including names with spaces and links with titles or nested punctuation.

– Describes scope of fix but no command or setting to act on v2.18.13

03 **Context menus on document tabs** Not checked yet NEW 40

Document tabs now have context menus, improving copy, reveal, and close actions.

– Names actions but no exact UI path or mechanism v2.18.13

04 **Persona selection removed** Not checked yet BREAKING 40

Persona selection and its library have been removed; Crew, project instructions, conversations, and model preferences remain.

– States what's removed and retained, no migration guidance v2.18.13

05 **Startup diagnostics and opt-out handling** Not checked yet IMPROVED

30

Adds independent desktop-server startup diagnostics and preserves usage-statistics opt-out during concurrent startup reporting.

— Brief description with no further mechanism or config named v2.18.13

└──▷ **BREAKING ON UPGRADE**

! Persona selection and its library have been removed; Crew, project instructions, conversations, and model preferences remain.

WAS THIS USEFUL? 👍 👎

infino

Feedback ⓘ v0.11.34 NOTES ↗ RANK

Infino is an embedded retrieval library built on Parquet. Fast, efficient, and scalable.

> Infino improved full-text ranked query performance with per-term level head tracking and sped up SQL equality lookups for both in-memory and on-disk data.

└──▷ **WHAT SHIPPED · 2 FEATURES** most completely described first ? what's the number?

01 **Per-term level heads for full-text search** Not checked yet

IMPROVED 40

Full-text search now tracks level heads per term and layer, improving performance for ranked full-text queries.

— Explains mechanism but no config or numbers given v0.11.34

02 **Faster SQL equality and IN lookups** Not checked yet IMPROVED

40

SQL `=` and `IN` predicates now look up a file's values in place, in memory and on disk, reducing query overhead.

— Names operators and mechanism but no benchmark or API v0.11.34

WAS THIS USEFUL? 👍 👎

◆ Document Parsing

anymd

Feedback ⓘ v8.5.14 NOTES ↗ RANK

anymd converts any file to clean Markdown for AI agents: PDF, Word, PowerPoint, Excel, EPUB, HTML and web pages.

> anymd adds a new `anymd cite-check` CLI command for citation verification and improves `anymd pro buy` with live pricing and automatic licence saving.

└──▷ WHAT SHIPPED · 2 FEATURES most completely described first ? what's the number?

01 **Citation verification via `anymd cite-check`** Not checked yet

NEW 80

The new `anymd cite-check` CLI command runs citation verification from the shell and now accepts `read` and `search` output directly, without requiring manual `bounding_box` coordinates.

Verify a citation directly from `read` output without supplying manual bounding-box coordinates.

```
$ anymd cite-check
```

— Names exact command and removes manual bounding_box step v8.5.14

02 **Live pricing and auto-save licence in pro buy** Not checked yet

IMPROVED 60

`anymd pro buy` now shows live pricing in the user's local currency before checkout, and automatically saves the offline licence upon purchase.

— Names command and behaviour but no further mechanism v8.5.14

WAS THIS USEFUL? 👍 👎

EVALUATE

◆ AI Observability & Evals

Profound

Feedback ⓘ announced · Nov 21, 2026 PRESS ↗

Profound is an AI search visibility platform that tracks how brands appear in answers from ChatGPT, Perplexity and other AI engines.

Profound adds Microsoft Teams presence and Microsoft 365 integrations, letting teams query AI Search visibility, run agents, and manage documents without leaving Teams.

NEW INTEGRATION Mention `@Profound` in a Microsoft Teams channel or direct message to query AI Search visibility, citation share, and accuracy changes, and get answers with full marketing data context. Not checked yet

Create and run agents or start a new Profound Project directly from a Teams message, then jump to the work via the Open project button. Not checked yet

Profound for Microsoft Teams is available to active subscribers and can be installed from the Microsoft Marketplace at `marketplace.microsoft.com`/en-us/product/wa200012602. Not checked yet

NEW INTEGRATION SharePoint and OneDrive integration

WAS THIS USEFUL? 👍 👎

Harbor Framework Harbor

Feedback ⓘ v0.27.0 NOTES ↗

RANK

Harbor is a framework for evaluating and improving AI agents through structured testing, metrics collection, and performance monitoring.

> Harbor v0.27.0 adds configuration options for the antigravity SDK and OpenCode v2 MCP servers, exposes new provider-level model limits, and migrates CLI artifact workflows onto API-managed SDKs.

└──▷ **WHAT SHIPPED · 4 FEATURES** most completely described first ? what's the number?

01 **System instructions option in antigravity SDK** Not checked yet

NEW 55

A new `system_instructions` option lets users pass system-level instructions to the agent via the antigravity SDK.

— Names the flag but not its format or limits v0.27.0

02 **Codemode option for OpenCode v2 task servers** Not checked yet

NEW 55

A new `codemode` option controls the coding mode used during task execution for OpenCode v2 task MCP servers.

— Names the flag but not its possible values v0.27.0

03 **Max output tokens and context window for custom endpoints**

Not checked yet NEW 50

The provider interface now exposes `max output tokens` and `context window` settings for custom endpoints.

— Names the settings but not where to configure them v0.27.0

04 **API-managed SDK transfers for CLI artifacts** Not checked yet

IMPROVED 50

CLI artifact workflows now move through API-managed SDK transfers instead of Hub and registry database calls, as part of completing the hosted, credentials, and authentication SDK migration with a new standalone Harbor API SDK.

— Explains the architectural change but no concrete surfaces v0.27.0

WAS THIS USEFUL?  

overmind

Feedback ⓘ v0.1.33 NOTES ↗ RANK

overmind is a platform for continuously improving AI agents through observability, evaluation, and fine-tuning.

> Overmind's v0.1.33 release adds a Reasoning Agent and rebuilt Data Workshop, expands the fine-tuning model catalog, and introduces consent-gated feedback loops from coding agents, while trimming supported inference providers.

└──▷ **WHAT SHIPPED · 6 FEATURES** most completely described first ? what's the number?

01 Consent-gated coding-agent feedback

Not checked yet NEW

60

Lets coding agents (Cursor , Claude Code , OpenCode , Codex) send feedback to Overmind with user consent, enabling agent-driven improvement loops.

— Names specific agents and consent mechanism, no config shown v0.1.33

THINNER COVERAGE BELOW

02 Baseten and Together AI removed as inference providers

Not checked yet BREAKING

55

Removes Baseten and Together AI as supported inference providers; any configuration pointing to them will stop working.

— Names providers and impact but no migration path given v0.1.33

03 Ornith, Fara, and Bolt E4B added to training catalog

Not checked yet NEW

50

Adds Ornith , Fara , and Bolt E4B to the training catalog as new base models available for fine-tuning.

— Names exact models but no usage detail v0.1.33

04 Reasoning Agent and Data Workshop Sandbox

Not checked yet NEW

45

Adds a Reasoning Agent and a Data Workshop Sandbox, extending the platform's agent-training capabilities with a dedicated reasoning component and an isolated execution environment.

— Names two new components but no mechanism detail v0.1.33

05 OpenRouter chat proxy disabled by default

Not checked yet

45

BREAKING

The OpenRouter chat proxy is now disabled by default, so it must be explicitly enabled before it activates.

— Clear behavior change but no config key given v0.1.33

06 **Data Workshop data-plane and dataset UI rebuild** Not checked yet

IMPROVED

40

Rebuilds the Data Workshop data-plane and dataset screens, replacing the previous UI with updated views for managing data pipelines and datasets.

— Describes scope but not specific UI changes v0.1.33

└──▷ **BREAKING ON UPGRADE**

! Baseten and Together AI are removed as inference providers; any configuration pointing to them will stop working.

WAS THIS USEFUL?  

Braintrust

Feedback ⓘ updated · Nov 21, 2026 NOTES ↗

Braintrust provides evaluation, tracing, and improvement workflows for AI applications.

Braintrust adds on-device trace translation, Loop pattern rules with SQL filters, scheduled team digests, a Decisions API for judges with confidence scores, and a grouped thread view PUBLIC PREVIEW.

Braintrust adds on-device whole-trace translation to English (including span fields, span names, thread messages, and custom trace views) using the browser's built-in translation API in desktop Chrome or Edge, keeping trace text on-device and consuming no model credits Not checked yet

Braintrust adds pattern rules so Loop knows what to investigate in each scheduled Patterns run, with optional SQL filters to narrow scope and a list of problems to watch for Not checked yet

Braintrust adds scheduled digests that send recurring project-activity summaries to a team, with configurable investigation description, compact or detailed writing style, schedule, and destinations Not checked yet

Braintrust adds a Decisions API for LLM-as-a-judge scorers and classifiers so the judge selects a choice directly rather than generating text to parse, recording confidence and per-choice probability for each result Not checked yet

Braintrust adds a grouped thread view PUBLIC PREVIEW that displays a multi-trace session as one ordered conversation, keeping tool calls with their originating request; grouping by `metadata.conversation_id` now includes traces that log the ID only on child spans Not checked yet

+5 more

NEW INTEGRATION Google Antigravity SDK tracing

NEW INTEGRATION Async Bedrock Runtime tracing via aiobotocore

NEW INTEGRATION OpenCode 2 tracing in bt CLI

NEW INTEGRATION Full-duplex GPT-Live Pipecat pipeline tracing

UPDATED INTEGRATION LiveKit Agents tracing extended with eou_wait spans and new model call spans

└──▷ BREAKING ON UPGRADE

! Terraform provider v0.0.9 requires Terraform or OpenTofu 1.11.1 or later.

! `braintrust_ai_secret` now requires `secret_version` alongside `secret`, and `braintrust_env_var` requires `value_version` alongside `value`; changing only the value no longer updates it, and references to `secret` or `value` return `null`.

- ! The `braintrust_org_automation` Terraform resource is removed; delete its blocks from your configuration and run `terraform state rm` on each address before running `terraform plan`.
- ! Python SDK v0.52.0: LiveKit Agents now uses `livekit.llm_response` for LLM calls it traces directly (replacing `LLM request` and `Realtime request`); for text-to-speech calls, `time_to_first_audio` moved from the `tts` span to `time_to_first_token` on the `livekit.tts` child — update saved queries, filters, and scorers accordingly.
- ! Python SDK v0.52.0: OpenAI spans no longer attach audio unless `BRAINTRUST_CAPTURE_USER_AUDIO_ATTACHMENTS` (input audio) or `BRAINTRUST_CAPTURE_AGENT_AUDIO_ATTACHMENTS` (generated audio) is set to `true`.

WAS THIS USEFUL? 👍 👎

GOVERN

◆ AI Runtime Defense & Guardrails

Prismor

Feedback ⓘ v1.80.0 NOTES ↗ RANK

Prismor is a self-hosted runtime control plane for AI agents. Observe or HITL approve or Block rogue tool calls before it executes: secret leaks, prompt injection, supply chain etc in a local dashboard. Agent agnostic (Claude, codex, langchain etc.)

> Prismor hardened secret-leak closures across several guard surfaces, added egress allowlist enforcement visibility, and expanded dashboard insight into MCP server activity and telemetry evidence.

└──▷ **WHAT SHIPPED · 4 FEATURES** most completely described first ? what's the number?

01 **Egress allowlist enforce mode and log-only indicator**

Not checked yet IMPROVED 55

Prismor's sandbox now indicates when the egress allowlist is running in log-only mode, and documentation adds the egress mode `enforce` setting.

– Names config setting but lacks full config path v1.79.0

02 **MCP server memory and usage visibility on dashboard**

Not checked yet NEW 40

Prismor's dashboard now shows MCP server memory usage and highlights rarely-used servers.

– Clear what it shows but no navigation detail v1.80.0

03 **Structured evidence file path in telemetry** Not checked yet

IMPROVED 40

Prismor's telemetry now ships the file path from structured evidence when full capture is enabled.

– Describes change but no exact field or flag name v1.79.0

04 **Secret-leak path closures across guards** Not checked yet IMPROVED

30

Prismor closed secret-leak paths across `env-guard`, `secret-guard`, MCP output and publish flows.

– Named guard surfaces but no mechanism or scope detail v1.80.0

WAS THIS USEFUL? 👍 👎

◆ Model & Agent Protection

ai-safe2-framework

Feedback ⓘ v3.5 NOTES ↗ RANK

Ai-Safe2-Framework provides governance, risk, and compliance controls for securing agentic AI systems and non-human identities.

> AI SAFE² v3.5 adds 12 new framework controls with 157 supporting clauses, three new CLI routes for querying and auditing the framework, an offline Rosetta Stone explorer with a much larger mapping dataset, and expanded MCP and threat-modeling coverage.

└──▷ WHAT SHIPPED · 9 FEATURES most completely described first ? what's the number?

01 **New CLI routes for framework queries and readiness**

Not checked yet NEW 82

Three new CLI commands let users query framework controls, audit mapping records, and run a scoped readiness workflow: `safe2 framework query --control CP.12`, `safe2 framework audit`, and `safe2 framework readiness --help`.

– Exact runnable commands given; underlying mechanism not detailed. v3.5

02 **Offline Rosetta Stone explorer dashboard** Not checked yet NEW

75

Open `dashboard/rosetta.html` from a local checkout to filter records by source, control, review status, and keyword, view record details, and export to JSON.

– Exact file path to open, with named filter and export capabilities. v3.5

03 **12 new framework controls across trust and execution**

Not checked yet NEW 75

Adds 12 new controls covering trust lifecycle, final-effect authorization, control evidence independence, organizational AI management governance, design-time coordination contracts, pre-trust initialization boundaries, aggregate consumable-authority ceilings, environment binding, idempotency and side-effect ledgers, verified outcome contracts, evidence-based failure attribution, and declared success criteria gates, with IDs `CP.11` – `CP.15`, `S1.8`, `F3.6`, `F3.7`, `A2.7` – `A2.9`, `E5.5`.

– Named control IDs and topics, but no usage workflow given. v3.5

04 **User Autonomy and Sovereignty component v2** Not checked yet

IMPROVED 63

Upgrades the User Autonomy and Sovereignty (UAS) component to v2, retaining its 27 IDs with revised acceptance interpretation for human understanding, appropriate

reliance, intervention, fairness, and evidence quality, plus reciprocal Cognitive Sovereignty Framework references.

— Describes scope and ID retention but no direct usage step. v3.5

05 **MCP profile and new deployment profiles** Not checked yet IMPROVED

60

Adds 12 child requirements to the MCP profile, which retains its 19 stable control IDs, alongside six new deployment profiles covering harnesses, gateways, evaluation, browser/computer use, finance, and physical systems.

— Names profile counts and scope areas, no config detail. v3.5

THINNER COVERAGE BELOW

06 **Rosetta Stone mapping layer expansion** Not checked yet IMPROVED

56

Adds 1,202 recovered source-to-control relationships, a 33-family source index, 64 discovery entries, and 16 audit-rule classes to the Rosetta Stone mapping layer.

— Quantified data expansion but no interface to act on it directly. v3.5

07 **Expanded clause and research documentation** Not checked yet

IMPROVED

50

Adds 157 explicit child clauses across 53 reviewed decisions, including two research-disposition clauses, deepening implementation and acceptance requirements for existing controls; also adds 53 research notes connecting source material, placement decisions, alternatives, enforcement boundaries, acceptance cases, and unresolved evidence to the research library.

— Quantified additions but no specific control IDs or navigation path. v3.5

08 **Three new threat profiles in AISM** Not checked yet NEW

45

Adds three threat profiles sharpening attention to assurance compromise, unsafe authorized execution, and human-agent trust exploitation within AISM.

— Names threat categories but no mechanism or location to find them. v3.5

09 **Machine-readable repository discovery and agent indexes**

Not checked yet NEW

40

Adds machine-readable repository discovery records, root and NEXUS agent indexes, and a complete page map to support both human navigation and agent workflows.

— Broad description without exact file paths or formats. v3.5

WAS THIS USEFUL?  

Minimal Software Research minimal Feedback ⓘ

release-14b57239 NOTES ↗ RANK

Minimal is a secure, declarative runtimes for Claude Code, Gemini, Codex, Pi and more, so you can run multiple agents safely on your stack from a verified software supply chain.

> This release adds a tamper-evident box audit trail, network and egress policy visibility and controls for box and task commands, richer `min box` introspection output, and smaller improvements to login, auth status, exec attach, diagnostics, and update checking.

└──▷ **WHAT SHIPPED · 8 FEATURES** most completely described first ? what's the number?

01 **Box audit trail with tamper detection** Not checked yet NEW

85

`min box audit` reads a box's full audit trail, `min box audit --parent` merges in the parent chain for nested boxes, and `min doctor` now reports when an audit chain has been edited.

Inspect the full audit trail for a box, including its parent chain, to review every egress decision made during a session.

```
$ min box audit mybox --parent
```

— Commands and a runnable example given; tamper-detection mechanism stays terse.

release-14b57239

02 **Network and egress policy controls for boxes and tasks**

Not checked yet NEW

70

`min box show <box> --network` displays a box's effective network policy, and `min task run` and `min box run` now take network and egress flags to control traffic at launch.

Show the effective network policy for a running box in JSON to pipe into downstream audit tooling.

```
$ min box show mybox --network -o json
```

— Named commands and a worked example; run-command flag names not listed.

release-14b57239

03 **Box listing, output formats, and store spec** Not checked yet NEW

65

`min box` gains list and show verbs for box types, with `min box show` accepting `-o json` or `-o yaml` output formats including exec details; `min box spec` lists each store reference along with its authorities.

— Named flags and subcommands but no worked example provided. release-14b57239

THINNER COVERAGE BELOW

04 **GitHub sign-in via min login** Not checked yet IMPROVED

50

`min login` is now the GitHub sign-in entry point, backed by a real browser page that lets users choose a repository.

— Describes the flow but gives no command example. release-14b57239

05 **Configurable detach keys for exec attach** Not checked yet NEW

40

Detaching from an exec attach session now supports configurable detach keys via `minimald`.

— Names minimald but not the config key or default binding. release-14b57239

06 **Vulnerability diagnostics flag** Not checked yet NEW 40

`min diag --security` reports a suspected vulnerability.

— Named flag but no detail on what vulnerabilities are checked. release-14b57239

07 **Proxy key naming in auth status** Not checked yet IMPROVED

35

`min auth status` now names each software proxy key.

— Single-line description with no further mechanism given. release-14b57239

08 **Daily release update check** Not checked yet NEW 35

`min` now checks once a day whether the install channel has a newer release and reports it.

— Behavior described but no config for frequency or channel. release-14b57239

WAS THIS USEFUL? 👍 👎

ToolHive

Feedback ⓘ v0.67.0 NOTES ↗ RANK

ToolHive runs MCP servers in isolated containers and enforces identity and access policies for local and Kubernetes deployments.

> ToolHive's v0.67.0 release adds a client_credentials OAuth grant and new association options for SPIFFE-authenticated workloads, structured findings for skill/plugin validation, and ships six breaking changes spanning SPIFFE resource scoping, transport defaults, and default container resource limits.

└──▷ **WHAT SHIPPED · 8 FEATURES** most completely described first ? what's the number?

01 Mandatory resource scoping for SPIFFE token exchange

Not checked yet 90

BREAKING

SPIFFE associations (`spiffeClientAuth`) now require a non-empty `resources` list; associations without `resources` are rejected and the auth server refuses to start without it. Token-exchange requests must send exactly one `resource` matching an entry in the association's `resources` ; sending `audience=` or no target now returns `invalid_target` .

— Clear before/after and exact field names, with migration step v0.67.0

02 Default resource limits on MCPRemoteProxy Not checked yet

90

BREAKING

`MCPRemoteProxy` now applies default CPU (`50m` request / `200m` limit) and memory (`64Mi` request / `256Mi` limit) resource constraints. Existing proxies whose working set exceeds 256Mi will be OOMKilled; operators should set explicit `spec.resources.limits` before upgrading.

— Concrete default values and migration note given v0.67.0

03 uvx transport default changed to stdio Not checked yet BREAKING

85

`thv run uvx://…` now defaults to the stdio transport instead of streamable-http; any `uvx://` invocation serving HTTP or SSE must add `--transport streamable-http` or `--transport sse` .

Run a uvx-packaged MCP server that speaks HTTP — explicitly override the new stdio default so the proxy can reach it.

```
$ thv run --transport streamable-http uvx://my-http-mcp-server --
--transport streamable-http
```

— Exact runnable command given for migration v0.67.0

04 **client_credentials grant and grantTypes for SPIFFE associations**

Not checked yet NEW 80

SPIFFE-authenticated workloads can now obtain access tokens using a `client_credentials` grant bound to their SPIFFE ID, with token lifetime capped at the SVID's remaining validity; the grant and SPIFFE auth methods are advertised in discovery. SPIFFE associations on `MCPExternalAuthConfig` and `VirtualMCPServer` gain a `grantTypes` field to select `client_credentials`, token exchange, or both, defaulting to token exchange.

— Names config field and grant semantics but no runnable command v0.67.0

05 **Structured findings in skill and plugin validation**

Not checked yet IMPROVED 75

`thv skill validate`, `thv ai-plugin validate` and their API endpoints now return a structured `findings` array with severity for skill and plugin validation results.

— Commands and field named, no sample findings shown v0.67.0

06 **act.sub claim now carries SPIFFE ID** Not checked yet BREAKING 75

The `act.sub` claim in a SPIFFE token exchange is now the workload's SPIFFE ID instead of the client ID. Cedar policies matching `claim_act.sub` against a SPIFFE client ID no longer match; they must be changed to match `claim_client_id` instead.

— Clear change and fix but no runnable example v0.67.0

07 **Token-exchange middleware moved in Go API** Not checked yet 75

BREAKING

The token-exchange middleware moved to `pkg/auth/tokenexchange`; it must be imported from there. `authserver.SPIFFEGrantTypeTokenExchange` should be replaced with `oauthproto.GrantTypeTokenExchange`.

— Exact import paths named for migration, developer-facing only v0.67.0

08 **Custom client name in DCR upstream config** Not checked yet NEW 60

Operators can set `clientName` in the DCR upstream configuration to control the name shown on the upstream IdP's authorization screen; it defaults to ToolHive MCP Client.

— Config key and default named, no example of use `v0.67.0`

└──▷ **BREAKING ON UPGRADE**

! SPIFFE associations (`spiffeClientAuth`) now require a non-empty `resources` list; associations without `resources` are rejected, and the auth server refuses to start without it. Add `resources` to every `spiffeClientAuth` entry.

! SPIFFE token-exchange requests must send exactly one `resource` matching an entry in the association's `resources`; sending `audience=` or no target now returns `invalid_target`.

! `thv run uvx://…` now defaults to the stdio transport instead of streamable-http. Add `--transport streamable-http` or `--transport sse` to any `uvx://` invocation that serves HTTP or SSE.

! The `act.sub` claim in a SPIFFE token exchange is now the workload's SPIFFE ID instead of the client ID. Cedar policies matching `claim_act.sub` against a SPIFFE client ID no longer match; change them to match `claim_client_id` instead.

! `MCPRemoteProxy` now applies default CPU (`50m` request / `200m` limit) and memory (`64Mi` request / `256Mi` limit) resource constraints. Existing proxies whose working set exceeds 256Mi will be OOMKilled; set explicit `spec.resources.limits` before upgrading.

! Go API: the token-exchange middleware moved to `pkg/auth/tokenexchange`; import it from there. Replace `authserver.SPIFFEGrantTypeTokenExchange` with `oauthproto.GrantTypeTokenExchange`.

WAS THIS USEFUL? 👍 👎

AI MODELS

◆ Open-Weight Models

ConvAI Innovations Laya

Feedback ⓘ v0.5.1 NOTES ↗

RANK

Laya is a non-autoregressive System 1 decision engine. Typed choice, score and yes/no decisions over any text in a single forward pass, in 100+ languages, with a router that picks the right checkpoint per request.

> Laya's only change this window was making ORT binary downloads optional in its Rust integration.

└──▷ WHAT SHIPPED · 1 FEATURE most completely described first ? what's the number?

01 **Optional ORT binary downloads in Rust integration** Not checked yet

IMPROVED 55

The Rust integration can now control whether ONNX Runtime binaries are fetched automatically, via the `download-binaries` setting.

— Names exact config flag but lacks default value or broader context. v0.5.1

WAS THIS USEFUL? 👍 👎

RadixArk DaisyRose 27B

Feedback ⓘ

released · Nov 21, 2026 MODEL ↗

DaisyRose 27B from RadixArk was released on 21 Nov 2026. 262K-token context; $0.15 in / $1.875 out per million tokens (OpenRouter). Not checked yet

WAS THIS USEFUL? 👍 👎

Thinking Machines Inkling 2 Mini

Feedback ⓘ

released · Nov 21, 2026 MODEL ↗

Inkling 2 Mini from Thinking Machines was released on 21 Nov 2026. 1049K-token context; $0.10 in / $0.40 out per million tokens (OpenRouter). Not checked yet

WAS THIS USEFUL? 👍 👎

Z.ai GLM 5.3 Flash

Feedback ⓘ price cut · Nov 20, 2026

MODEL ↗

GLM 5.3 Flash is a reasoning model from Z.ai.

Price cut on OpenRouter (z-ai/glm-5.3-flash): $0.15 → $0.0399 in, $0.50 → $0.0784 out per million tokens (-84%), first seen 20 Nov 2026. Not checked yet

WAS THIS USEFUL? 👍 👎

inclusionAI Ling 3.1 Flash

Feedback ⓘ

price increase · Nov 20, 2026 MODEL ↗

Ling 3.1 Flash is a chat model from inclusionAI.

Price increase on OpenRouter (inclusionai/ling-3.1-flash): $0.045 → $0.06 in, $0.135 → $0.18 out per million tokens (+33%), first seen 20 Nov 2026. Not checked yet

WAS THIS USEFUL? 👍 👎

NVIDIA Nemotron

Feedback ⓘ announced · Nov 19, 2026 PRESS ↗

RANK

Nemotron is NVIDIA's family of reasoning, vision, speech, and retrieval models, with training recipes and deployment guides.

> Nemotron shipped a broad model-family launch on Microsoft Foundry, adding new reasoning, retrieval, parsing and safety models, extending Nemotron 3.5 support to Foundry Local, and enabling custom model weight deployment on Foundry Managed Compute.

└──▷ **WHAT SHIPPED · 3 FEATURES** most completely described first ? what's the number?

01 New Nemotron model lineup on Microsoft Foundry

Not checked yet

NEW

63

Nemotron 3.5 Super is now available on Foundry Managed Compute for long-running agents across coding, research, and business workflows, and Nemotron 3.5 Lightning is available on Foundry Managed Compute for applications needing fast responses and tool use. Nemotron Embed and Rerank models are now available to help agents find relevant information across enterprise data, Nemotron Parse is now available to extract structured information from documents, and Nemotron Content Safety is now available to detect harmful content. Nemotron 3.5 Bolt (a high-throughput, low-latency reasoning model) and Nemotron Transcribe (a speech recognition model) are coming soon to Foundry.

— Names eight distinct models and their use cases but no mechanism or exact deployment steps.

launch-20261119-e755702c

THINNER COVERAGE BELOW

02 Nemotron 3.5 support in Foundry Local

Not checked yet NEW

55

Nemotron 3.5 is now supported in Foundry Local for reasoning, coding, and tool-using agents in Azure Local environments, with accompanying retrieval models covering text, document images, charts, and tables.

— Describes scope and modalities but gives no setup or configuration detail.

launch-20261119-e755702c

03 Custom model weight deployment on Foundry Managed Compute

Not checked yet NEW

35

Foundry Managed Compute now supports deploying custom model weights, connecting model customization with deployment.

— Brief statement of capability with no mechanism, limits, or steps given.

launch-20261119-e755702c

WAS THIS USEFUL? 👍 👎

◆ Frontier Models

Particle AI Particle-1 Feedback ⓘ

price increase · Nov 20, 2026 MODEL ↗

Particle-1 is a multimodal model from Particle AI.

Price increase on OpenRouter (particle-ai/particle-1): $0.06 → $0.07 in, $0.50 → $0.50 out per million tokens (+17%), first seen 20 Nov 2026. Not checked yet

WAS THIS USEFUL? 👍 👎

OTHER / UNCATEGORIZED

◆ UNCATEGORIZED

Tokenrip

Feedback ⓘ changes since 2026-10-22 CODE ↗

Tokenrip's command-line client.

Tokenrip: Delete a row by argument with the updated `rip` row delete command.

Create limited API keys from the CLI. Not checked yet

Delete a job with `rip job delete`. Not checked yet

Pass `--json` for machine-readable output or set `TOKENRIP_OUTPUT=json` to default all output to JSON. Not checked yet

Set `TOKENRIP_API_KEY` and `TOKENRIP_API_URL` environment variables to override config-file credentials. Not checked yet

└──▷ **BREAKING ON UPGRADE**

! The `--audience`, `--membership`, and folder `share-contents` options are removed (2.6.24).

! `rip comment` is replaced by `rip message` for workspace threads (2.6.24); scripts calling `rip comment` will break.

! `cancel` replaces `decline` for review round operations (2.6.12).

! `billingUrl` is dropped from `DomainErrorDetails` (2.6.13).

WAS THIS USEFUL? 👍 👎

// END OF TRANSMISSION

The AI Toolchain · curated daily · November 21, 2026
Releases & features from the open-source and commercial AI tools that practitioners actually run.

▷ Subscribe Feedback on this issue